# A MULTIVARIATE APPROACH TO DURATION ANALYSIS

Robert R. Reitano, Ph.D., F.S.A.

## ABSTRACT


Traditionally, the study of the interest rate sensitivity of the price of a portfolio of assets or liabilities has been performed using single variable price functions and a corresponding one variable duration analysis. This unique variable was originally defined as the yield to maturity of the portfolio, and later generalized to reflect "parallel" changes in the underlying yield curve. That is, a change in which each yield point moves by the same amount. Still later, this parallel shift model was generalized to linear shifts, reflecting changes in both the level and slope of the yield curve, as well as to other mathematical models of the manner in which a yield curve is assumed to move.

In general, the ability of such a model to predict price sensitivity is dependent on the validity of this underlying yield curve assumption. For general yield curve shifts, large errors are possible. In practice, this will happen to a greater extent when the portfolio contains both "long" and "short" positions, as is the case for surplus or net worth. A classical duration analysis can greatly understate price sensitivity to nonparallel yield curve shifts in this case. Consequently, surplus changes can appear unpredictable, and duration matching strategies unsuccessful.

In this paper, a general multivariate duration analysis is introduced that does not depend on a mathematical formulation of the way in which a yield curve moves. Consequently, complete price sensitivity information is derived which is equally applicable in all yield curve environments. In addition, this model is practical and relatively easy to apply.

To motivate the multivariate approach, the one variable model is analyzed in theory and through examples, with emphasis on its effectiveness and limitations. Some new results are introduced in this classical setting. The limitations of this model are seen to be overcome by a more general multivariate analysis, and these models are then developed in detail. Examples are utilized throughout to make the theory more accessible. The last section focuses on applications of these models, as well as a variety of practical considerations.

## TABLE OF CONTENTS

# 0. INTRODUCTION

The concept of duration has received a great deal of attention during its relatively short history. Bierwag, Kaufman and Khang [3] and Ingersoll, Skelton and Weil [12] present interesting historic summaries of this activity through 1977, while the newer Bierwag [1] provides additional information on more recent developments. In addition, these sources contain extensive references to the literature, which will be only highlighted here.

The notion of duration was independently discovered by at least four authors. The earliest source is Macaulay [15], who coined the term "duration" in 1938 as a refinement of maturity for quantifying the length of a payment stream, such as a bond. His focus was on better defining the mean time to prepayment. At about the same time, Hicks [10] developed the same duration formula, naming it the "average period," by analyzing the price sensitivity of an income stream to changes in the underlying interest rate. Specifically, the Macaulay duration equals the elasticity of the price of a bond with respect to $v = (1 + i)^{-1}$.

A number of years later, Redington [16] and Samuelson [17] again discovered this formula by analyzing questions in what has come to be known as immunization theory. Redington sought to "immunize" a liability stream with an asset stream, which meant that each was to be equally responsive to changes in the underlying interest rate. This was accomplished by equalizing first derivatives of the associated present value functions, thereby introducing this particular approach to the definition of a duration which has come to be known as "modified duration." Similarly, Samuelson's focus was on immunization, analyzing the sensitivity of a firm's net worth to changes in the underlying interest rate.

For the above one variable formulations, duration was defined in terms of "the interest rate," which typically equalled the yield to maturity. This approach was also followed in Vanderhoof [19], [20] which adapted the Redington model and became what to most actuaries represented "the" introduction to this field of thought. At about the same time, Fischer and Weil [9] generalized the definition of duration to reflect a complete yield curve, rather than the yield to maturity. There, a change in yields was modelled in terms of a parallel shift, whereby each yield rate is changed by the same amount. This duration measure is often referred to as $D_2$, to distinguish it from the Macaulay duration, denoted $D_1$. Corresponding to other models of yield curve dynamics, other duration measures have been defined (see [1], [2], [3], [4], [13] and [14], for example).

More recently, Stock and Simonson [18] have analyzed after-tax adjustments to price sensitivity, while Chambers, Carleton and

McEnally [6] have explored the notion of a duration vector in immunizing bond portfolios. There, the various components of the duration vector correspond intuitively to weighted averages of the adjusted times to maturity raised to various powers. The first component is similar to $D_2$, while the second reflects a measure of the average time squared, then average time cubed, etc.. The adjustment made to the time values is a reduction of one period.

In this paper, a general multivariate approach to duration analysis and price sensitivity is developed which is applicable to virtually any model of yield curve movements. Of course, multivariate models have been used elsewhere (see [1], for example). The purpose here is to explore the general mathematical theory and its application in some detail.

To motivate the approach taken and introduce some new notions in a familiar environment, section 1 focuses on the one variable models in theory and through examples. Here, duration and convexity are defined and used to estimate relative price sensitivity based on the well-known Taylor series approximations. In addition, exponential approximation models are developed based on an identity between price changes and the integral of duration, which is reminiscent of similar identities involving the force of interest, or, the force of mortality.

The various approximation formulas are also compared and characterized in terms of their underlying assumptions concerning the functional form of duration. In addition, the various exponential formulas are seen to be limiting cases of the more traditional formulations. The notion of a "compound duration," or the "duration of duration," is also introduced and the second order approximations are seen to be equivalent to intuitively appealing composites of first order approximations. The examples developed illustrate the effectiveness of these models to approximate relative price changes when price is truly a function of a single yield variable.

Section 2 focuses on the limitations of these models to estimate price changes in the real world, where yields defined from a yield curve and the associated yield changes are truly multivariate. Examples are developed corresponding to the yield to maturity approach, and the parallel yield curve shift approach. In each case, apparently anomolous price behavior is exemplified. In the first example, the units used to define yield changes are seen to have a material effect on price sensitivity conclusions. For the second example, it is shown that for yield curve shifts which are not parallel, the standard formulas can produce estimates which are orders of magnitude in error.

As part of the analysis of this second example, the notions of "directional duration" and "directional convexity" are introduced. Intuitively, these measures reflect price sensitivity to yield curve shifts which in a multivariate sense, are in directions other than the parallel shift vector, $\mathbf{M} = (1, 1, \ldots, 1)$.

Here, each component of the shift vector is interpreted as the change in the corresponding yield curve point.

Section 3 then develops a multivariate duration calculus in detail. Starting with formal definitions of the directional measures noted above, properties are developed which parallel the single variable case of section 1. In particular, polynomial approximations analogous to the traditional formulas are established, as well as exponential approximations based on an exponential identity. Bounds are also determined for the size of directional durations, based on the familiar estimates involving the gradient of a multivariate function.

The concepts of "partial duration" and "partial convexity" are next developed, as well as the corresponding "total duration vector" and "total convexity matrix." Again, polynomial approximations follow, as do exponential approximations based on an exponential identity. These formulations are shown to reduce to the one variable formulas when yield curve shifts are parallel, and this corresponds to the results that duration equals the sum of the partial durations, and similarly, convexity equals the sum of the partial convexities. The examples from section 2 are then revisited and more formally analyzed in the context of these models.

A variety of results are then derived between the partial models and directional models. Not surprisingly, as is true for partial and directional derivatives, the directional duration and convexity values can be readily calculated from the corresponding partial duration and convexity values. Directional duration bounds are revisited in the more natural context of the total duration vector, which is also analyzed in terms of its potential length. Derivatives of the various durations are also derived, as are the associated compound duration concepts. As was the case in section 1, second order multivariate approximations are seen to reduce to natural composites of first order approximations via these compound duration values.

Section 4 then develops some applications in more detail. For noncallable bonds, partial duration and convexity formulas are seen to naturally decompose the classical duration and convexity formulas. For securities which contain options, the standard derivative formulas are inappropriate. Consequently, finite difference formulas are reviewed which are suitable for use with option pricing models. These formulas are formally analyzed with respect to their estimation errors, although in practice, the appropriate difference interval will often be chosen based on trial and error, and judgement.

The price sensitivity implications of the estimated duration values are next explored. The concept of "durational leverage" is introduced and proves to be a useful quantitative measure for understanding the potential price sensitivity compared with that implied by the traditional duration value.

Finally, two yield curve slope models are developed and shown to be easily analyzed with the durational calculus developed in section 3. The first model corresponds to the now relatively common generalization of traditional duration, whereby parallel yield curve shifts are generalized to include affine or linear shifts (Bierwag [2]). That is, where both the level and slope of the yield curve change. The second model is more general, in that the yield curve is reparametrized in terms of its various interpoint slopes.

## a. Duration

Let $P(i)$ denote the price function which assigns to each interest rate $i \geq 0$, the present value of a given collection of future cash flows. The actual rate $i$ can be defined within any system of units: annual, semi-annual, continuous, etc., and will generally follow from the context of the problem. Also, the future cash flows can be positive or negative, fixed or dependent on $i$. However, we will always assume that $P(i)$ is at least twice differentiable, and has a continuous second derivative.

As an example, if $i = .08$ is a semi-annual rate, and future cash flow equals 5 at time 1 year, and 10 at time 5 years, we have:

(1.1) $$P(i) = 5v^2 + 10v^{10},$$
$$P(.08) = 11.38,$$

where $v = (1 + i/2)^{-1}$.

**Definition 1.1** Given a price function $P(i)$, the **(modified) duration function**, $D(i)$, is defined for $P(i) \neq 0$ as follows:

(1.2) $$D(i) = -\frac{dP}{di} \Big/ P(i).$$ ||

For the price function given in (1.1),

(1.3) $$D(i) = (5v^3 + 50v^{11}) \; / \; P(i); \quad D(.08) = 3.25.$$

As defined above, the duration function quantifies the approximate relative change in price caused by a given change in interest rates. This is because the Taylor series approximation:

(1.4) $$P(i) \approx P(i_0) + P'(i_0)(i - i_0),$$

can be rewritten:

(1.5) $$P(i)/P(i_0) \approx 1 - D(i_0)\,\Delta i,$$

where $\Delta i = i - i_0$.

In the above example with $i_0 = .08$ and $i = .085$, the actual relative price change is .9840, which reflects a decrease of 1.60%, while the approximation in (1.5) gives .9838, for a decrease of 1.62%.

Of course, the approximation given by (1.4) is just the traditional tangent line approximation to $P(i)$ at $i_0$. In this light, the duration $D(i_0)$ is seen to be -1 times the slope of the tangent line to $P(i)/P(i_0)$ at $i_0$. Intuitively, $D(i_0)$ approximates the percentage change in price due to a yield change of 100 basis points, or $\Delta i = .01$. For positive $D(i_0)$, price decreases are associated with yield increases, and conversely. For negative $D(i_0)$, price and yield changes move with the same orientation. For the above example, (1.3) therefore implies that the price function in (1.1) will change about 3.25% for a 100 basis point change in rates from $i_0 = .08$. The actual relative change is calculated to be -3.17% for a 100 basis point increase, and 3.32% for a rate decrease.

When the cash flows are fixed and independent of interest rates, another interpretation of duration is possible which relates to the timing of the cash flows. In particular, the duration function in (1.2) is proportional to the weighted average of the times to receipt of the various cash flows. Here, each weight equals the proportion of the total price encompassed by the given cash flow, and the proportionality constant is 1 for continuous i, and $(1 + i/m)^{-1}$ for nominal i compounded m times per year. In the above example with $i_0 = .08$, the weighting on the first cash flow is .41, that on the second is .59, and the weighted average time to receipt is 3.37 years. Scaling by $(1.04)^{-1}$ produces the duration calculated above.

As noted in the Introduction, this "weighted average time" concept is the basis of the original definition of duration, today referred to as the Macaulay Duration, while the definition given in (1.2) is now known as the Modified Duration. The appeal of the original definition is that in terms of average time, no proportionality constant is necessary. In particular, the duration of a single cash flow equals the time to receipt of that cash flow.

The disadvantage of the Macaulay Duration is that to estimate a relative change in price, its value would have to be scaled before applying (1.5).

In addition to the standard approximation given in (1.5), duration can also be used as part of an exponential approximation to P(i). To this end, we have the following:

**Proposition 1** Let P(i) be a price function which is non-zero in an interval I. Then for $i_0$, $i \in I$:

$$(1.6) \qquad P(i)/P(i_0) = \exp\left[-\int_{i_0}^{i} D(y)\, dy\right].$$

**Proof** Because $P(i) \neq 0$, we have that:

$$(1.7) \qquad D(i) = -\frac{d}{di} \ln |P(i)|,$$

for $i \in I$. Integrating (1.7) between $i_0$ and i, and exponentiating the result produces (1.6). ||

Proposition 1 motivates the approximation:

$$(1.8) \qquad P(i)/P(i_0) \approx \exp\left[-D(i_0)\, \Delta i\right],$$

where $\Delta i = i - i_0$. For small values of $\Delta i$, this exponential approximation will produce values which are close to those based on the traditional formula (1.5). This is easily verified by considering the Taylor series expansion of the exponential function in (1.8).

Applying (1.8) to the example in (1.1) yields better approximations than those produced by the traditional approximation (1.5). For example, given a 100 basis point yield change, we would estimate a price change of -3.20% if positive, and 3.30% if negative, based on (1.8). These values compare more favorably to the actual respective values of -3.17% and 3.32%, than the traditional estimate of ±3.25%.

## b. Convexity

The fact that neither approximation (1.5) nor (1.8) tends to produce exact answers suggests that price functions tend to be more complicated than linear or simple exponential models can reflect. More formally, it is virtually always the case that the second derivative of the price function, $P''(i)$, is not identically 0. One exception is given by a simple discount price function with one cash flow, $P(0)$, at time equal to duration, $D(0)$. That is, $P(i) = P(0)(1 - D(0)i)$.

To accomodate the effect of the second derivative of $P(i)$, the concept of convexity is defined analogously to duration, as a relative change function.

**Definition 1.2** Given P(i), the **convexity function**, C(i), is defined for $P(i) \neq 0$ as follows:

(1.9) $$C(i) = \frac{d^2P}{di^2} / P(i). \quad ||$$

For the price function given in (1.1),

(1.10) $$C(i) = (7.5v^4 + 275v^{12}) / P(i); \quad C(.08) = 15.66.$$

Using the second order Taylor series:

(1.11) $$P(i) \approx P(i_0) + P'(i_0)(i - i_0) + ½P''(i_0)(i - i_0)^2,$$

we get the following quadratic generalization of (1.5):

(1.12) $$P(i)/P(i_0) \approx 1 - D(i_0)\Delta i + ½C(i_0)(\Delta i)^2.$$

For example (1.1) with $i_0 = .08$ and $i = .085$, a calculation produces an exact relative price change of .9840 (-1.60%), while the approximation in (1.12) gives .9842 (-1.58%). In this example, the absolute error of the second order approximation is no better than that produced by the first order estimate; both are .02%. For small values of $\Delta i$ in general, the sign of the error associated with a given Taylor series approximation is equal to the sign of the next higher order term. That is, the sign of the product of:
(a) the next derivative evaluated at $i_0$, and (b) $\Delta i$ raised to the corresponding power. Consequently, because convexity is positive in the above example and $(\Delta i)^2$ is always positive, the duration

approximation in (1.5) of .9838 understated the actual relative price change of .9840.

As for the second order approximation using (1.12), the sign of the error depends on both the sign of the third derivative of $P(i)$, and the sign of $\Delta i$, since its exponent will be odd. For the above example the third derivative is negative, so it is predictable that (1.12) will overstate price changes associated with small interest rate increases, and understate these changes for small decreases.

In order to develop a second order counterpart to the exponential approximation in (1.8), we need to first expand the exponent function in (1.6) into a Taylor series. In particular, let:

$$f(i) = \int_{i_0}^{i} D(y)\, dy. \tag{1.13}$$

We then have:

$$f'(i) = D(i), \quad f''(i) = D^2(i) - C(i). \tag{1.14}$$

The second derivative is easily obtained by differentiating the identity, $P' = -DP$, and solving for $D'$. Consequently,

$$f(i) \approx D(i_0)(i - i_0) + \tfrac{1}{2}[D^2(i_0) - C(i_0)](i - i_0)^2. \tag{1.15}$$

Substituting (1.15) into (1.6):

$$P(i)/P(i_0) \approx \exp\{-D(i_0)\Delta i + \tfrac{1}{2}[C(i_0) - D^2(i_0)](\Delta i)^2\}. \tag{1.16}$$

When the approximation in (1.16) is applied to the price function in (1.1), the price change predicted due to an increase in interest rates from .08 to .085 is .9842. This compares to the correct answer of .9840, and equals the quadratic estimate using (1.12).

**Proposition 2** Let $P(i)$ be a price function which is nonzero at $i_0$. Then for $\Delta i$ sufficiently small:

$$\begin{array}{lllll} & \exp(-D(i_0)\Delta i) < P(i)/P(i_0) & & & C > D^2 \\ (1.17) & 1 - D(i_0)\Delta i < P(i)/P(i_0) < \exp(-D(i_0)\Delta i) & & & 0 < C < D^2 \\ & P(i)/P(i_0) < 1 - D(i_0)\Delta i & & & C < 0 \end{array}$$

where $\Delta i = i - i_0$, $D = D(i_0)$, $C = C(i_0)$.

**Proof** Clearly, the bounds in (1.17) correspond to the linear and first order exponential approximations in (1.5) and (1.8). The sign of the error in these first order approximations equals the sign of the second order terms in the respective expansions in (1.12) and (1.16). For the linear approximation, this term has the sign of $C(i_0)$, while for the exponential approximation, this term has the sign of $C(i_0) - D^2(i_0)$. The bounds in (1.17) follow from this and the observation that $1 + x \le e^x$. ||

c. <u>The Duration of Duration</u>

As implied by (1.14), the derivative of the duration function is related to the convexity function. More formally:

(1.18) $$D'(i) = D^2(i) - C(i).$$

Using this expression in the first order Taylor series for D(i):

(1.19) $$D(i) \approx D(i_0) + [D^2(i_0) - C(i_0)]\Delta i.$$

For the example in (1.1), we have D'(.08) = −5.10. Consequently, the approximation in (1.19) predicts that a yield increase of 100 basis points would decrease the duration by about .05. An actual calculation shows that D(.09) = 3.19, for a decrease of .06.

From (1.19), one can conclude that for a small increase in i, D(i) will <u>increase</u> only if $D^2(i_0)$ is larger than $C(i_0)$. Consequently, if convexity is negative at $i_0$, D(i) will always be an increasing function locally about $i_0$. For positive convexity, D(i) will be an increasing function only if $C(i_0) < D^2(i_0)$, and will be a decreasing function, as in the above example, if $C(i_0) > D^2(i_0)$.

By introducing the notion of the duration of duration, the second order approximations in (1.12) and (1.16) can be interpreted naturally as the corresponding first order approximations in (1.5) and (1.8), with an "adjusted duration" value. To this end, we formalize this notion of a compound duration in the natural way:

**Definition 1.3** Given a duration function D(i), the **duration of duration function**, DD(i), is defined for $D(i) \neq 0$ as follows:

(1.20) $$DD(i) = -\frac{dD}{di} / D(i). \quad ||$$

From (1.18), we have that:

(1.21) $$DD(i) = C(i)/D(i) - D(i).$$

Also, the following version of Proposition 1 holds within any interval in which D(i) does not change sign:

(1.22) $$D(i)/D(i_0) = \exp\left[-\int_{i_0}^{i} DD(y)\, dy\right].$$

Rewriting the first order Taylor expansion in (1.19), we get:

(1.23) $$D(i) \approx D(i_0)[1 - DD(i_0)\Delta i],$$

which is functionally equivalent to (1.5). Using this expression in (1.6) and integrating, we get:

(1.24) $$P(i)/P(i_0) \approx \exp\left[-\Delta i\, D(i_0)[1 - DD(i_0)\,\Delta i/2]\right].$$

This approximation for price change is equivalent to the second order exponential obtained in (1.16), as a calculation shows. However, this format is intuitively more appealing to use, since it can be interpreted as an application of the first order exponential approximation in (1.8) with an adjusted duration value. The adjusted duration equals the approximation in (1.23) for $D(i_0 + \Delta i/2)$.

For example, consider the price function in (1.1). A calculation shows that DD(.08) = 1.57. Consequently, if yields increased 100 basis points, the adjusted duration, $D(i_0)[1 - DD(i_0)\Delta i/2]$, will equal the original duration of 3.25, decreased by ½(1.57)% to 3.22. Using this adjusted value in (1.8) is equivalent to applying (1.24) directly, and a price decrease of 3.17% is estimated.

The quadratic approximation in (1.12) can also be rewritten in terms of DD(i) as follows:

(1.25) $$P(i)/P(i_0) \approx [1 - \Delta i\, D(i_0)[1 - (DD(i_0) + D(i_0))\Delta i/2]].$$

Analogous to (1.24), the expression in (1.25) can be interpreted as an application of the standard linear approximation in (1.5) with an adjusted duration. Here, however, the duration adjustment differs from that in (1.24), reflecting both $DD(i_0)$ and $D(i_0)$. Applying (1.25) to the price function in (1.1), DD(.08) + D(.08) = 4.82, so the adjusted duration corresponding to $\Delta i$ = 100 basis points equals the original duration of 3.25, decreased by ½(4.82)% to 3.17. Using this adjusted value in (1.5) produces an estimated price decrease of 3.17%.

d. A Characterization of the Approximations for $P(i)/P(i_0)$

It is interesting to observe that the fundamental difference between the various approximations for $P(i)/P(i_0)$ is the underlying assumption regarding the behavior of $D(i)$ near $i_0$. For the exponential approximations in (1.8), (1.16) and (1.24), this assumption is explicitly based on the identity in (1.6). Namely,

| | Exponential Approximation | | Model for D(i) |
|---|---|---|---|
| (1.26) | (1.8) | 1st Order | $D$ |
| | (1.16), (1.24) | 2nd Order | $D + [D^2 - C]\Delta i$ |

where $D = D(i_0)$ and $C = C(i_0)$.

That is, the first order approximation reflects the assumption that $D(i)$ is constant, while for the second order version, it is assumed that $D(i)$ varies linearly according to its tangent line approximation in (1.19). Hence, if $D(i)$ is constant or linear, the corresponding approximation will be exact.

Turning to the polynomial approximations in (1.5), (1.12) and (1.25); while they may appear more natural than their exponential counterparts, they imply less natural, and sometimes counter-intuitive assumptions about $D(i)$. These assumptions can be determined by equating the exact value of $P(i)/P(i_0)$ as given in (1.6) to the respective approximations, and solving for $D(i)$.

Although integral equations are encountered, these are easily solved by first taking logarithms, then differentiating with respect to i. The following relationships then result:

| | Polynomial Approximation | | Model for D(i) |
|---|---|---|---|
| (1.27) | (1.5) | 1st Order | $D / (1 - D\Delta i)$ |
| | (1.12), (1.25) | 2nd Order | $(D - C\Delta i) / (1 - D\Delta i + \frac{1}{2}C(\Delta i)^2)$ |

The underlying model for D(i) in (1.5) can be counter-intuitive. For example, a calculation shows that this function is an increasing function of $\Delta i$, while as noted above, D(i) is an increasing function locally only when $D^2(i_0)$ exceeds $C(i_0)$. While somewhat more complicated, the model for D(i) underlying (1.12) and (1.25) does not have this potential problem, in that it too will be an increasing function locally only when $D^2(i_0)$ exceeds $C(i_0)$.

## e. Other Relationships

As shown in section 1.d. above, the various approximations for $P(i)/P(i_0)$ can be interpreted in terms of the underlying assumptions regarding the behavior of the duration function, D(i), near $i_0$. In addition, each of the exponential approximations can be shown to be the limiting case of applying the linear approximation in (1.5) to ever finer subdivisions of the interval from $i_0$ to i.

To see this, let $i_0$ and $i > i_0$ be given, and define a subdivision of the corresponding interval by:

(1.28) $$i_j = i_0 + j/n\,(i - i_0), \qquad j = 0,\ldots,n.$$

Clearly, we have that:

(1.29) $$\frac{P(i)}{P(i_0)} = \prod_{j=1}^{n} \frac{P(i_j)}{P(i_{j-1})}.$$

Applying the linear approximation in (1.5) to each term in this product, we get:

(1.30) $$\frac{P(i)}{P(i_0)} \approx \prod_{j=1}^{n} (1 - D(i_{j-1})\Delta i/n).$$

In (1.30), if it is assumed that $D(i_j) = D(i_0)$ for all $j$, the resulting product converges to the first order exponential approximation in (1.8) as $n \to \infty$. If it is assumed that $D(i_j)$ is given by the linear approximation in (1.19), the resulting product converges to the second order exponential approximation in (1.16). If exact $D(i_j)$ is used, the exponential identity in (1.6) results. Interestingly, if the quadratic approximation in (1.12) is used in (1.29), and the convexity function is assumed constant, i.e. $C(i_j) = C(i_0)$, the products again converge to the two exponential approximations depending only on whether we assume $D(i_j)$ to be constant or linear. Similarly, for exact $C(i_{j-1})$ and $D(i_{j-1})$ the product again converges to the identity in (1.6). See the Appendix for a proof of these relationships.

## 2. LIMITATIONS OF THE ONE VARIABLE MODELS

The example developed throughout section 1 illustrates the effectiveness of one variable models to approximate the relative change in price due to a change in the interest rate. Unfortunately, it is difficult to reduce the real world financial markets to such a unique interest rate. In practice, therefore, the use of one variable models is not without its limitations, as the following two examples demonstrate.

### a. An Example - Yield to Maturity Approach

Assume that we have a simple portfolio of three cash flows equal to 20, -20 and 11 at times 0, 1 year and 2 years, respectively. Also, assume that the one year spot rate is .105, and the two year spot rate is .10. For simplicity, such a spot rate curve will be denoted (.105, .10). At these rates, the current price is easily calculated to be 10.99136.

One traditional approach to applying the one variable model is the yield to maturity (YTM) approach whereby the price function P(i) is modelled as follows:

(2.1) $$P(i) = 20 - 20v + 11v^2, \quad v = (1 + i)^{-1}.$$

The equation $P(i) = 10.99136$ has two solutions: .00445 and .21565, and one logical approach to choosing between these values is to check the behavior of $P(i)$ nearby. A simple calculation shows that $P(i)$ is a decreasing function near .00445, and an increasing function near .21565. However, if spot rates increased 100 basis points to (.115, .11), the portfolio value would decrease to 10.99063. Consequently, it is more intuitively appealing to use a decreasing function, so we choose the smaller YTM of .00445. The duration of $P(i)$ at this point is calculated to be .172, and the convexity equals 2.308.

Using the linear approximation for $P(i)$, we get:

(2.2) $$P(i)/P(.00445) \approx 1 - .172(i - .00445).$$

Now, if the yield curve increased uniformly by .01 to (.115, .11), the use of .01445 (i.e. .00445 + .01) for i in (2.2) would yield a very poor approximation. The actual portfolio decrease in this case is .0067%, while this linear approximation and i value would predict a decrease of .17%. Making the adjustment for the convexity value of 2.308 improves the approximation slightly to a predicted decrease of .16%, still orders of magnitude from the correct answer.

Of course, the problem here is one of units; yield curve units versus YTM units. The proper value to use for i in (2.2) is not .01445, but the YTM corresponding to the yield curve (.115,.11). A calculation shows this value to be .00485. That is, the .01 change in yields corresponds to only a .0004 change in YTM, so it is obvious why the above initial approximation was so poor. Using the new YTM in (2.2) produces a predicted decrease of .0069%, and this compares quite favorably to the actual value of .0067%. Here, the convexity adjustment is 0 to four decimal places (in percentage units). Using the exponential approximations provide similar results because the duration and convexity values are relatively small.

It should be noted that if we had chosen the larger YTM value of .21565, its counterintuitive negative duration of -.117 can also be interpreted as a problem of units. That is, an increase in spot yields corresponds to a decrease in YTM, thereby correcting for both the wrong sign and the wrong order of magnitude. Specifically, the yield increase of .01 corresponds to a YTM change of -.0006.

Consequently, one can often correct for the scaling problem inherent with the YTM approach by developing an appropriate conversion formula (see section 2.c). However, the YTM approach also has the uncorrectable problem of nonexistence of solutions. For example, the yield curve (.109,.110) produces a price for the above cash flows of 10.8936, which is below the minimum value in (2.1) of 10.909. Hence, no YTM exists, nor does an estimable $\Delta i$.

b. An Example - Parallel Shift Approach

The commonly used alternative to the YTM approach is the parallel shift approach, whereby the interest rate parameter is defined directly in terms of the change in the yield curve. The restriction here is that the original yield curve of (.105,.10) moves only "in parallel." That is, each yield rate changes by the same amount. Specifically, the price function for the above cash flows is modelled as follows:

(2.3) $$P(i) = 20 - 20v + 11w^2; \quad v = (1.105 + i)^{-1}, \; w = (1.10 + i)^{-1}.$$

The equation $P(i) = 10.99136$ now has the obvious solution of $i = 0$. A calculation produces $D(0) = .0136$, $C(0) = 1.404$, and $P(i)$ is linearly approximated by:

(2.4) $$P(i)/P(0) \approx 1 - .0136i,$$

or, by the corresponding second order estimate which adds $\frac{1}{2}C(0)i^2$. For a parallel yield curve increase of .01 to (.115,.11), the approximation in (2.4) predicts a portfolio decrease of .0136%, which overstates the actual decrease of .0067%. The convexity adjustment improves the approximation from .0136% to .0066%, which is quite good. Using the exponential approximations provides virtually identical results in this case, because D(0) and C(0) are small.

The primary limitation of the parallel shift approach is that yield curve shifts are often not parallel, and the above model can provide poor approximations. Consider, for example, an increase in yields from (.105,.10) to (.1075,.1075). That is, an increase of 25 basis points in the one year spot yield, and 75 basis points in the two year value. Since the duration of the portfolio is positive at .0136, one might expect that an increase in yields should decrease the portfolio value. In this case, this does indeed occur and this nonparallel increase in yields causes a decrease in the portfolio value of .745%.

However, this actual decrease would not have been predicted from the first or second order approximations for $P(i)/P(0)$, *choosing* $i$ to be in the range from 25 to 75 basis points. The best of the four approximations would predict a portfolio decrease of only .010%; a poor estimate for the actual decrease of .745%. It appears that for this nonparallel yield curve change, the portfolio is far more sensitive than $D(0) = .0136$ and $C(0) = 1.404$ imply. This problem has little to do with the order of magnitude of the yield curve shift. That is, the problem is not that shifts of 25 basis points or 75 basis points are too large for the approximation to work well.

For example, assume that the yield curve had increased only slightly from (.105,.10) to (.1052,.1001). This shift is positive and nearly parallel, so given that $D(0) = .0136$, a portfolio decrease is expected. However, the portfolio value actually increases in this case by .015%. Both linear and quadratic approximations predict decreases at both 1 and 2 basis points. The best of these approximations calls for a decrease of .0001%. As before, the sensitivity of the portfolio to this non-parallel shift appears much greater than D(0) and C(0) imply. Unlike before, not even the sign of the sensitivity is accurately predicted.

### c. An Analysis - Yield to Maturity Approach

As the example in section 2.a. shows, the YTM approach can often be used effectively to gauge portfolio sensitivity to parallel yield curve shifts. What is necessary, however, is an appropriate conversion formula to estimate the change in the YTM caused by the given parallel shift in the yield curve.

To this end, let $\mathbf{i}_0$ denote the initial yield curve in common vector notation, and $I_0$ the corresponding YTM, so that $P(\mathbf{i}_0) = P(I_0)$. For the above example, $\mathbf{i}_0 = (.105, .10)$ and $I_0 = .00445$. Also, let $\Delta\mathbf{i}$ denote the parallel shift in the yield curve, and $\Delta I$ the corresponding shift in the YTM, so that $P(\mathbf{i}_0 + \Delta\mathbf{i}) = P(I_0 + \Delta I)$. Expanding each of these functions as first order Taylor series, we get:

(2.5) $$P(\mathbf{i_0} + \Delta i) \approx P(\mathbf{i_0}) + P'(\mathbf{i_0})\,\Delta i,$$

(2.6) $$P(I_0 + \Delta I) \approx P(I_0) + P'(I_0)\,\Delta I.$$

Equating these expressions, and recalling that $P(\mathbf{i_0}) = P(I_0)$, we derive the first order estimate when $D(I_0) \neq 0$:

(2.7) $$\Delta I \approx \frac{D(\mathbf{i_0})}{D(I_0)}\,\Delta i.$$

Note that the proportionality constant in (2.7) is the ratio of $D(\mathbf{i_0})$, the duration of the price function evaluated on the initial yield curve, to $D(I_0)$, the duration evaluated at the initial YTM. For the example in section 2.a. above, this constant is .079. Consequently, a 100 basis point parallel shift corresponds to about an 8 basis point change in YTM. As was noted above, the actual YTM change is about 4 basis points for a .01 parallel increase.

To develop a second order estimate for $\Delta I$, the Taylor series in (2.5) and (2.6) are expanded to include second derivatives. The corresponding quadratic equation in $\Delta I$ is then solved with the quadratic formula, producing:

(2.8) $$\Delta I \approx \left\{D(I_0) - \sqrt{[D^2(I_0) - 2C(I_0)D(\mathbf{i_0})\,\Delta i + C(I_0)C(\mathbf{i_0})(\Delta i)^2]}\right\}/C(I_0).$$

The negative square root is taken in (2.8) to satisfy the initial condition that $\Delta I = 0$ when $\Delta i = 0$.

Applying (2.8) to the example in section 2.a. above, with $i_0 = (.105,.10)$ and $I_0 = .00445$, one calculates that $\Delta I \approx .0004$ for $\Delta i = .01$, a good estimate. Unlike the linear estimate in (2.7), the approximation given by (2.8) is not symmetric in $\Delta i$. This asymmetry is often needed. In the above example, a .01 parallel decrease in $i_0$ corresponds to a .0012 decrease in the YTM. Using (2.8), we estimate that for $\Delta i = -.01$, $\Delta I \approx -.0012$.

As noted above, although the YTM approach can often be used effectively for parallel shifts when the units are properly converted, at least two serious problems persist:

a). Non-existence of YTMs : if there is no exact YTM corresponding to the parallel shifted yield curve $i_0 + \Delta i$, the above conversion formulas for $\Delta I$ may not provide good results. That is, $P(I_0 + \Delta I)$ will not necessarily give a good approximation to $P(i_0 + \Delta i)$.

b). Non-parallel shifts: for yield curve shifts which are not parallel, the above conversion formulas for $\Delta I$ will generally provide unreliable results.

Clearly, the nonexistence problem is unavoidable. However, the problem of non-parallel shifts can be accommodated with more general conversion formulas. These will be developed in section 3.d.

### d. An Analysis - Parallel Shift Approach

We next turn our attention to the example in section 2.b. of the parallel shift approach. As was demonstrated, the sensitivity of the portfolio value to non-parallel shifts, even slightly non-parallel, could be much different from what would have been inferred from the given duration and convexity values.

As was the case for the YTM approach, the problem here is again a problem of units. The various approximation formulas for P(i) reflect the sensitivity of price to parallel shifts of the yield curve of $\Delta i$. This parallel shift of $\Delta i$ is really a vector shift of $\mathbf{\Delta i}$. That is, $\mathbf{\Delta i} \equiv (\Delta i, \Delta i)$ represents a yield change vector which moves the yield curve from $\mathbf{i}$ to $\mathbf{i} + \mathbf{\Delta i} \equiv$ $(i_1 + \Delta i,\ i_2 + \Delta i)$. Looked at this way, the shift vector $\mathbf{\Delta i}$ can be decomposed into a "magnitude," $\Delta i$, and a direction, $\mathbf{N} = (1,1)$:

(2.9) $$\mathbf{\Delta i} = \Delta i(1,1).$$

In addition, the various approximation formulas for $P(\mathbf{i_0} + \Delta i)$ can be interpreted as reflecting the change in price due to a change in yields of $\Delta i$, where this change is in the direction of the vector (1,1).

Decomposing the various shifts in section 2.b., we get:

$$(.01,.01) = .01\ (1,1) \tag{2.10a}$$

$$(.0025,.0075) = .0025\ (1,3) \tag{2.10b}$$

$$(.0002,.0001) = .0001\ (2,1). \tag{2.10c}$$

Of course, other decompositions are also possible. The approximation formulas worked well for shift (2.10a) because the direction of change was (1,1), the direction implicitly assumed in the derivation of these formulas. Non-parallel shifts (2.10b-c) caused poor estimates because their directions did not equal (1,1), and for the cash flows underlying P(i), this difference in directions was very important.

For notational convenience here, let $D_{(1,1)}$ denote the duration as defined in (1.3), with the underlying direction vector of (1,1) explicitly displayed. For the example in section 2.b., we had $D_{(1,1)} = .0136$ evaluated on the initial yield curve, $i_0 = (.105,.10)$. In the next section, duration and convexity will be formally defined with respect to directions other than (1,1). With those definitions, one can calculate:

| | | |
|---|---|---|
| (2.11a) | $D_{(1,1)} = .0136$ | $C_{(1,1)} = 1.404$ |
| (2.11b) | $D_{(1,3)} = 3.0212$ | $C_{(1,3)} = 34.214$ |
| (2.11c) | $D_{(2,1)} = -1.4767$ | $C_{(2,1)} = -6.688$ |

For this example, these duration and convexity values reflect the price sensitivity to yield curve shifts in various directions, and are seen to differ greatly.

Once such <u>directional durations and convexities</u> have been defined and calculated, one can develop the corresponding approximation formulas, such as the counterpart to (1.12):

(2.12) $$P(i_0 + \Delta i N)/P(i_0) \approx 1 - D_N(i_0)\Delta i + \tfrac{1}{2}C_N(i_0)(\Delta i)^2,$$

as well as the analogous first order counterpart to (1.5). Utilizing (2.12) and the directional values in (2.11), the following improved estimates can be obtained:

| | Shift | First Order | Second Order | Exact Value |
|---|---|---|---|---|
| | (.01,.01) | -.0136% | -.0066% | -.0067% |
| (2.13) | (.0025,.0075) | -.7533% | -.7446% | -.7447% |
| | (.0002,.0001) | +.0148% | +.0148% | +.0148% |

In section 3, this multivariate approach to duration and convexity will be explored in detail.

## a. Directional Durations and Convexities

Let $\mathbf{i}_0 = (i_{01}, i_{02}, \ldots, i_{0m})$ represent an m-point yield curve on which the portfolio is valued. Typically, the components of this yield vector would correspond to the yield curve pivotal points. For example, yields for terms: .25, .5, 1, 3, 5, 7, 10, 20, and 30 years. Such pivotal points are truly the external variables on a yield curve since they are observed from market activity. The other yield values are typically interpolated and therefore, internally generated and dependent. Also, let $\mathbf{N} = (n_1, \ldots, n_m)$ be a direction vector, $\mathbf{N} \neq \mathbf{0}$, and $|\mathbf{N}| = (\Sigma n_i^2)^{\frac{1}{2}}$ denote its length.

Consider $f(t) = P(\mathbf{i}_0 + t\mathbf{N})$, where $P(\mathbf{i})$ is a multivariate price function, assumed to be twice continuously differentiable. Clearly, this function defines the price of the portfolio as the initial yield curve $\mathbf{i}_0$ is shifted various units in the direction of $\mathbf{N}$. That is, where $i_{01}$ is shifted $tn_1$ units, $i_{02}$ is shifted $tn_2$ units, etc.. Using a Taylor series expansion, we can approximate $f(t)$ to first and second order in t as follows:

(3.1a) $$f(t) \approx f(0) + f'(0)t,$$

(3.1b) $$f(t) \approx f(0) + f'(0)t + \tfrac{1}{2}f''(0)t^2.$$

In order to calculate the derivatives of $f(t)$ needed in (3.1), it is necessary to recognize that the price function is actually a function of m variables, the shifted yield curve points, and each of these variables is a function of t. Let $P_j(i)$ denote the jth partial derivative of the price function. Similarly, let $P_{jk}(i)$ denote the corresponding mixed second order partial derivative. Then,

(3.2a) $$f'(t) = \Sigma n_j P_j(i_0 + tN),$$

(3.2b) $$f''(t) = \Sigma n_j n_k P_{jk}(i_0 + tN).$$

Evaluated at t=0, the expressions in (3.2) are seen to be the first and second order directional derivatives of the price function P(i). That is;

(3.3a) $$f'(0) \equiv \frac{\partial P}{\partial N}\bigg|_{i_0} = \Sigma n_j P_j(i_0),$$

(3.3b) $$f''(0) \equiv \frac{\partial^2 P}{\partial N^2}\bigg|_{i_0} = \Sigma n_j n_k P_{jk}(i_0).$$

Considering (3.1) and (3.3), the following definitions are motivated:

__Definition 3.1__ Let P(i) be a multivariate price function and $N \neq 0$ a direction vector. The __directional duration function__ in the direction of N, $D_N(i)$, is defined for $P(i) \neq 0$ as follows:

(3.4) $$D_N(i) = -\frac{\partial P}{\partial N} / P(i). \quad ||$$

**Definition 3.2** Given the assumptions of Definition 3.1, the **directional convexity function** in the direction of N, $C_N(i)$, is defined for $P(i) \neq 0$ as follows:

(3.5) $$C_N(i) = \frac{\partial^2 P}{\partial N^2} / P(i). \quad ||$$

Substituting (3.3) into (3.1), the following counterparts to (1.5) and (1.12) are produced, as noted in (2.12):

(3.6) $$P(i_0 + \Delta i N)/P(i_0) \approx 1 - D_N(i_0)\Delta i,$$

(3.7) $$P(i_0 + \Delta i N)/P(i_0) \approx 1 - D_N(i_0)\Delta i + \tfrac{1}{2}C_N(i_0)(\Delta i)^2.$$

As an example, consider the price function in (2.3) explicitly expressed as a function of two variables:

(3.8) $$P(i_1, i_2) = 20 - 20v + 11w^2;$$

where $v = (1 + i_1)^{-1}$, $w = (1 + i_2)^{-1}$. The various partial derivatives of $P(i_1, i_2)$ are easily calculated to be:

(3.8a) $$P_1(i_1, i_2) = 20v^2; \qquad P_2(i_1, i_2) = -22w^3$$

(3.8b) $$P_{11}(i_1, i_2) = -40v^3; \quad P_{22}(i_1, i_2) = 66w^4; \quad P_{12} = P_{21} \equiv 0.$$

Evaluating these derivatives at $i_0 = (.105, .10)$, and performing the necessary weighted summations in (3.3), the directional durations and convexities displayed in (2.11) can be readily verified.

Before continuing, it is worth noting that:

1) If $\mathbf{N} = (1,\ldots,1)$, the parallel shift direction vector, $D_N(\mathbf{i_0})$ equals the traditional value of $D(0)$, and $C_N(\mathbf{i_0}) = C(0)$. Here, these traditional values are calculated utilizing the parallel shift approach (see Proposition 6, below).

2) Formulas (3.6) and (3.7) are consistent even though there are infinitely many ways to specify the direction vector $\mathbf{N}$. For example, given $\mathbf{N}$, let $\mathbf{N'} = \frac{1}{2}\mathbf{N}$. The corresponding shift magnitudes satisfy: $\Delta i' = 2\Delta i$. The estimates in (3.6) and (3.7) will then be the same for $\mathbf{N}$ and $\mathbf{N'}$, because $D_{N'} = 1/2\ D_N$, and $C_{N'} = 1/4\ C_N$ by (3.3).

To make this more well-defined, it is possible to normalize the model by requiring the direction vector $\mathbf{N}$ to satisfy $|\mathbf{N}| = 1$. The magnitude variable, $\Delta i$, is then uniquely defined as the length of the shift vector $\Delta i\mathbf{N}$. However, whether $\mathbf{N}$ is normalized or not, consistent estimates are produced.

<u>**Proposition 3**</u> Let $P(\mathbf{i})$ be a multivariate price function and $\mathbf{N}$ a direction vector with $P(\mathbf{i_0} + \Delta i\mathbf{N}) \neq 0$ for $|\Delta i| \leq K$. Then,

$$P(\mathbf{i_0} + \Delta i\mathbf{N})/P(\mathbf{i_0}) = \exp\left[-\int_0^{\Delta i} D_N(\mathbf{i_0} + t\mathbf{N})\,dt\right], \tag{3.9}$$

for $|\Delta i| \leq K$.

**Proof** Define $f(t) = \ln|P(i_0 + tN)|$. Then $-f'(t) = D_N(i_0 + tN)$, which can be integrated and exponentiated to produce (3.9). ||

From (3.9), the following first order exponential approximation is transparent:

$$P(i_0 + \Delta iN)/P(i_0) \approx \exp(-D_N(i_0)\,\Delta i). \tag{3.10}$$

As was true in section 1, for small values of $D_N(i_0)$ this exponential approximation will yield results which are close to those produced by the more traditional-looking approximation (3.6). In order to develop the second order exponential formula, we must expand the exponent function in (3.9) as a Taylor series in $\Delta i$. To do this, the directional derivative of $D_N$ at $i_0$ is needed. Analogous to (1.18), we have:

$$\frac{\partial D_N}{\partial N} = D_N^2(i) - C_N(i). \tag{3.11}$$

This formula is readily verified by taking directional derivatives of the identity, $\frac{\partial P}{\partial N} = -D_N P$.

Proceeding as in the derivation of (1.16), we obtain:

$$P(i_0 + \Delta iN)/P(i_0) \approx \exp[-D_N(i_0)\,\Delta i + ½(C_N(i_0) - D_N^2(i_0))(\Delta i)^2]. \tag{3.12}$$

b. **Bounds for Directional Durations**

Given a price function, $P(\mathbf{i})$, and a yield curve vector $\mathbf{i}_0$, it is natural to inquire as to the existence of direction vectors which either minimize or maximize $D_N(\mathbf{i}_0)$. In light of (3.6), such direction vectors will represent critical yield curve shift directions for $P(\mathbf{i})$. As noted in section 3.a. above, this question will not be well posed unless some restriction is put on the length of $\mathbf{N}$. This is because if $\mathbf{N}' = \alpha\mathbf{N}$, $D_{N'}(\mathbf{i}_0) = \alpha D_N(\mathbf{i}_0)$. Consequently, we can always increase a positive $D_N(\mathbf{i}_0)$ by increasing the length of $\mathbf{N}$. Restricting our attention to normalized direction vectors $\mathbf{N}$ satisfying $|\mathbf{N}| = 1$, we have the following (see also Proposition 10):

**Proposition 4** Let $P(\mathbf{i})$ be a price function, $\mathbf{i}_0$ a yield vector with $P(\mathbf{i}_0) > 0$, and $\mathbf{N}_0 = -(P_1(\mathbf{i}_0),\ldots,P_m(\mathbf{i}_0))/|\mathbf{P}'(\mathbf{i}_0)|$, where $|\mathbf{P}'(\mathbf{i}_0)|^2 = \Sigma P_j^2(\mathbf{i}_0)$ is assumed to be non-zero. Then for all direction vectors $\mathbf{N}$ satisfying $|\mathbf{N}| = 1$, we have:

(3.13) $$-\frac{|\mathbf{P}'(\mathbf{i}_0)|}{P(\mathbf{i}_0)} \le D_N(\mathbf{i}_0) \le \frac{|\mathbf{P}'(\mathbf{i}_0)|}{P(\mathbf{i}_0)}.$$

Further, the limits in (3.13) are attained for $\mathbf{N} = \pm\mathbf{N}_0$, with the upper limit corresponding to $\mathbf{N}_0$, and conversely.

If $P(\mathbf{i}_0) < 0$, the inequalities in (3.13) are reversed, and the upper limit is attained at $-\mathbf{N}_0$.

**Proof** This proposition is nothing more than a restatement of the classic result regarding a directional derivative; that it is maximized in the direction of its gradient, and minimized in the opposite direction. Here of course, $N_0$ is -1 times the normalized gradient of $P(i)$ at $i_0$. ||

**Proposition 5** Let $P(i)$ and $i_0$ be given and assume that $|P'(i_0)| = 0$. Then for all $N$,

(3.14) $$D_N(i_0) = 0.$$

**Proof** This result is clear from the definition of $D_N(i_0)$ and (3.3a), since $|P'(i_0)| = 0$ if and only if $P_j(i_0) = 0$ for all $j$. ||

Returning to the example of (3.8), one readily calculates from (3.8a) that $|P'(i_0)| = 23.27$ and $N_0 = (-.704, .710)$. Evaluating the critical values of $D_N(i_0)$ by (3.13), we get:

(3.15) $$-2.12 \leq D_N(i_0) \leq 2.12, \quad |N| = 1.$$

Finally, a calculation shows that $D_N(i_0) = \pm 2.12$ at $\pm N_0$, respectively.

As a final comment regarding Proposition 4, it should be noted that for $|N| \neq 1$, the bounds in (3.13) are readily generalized. For example, for $P(i_0) > 0$,

(3.13)' $$-\frac{|P'(i_0)|}{P(i_0)} |N| \leq D_N(i_0) \leq \frac{|P'(i_0)|}{P(i_0)} |N|.$$

### c. Partial Durations and Convexities

As shown in section 3.b., the classical duration and convexity analysis of section 1 can be readily generalized to include yield curve shifts which are not parallel. An alternative model would be one which more explicitly recognizes the multivariate nature of yield curve changes. That is, a model which estimates $P(\mathbf{i_0} + \Delta\mathbf{i})$ directly, where $\mathbf{i_0}$ is the initial yield curve vector, and $\Delta\mathbf{i} = (\Delta i_1, \ldots, \Delta i_m)$ is a yield change vector.

To this end, consider the following m-dimensional versions of the first and second order Taylor series:

(3.16a) $$P(\mathbf{i_0} + \Delta\mathbf{i}) \approx P(\mathbf{i_0}) + \Sigma P_j(\mathbf{i_0}) \Delta i_j,$$

(3.16b) $$P(\mathbf{i_0} + \Delta\mathbf{i}) \approx P(\mathbf{i_0}) + \Sigma P_j(\mathbf{i_0}) \Delta i_j + \tfrac{1}{2}\Sigma\Sigma P_{jk}(\mathbf{i_0}) \Delta i_j \Delta i_k.$$

These approximations naturally motivate the following definitions:

**Definition 3.3** Given a multivariate price function $P(\mathbf{i})$, the **jth partial duration function**, denoted $D_j(\mathbf{i})$, is defined for $P(\mathbf{i}) \neq 0$ as follows:

(3.17) $$D_j(\mathbf{i}) = -P_j(\mathbf{i})/P(\mathbf{i}), \quad j = 1, \ldots, m. \quad ||$$

**Definition 3.4** Given the price function $P(i)$, the **jkth partial convexity function**, denoted $C_{jk}(i)$, is defined for $P(i) \neq 0$ as follows:

$$C_{jk}(i) = P_{jk}(i)/P(i), \quad j,k = 1,\ldots,m. \quad || \tag{3.18}$$

**Definition 3.5** Given the above definitions, the **total duration vector**, denoted $\mathbf{D}(i)$, and the **total convexity matrix**, denoted $\mathbf{C}(i)$, are defined as follows:

$$\mathbf{D}(i) = (D_1(i),\ldots,D_m(i)), \tag{3.19}$$

$$\mathbf{C}(i) = \begin{vmatrix} C_{11}(i) & \ldots & C_{1m}(i) \\ \vdots & & \vdots \\ C_{m1}(i) & \ldots & C_{mm}(i) \end{vmatrix}. \quad || \tag{3.20}$$

Note that $\mathbf{D}(i)$ is to be interpreted as a row vector. Utilizing these definitions in (3.16), the following generalizations of (1.5) and (1.12) are produced:

$$P(i_0 + i)/P(i_0) \approx 1 - \mathbf{D}(i_0)\cdot\Delta i \tag{3.21}$$

$$P(i_0 + i)/P(i_0) \approx 1 - \mathbf{D}(i_0)\cdot\Delta i + \tfrac{1}{2}\Delta i^T \mathbf{C}(i_0)\Delta i. \tag{3.22}$$

To simplify notation, (3.21) utilizes the well known *dot product* or *inner product* notation, whereby if $x$ and $y$ are m-vectors, $x \cdot y$ is defined:

(3.23) $$x \cdot y = \Sigma x_j y_j.$$

Similarly, the last term in (3.22) is expressed in <u>matrix product notation</u>, or more specifically, as a <u>quadratic form</u> in $\Delta i$. By convention, $\Delta i$ is interpreted as a column vector, and $\Delta i^T$ is the corresponding row vector, or <u>transpose</u> of $\Delta i$. Standard matrix calculations then produce:

(3.24) $$x^T C x = \Sigma\Sigma c_{jk} x_j x_k.$$

It should be noted that for smooth price functions,

(3.25) $$C_{jk}(i) = C_{kj}(i),$$

because of the corresponding property for mixed partial derivatives. Consequently, $C(i)$ is a symmetric matrix in this case. That is,

(3.26) $$C(i) = C(i)^T.$$

It should also be noted that the dot product in (3.23) can also be expressed in matrix notation as $x^T y$ ($x, y$ column matrices), or $xy^T$ ($x, y$ row matrices).

Again returning to the example in (2.3), where $P(i_1, i_2) = 20 - 20v + 11w^2$, and $i_0 = (.105, .10)$, the partial derivatives in (3.8) imply:

$$(3.27a) \qquad D_1(i_0) = -1.4902, \quad D_2(i_0) = 1.5038,$$

$$(3.27b) \qquad C_{11}(i_0) = -2.697, \quad C_{22}(i_0) = 4.101, \quad C_{12} = C_{21} = 0.$$

Hence, the first order approximation in (3.21) becomes:

$$(3.28) \qquad P(i_0 + \Delta i) \approx 10.99136(1 + 1.4902\,\Delta i_1 - 1.5038\,\Delta i_2).$$

Looking at the functional form of (3.28), it is little wonder that for nonparallel yield curve shifts, $\Delta i_1 \neq \Delta i_2$, the price function changed in ways not anticipated by the traditional approximation (2.4). Namely, this price function is relatively sensitive to movements in $\Delta i_1$ and $\Delta i_2$ separately. However, because these sensitivities are of opposite sign and similar magnitude, the traditional approximation, which assumes $\Delta i_1 = \Delta i_2$, produces an apparent sensitivity of only .0136.

Similarly, the traditional convexity value of 1.404 disguises the greater sensitivities implied by the partial convexities in (3.27b). That is, expanding (3.28) to second order terms as in (3.22), we get:

$$(3.29) \qquad P(i_0 + \Delta i) \approx 10.99136\,[1 + 1.4902\,\Delta i_1 - ½(2.697)(\Delta i_1)^2 - 1.5038\,\Delta i_2 + ½(4.101)(\Delta i_2)^2].$$

Again, depending on the relationship between $\Delta i_1$ and $\Delta i_2$, this price function will behave in ways not anticipated by the traditional approximation which assumes $\Delta i_1 = \Delta i_2$.

Implicit in the above discussion is the assumption that when a multivariate approximation is restricted to parallel shifts, i.e. $\Delta i_j = \Delta i$ for all $j$, the corresponding one variable approximation from section 1 is produced. For example, (3.21) reduces to (1.5). For this to be so, it is necessary and sufficient that duration equals the sum of the partial durations, and convexity equals the sum of the partial convexities.

The following proposition formalizes this result:

__Proposition 6__ Let $\mathbf{i_0}$ be a yield curve vector and $D(\mathbf{i_0})$ and $C(\mathbf{i_0})$ denote the duration and convexity values calculated according to the "parallel shift" approach. Then:

(3.30) $$D(\mathbf{i_0}) = \Sigma D_j(\mathbf{i_0}),$$

(3.31) $$C(\mathbf{i_0}) = \Sigma\Sigma C_{jk}(\mathbf{i_0}).$$

__Proof__ Let $\mathbf{M} = (1,\ldots,1)$, the parallel shift direction vector and define the price function $P(i) = P(\mathbf{i_0} + i\mathbf{M})$. The chain rule then gives:

$$(3.32a) \qquad P'(i) = \Sigma P_j(i_0 + iM)$$

$$(3.32b) \qquad P''(i) = \Sigma\Sigma P_{jk}(i_0 + iM).$$

Evaluating (3.32) at $i = 0$, and dividing by $P(0) = P(i_0)$, completes the proof. !!

Turning next to the exponential models, we have the following:

<u>Proposition 7</u> Let $\Gamma(t)$ be a smooth parametrization of yield curve vectors defined on $[0,1]$ so that $\Gamma(0) = i_0$, $\Gamma(1) = i_0 + \Delta i$. Also, assume that $P(\Gamma(t)) \neq 0$ for $0 \leq t \leq 1$. Then:

$$(3.33) \qquad P(i_0 + \Delta i)/P(i_0) = \exp[-\int_0^1 D(\Gamma(t))\cdot\Gamma'(t)\,dt],$$

where $\Gamma'(t)$ denotes the ordinary derivative of this vector valued function.

<u>Proof</u> Define $f(t) = \ln|P(\Gamma(t))|$. A calculation shows that $f'(t) = -D(\Gamma(t))\cdot\Gamma'(t)$, which can be integrated and exponentiated to complete the proof. !!

From Proposition 7, the following approximation results:

(3.34) $$P(i_0 + \Delta i)/P(i_0) \approx \exp[-D(i_0)\cdot\Gamma'(0)].$$

In the special case where $\Gamma(t)$ is linear, $\Gamma(t) = i_0 + t\Delta i$, the more general formulas in (3.33) and (3.34) are easily seen to reduce to the directional derivative counterparts in (3.9) and (3.10), with $\Delta i$ here corresponding to $\Delta iN$ above.

In order to develop the second order exponential approximation, partial derivatives of the various partial durations are required. Analogous to (1.18) and (3.11), we have:

(3.35) $$\frac{\partial D_j}{\partial i_k} = D_k D_j - C_{jk},$$

which is derived by differentiating the identity $P_j = -PD_j$, with respect to $i_k$. Proceeding as before, one can expand the exponent function in (3.33) as a one variable Taylor series by replacing the upper limit of integration with s, say, then substituting $s = 1$ into the second order Taylor expansion to obtain:

(3.36) $$P(i_0 + i)/P(i_0) \approx \exp\{-D(i_0)\cdot\Gamma'(0) + \tfrac{1}{2}[\Gamma'(0)^T(C(i_0) - D(i_0)^T D(i_0))\Gamma'(0) - D(i_0)\cdot\Gamma''(0)]\}.$$

In the special case where $\Gamma(t)$ is linear, $\Gamma''(0) \equiv 0$, and (3.36) reduces to the directional derivative counterpart in (3.12).

d. <u>YTM Approach Revisited</u>

In section 2.c, approximation formulas were developed in (2.7) and (2.8) which illustrated the sensitivity of the yield to maturity to parallel shifts in the yield curve. In this section, these results will be generalized to include non-parallel shifts.

As before, let $\mathbf{i_0}$ be a yield curve vector, and $I_0$ the equivalent YTM so that $P(\mathbf{i_0}) = P(I_0)$. Expanding into the respective first order Taylor series,

(3.37a) $$P(\mathbf{i_0} + \Delta\mathbf{i}) \approx P(\mathbf{i_0})[1 - \mathbf{D}(\mathbf{i_0}) \cdot \Delta\mathbf{i}],$$

(3.37b) $$P(I_0 + \Delta I) \approx P(I_0)[1 - D(I_0)\Delta I].$$

Equating these values, we can solve for $\Delta I$ when $D(I_0) \neq 0$, obtaining:

(3.38) $$\Delta I \approx \frac{\mathbf{D}(\mathbf{i_0}) \cdot \Delta\mathbf{i}}{D(I_0)}.$$

This equation reduces to (2.7) when $\Delta\mathbf{i}$ is a parallel shift, since $D(\mathbf{i_0}) = \Sigma D_j(\mathbf{i_0})$.

As an example, recall the price function of section 2.a., where the initial yield curve, $\mathbf{i_0} = (.105,.10)$, was seen to be equivalent to the yield to maturity, $I_0 = .00445$. That is, both yielded an initial price of 10.99136. Consider the small nonparallel yield curve shift, $\mathbf{\Delta i} = (.0005, .001)$. Based on (3.38), one approximates the associated change in the yield to maturity, $\Delta I \approx .00442$, using the duration values from (2.2) and (3.27). Estimating $\Delta I$ directly proves this result to be a little understated, in that $\Delta I = .00455$.

Consider next the larger nonparallel shift of $\mathbf{\Delta i} = (.005, .01)$. Because this shift flattens the original yield curve to $(.11, .11)$, it is obvious that the new corresponding YTM equals $.11$, and that we should find that $\Delta I = .10555$. The approximation based on (3.38) equals $.0442$, an apparently significant error. However, it must be kept in mind that the approximation produced by (3.38) for $\Delta I$, used in conjunction with $D(I_O)$ in (3.37b), will produce the same estimate for $P(.11, .11)$ as will (3.37a) using the actual $\mathbf{\Delta i}$ and the partial durations.

By expanding the Taylor series in (3.37) to include second order terms, $\Delta I$ can be estimated using the quadratic formula, producing the following generalization of (2.8):

$$(3.39) \qquad \Delta I \approx \left\{ D_O - \sqrt{[D_O^2 - 2C_O \mathbf{D} \cdot \mathbf{\Delta i} + C_O \mathbf{\Delta i}^T \mathbf{C} \mathbf{\Delta i}]} \right\} / C_O,$$

where $D_O = D(I_O)$, $C_O = C(I_O)$, $\mathbf{D} = \mathbf{D}(\mathbf{i_O})$, and $\mathbf{C} = \mathbf{C}(\mathbf{i_O})$.

This formula generalizes (2.8) to allow nonparallel yield curve shifts, and as was the case there, the negative square root is used to satisfy the initial condition that $\Delta I = 0$ when $\mathbf{\Delta i} = \mathbf{0}$.

Recalling the partial duration and convexity values in (3.27), this quadratic formula can be used to estimate the $\Delta I$ associated with $\boldsymbol{\Delta i}$ = (.0005,.001) in the example above. In this case, the estimate for $\Delta I$ is improved compared with the linear estimate, reproducing the exact value of $\Delta I$ = .00455 to five decimal places. For the larger shift of $\boldsymbol{\Delta i}$ = (.005,.01), a negative value is produced under the square root. That is, there is no real number, $\Delta I$, for which the one variable second order Taylor series equals the multivariable series which reflects $\boldsymbol{\Delta i}$, $D(\mathbf{i})$, and $C(\mathbf{i})$. A calculation shows that this latter value is .99258, while the minimum value of the one variable quadratic is .99362, which is achieved at $\Delta I$ = .07435.

In this case, although an improved estimate for $\Delta I$ can be obtained by this critical value analysis, its use in the associated second order Taylor series does not produce a good estimate for the change in price. Specifically, this second order analysis would produce a relative change of .99362, while the first order analysis with $\Delta I$ = .04422 produces a relative price change of .99241, which is significantly closer to the actual value of .99258.

## e. Parallel Shift Approach Revisited

Considering next the parallel shift analysis of section 2.d, recall that it was shown that non-parallel shifts could be handled by redefining duration and convexity to reflect these non-parallel directions. Alternatively, non-parallel shifts can be accommodated using the standard section 1 formulas, if the parallel shift parameter, $\Delta i$, is properly constructed as a function of the actual shift, $\boldsymbol{\Delta i}$.

To this end, the first order expansion of $P(i_0 + \boldsymbol{\Delta i})$ in (3.37a) must be used twice, once for the general $\boldsymbol{\Delta i}$, and once for the parallel shift vector, $\boldsymbol{\Delta i} = \Delta i \mathbf{M}$, where $\mathbf{M} = (1,...,1)$. Equating these approximations, we can solve for $\Delta i$ when $D(i_0) \neq 0$, obtaining:

(3.40) $$\Delta i \approx \frac{\mathbf{D}(i_0)\cdot\boldsymbol{\Delta i}}{D(i_0)}.$$

Unlike the YTM counterpart formula in (3.38), here $\Delta i$ is seen to be a weighted average of the various component $\Delta i_j$ values since $\Sigma D_j(i_0) = D(i_0)$.

Using the partial durations in (3.27a), we can apply (3.40) to the non-parallel shifts in (2.10), to obtain:

| $\boldsymbol{\Delta i}$ | "Equivalent" $\Delta i$ |
|---|---|
| (.0025,.0075) | .5554 |
| (.0002,.0001) | -.0109 |

A calculation shows that using these parallel shift equivalents in the standard first order formula (2.4) produces identical first order results to those displayed in (2.13) produced with directional derivatives.

Interpreted this way, we see that the traditional formulas can provide poor estimates for non-parallel shifts because the units of the associated parallel shift $\Delta i$, can be orders of magnitude larger, and/or of a different sign, than may be inferred from the various non-parallel shift values of $\Delta i_j$. This cannot happen if all $D_j(i_0)$ values have the same sign, for example, as is true for a noncallable bond (see (4.2)). In such cases, the equivalent $\Delta i$ will be within the range of $\Delta i_j$ values, as is easily seen.

The second order counterpart to (3.40) is identical to (3.39), only with $D_0 = D(i_0)$ and $C_0 = C(i_0)$.

### f. <u>Duration and Convexity Relationships</u>

Relationships between the various duration and convexity values defined in the previous sections are developed in the following propositions:

<u>Proposition 8</u> Let $N \neq 0$ be a direction vector. Then:

(3.41) $$D_N(i_0) = N \cdot D(i_0) = \Sigma n_j D_j(i_0),$$

(3.42) $$C_N(i_0) = N^T C(i_0) N = \Sigma\Sigma n_j n_k C_{jk}(i_0).$$

**Proof** Both formulas are restatements of the definitions of $D_N(i_0)$ and $C_N(i_0)$, reflecting the directional derivative identities in (3.3). ||

Before continuing, it should be noted that for $M = (1,\ldots,1)$, we have by Propositions 6 and 8, the expected results:

(3.43) $$D_M(i_0) = D(i_0),$$

(3.44) $$C_M(i_0) = C(i_0).$$

The following proposition summarizes a number of results regarding derivatives of the various duration functions.

**Proposition 9** Let $N \neq 0$ be a direction vector. Then:

(3.45) $$\frac{d}{di} D(i_0) = D^2(i_0) - C(i_0),$$

(3.46) $$\frac{\partial}{\partial N} D_N(i_0) = D^2{}_N(i_0) - C_N(i_0),$$

(3.47) $$\frac{\partial}{\partial i_j} D_k(i_0) = D_j(i_0)D_k(i_0) - C_{jk}(i_0),$$

(3.48) $$\frac{\partial}{\partial i_j} D(i_0) = D(i_0)D_j(i_0) - \sum_k C_{jk}(i_0).$$

**Proof** Let $P(i) = P(i_0 + iM)$. Relationship (3.45) is derived by differentiating the identity, $P'(i) = -P(i)D(i)$, solving for $D'(i)$, and substituting $i = 0$. Similarly, (3.46) is derived from the identity, $P_N(\mathbf{i}) = -P(\mathbf{i})D_N(\mathbf{i})$, where $P_N(\mathbf{i})$ denotes the directional derivative of $P(\mathbf{i})$. Here, however, it is the directional derivatives which are taken.

Similarly, differentiating the identity, $P_k(\mathbf{i}) = -P(\mathbf{i})D_k(\mathbf{i})$ with respect to $i_j$ leads to (3.47), while summing this result with respect to k and using (3.30) produces (3.48). ||

Returning now to bounds for directional derivatives, we have:

Proposition 10 Let $P(\mathbf{i})$ be a price function and $\mathbf{D}(\mathbf{i}_0)$ its total duration vector evaluated at $\mathbf{i}_0$. Then for all duration vectors, $\mathbf{N}$,

(3.49) $$-|\mathbf{D}(\mathbf{i}_0)||\mathbf{N}| \leq D_N(\mathbf{i}_0) \leq |\mathbf{D}(\mathbf{i}_0)||\mathbf{N}|,$$

where | | denotes the length of the given vectors. Further, the upper bound in (3.49) is achieved for all positive multiples of the unit vector:

(3.50) $$\mathbf{N}_0 = \mathbf{D}(\mathbf{i}_0)/|\mathbf{D}(\mathbf{i}_0)|.$$

Similarly, the lower bound is achieved for all positive multiples of $-\mathbf{N}_0$.

Proof By multiplying the numerator and denominator of $\mathbf{N}_0$ in (3.50) by $P(\mathbf{i}_0)$, it becomes clear that this unit vector equals $\mathbf{N}_0$ of Proposition 4. By evaluating $D_N(\mathbf{i}_0)$ for $\mathbf{N} = \pm\mathbf{N}_0$ by (3.41), the bounds in (3.49) are seen to be a simplified restatement of (3.13) and (3.13)', since the sign of $P(\mathbf{i_0})$ becomes transparent. ||

It should be noted that by Proposition 10, if $D_j(\mathbf{i_0}) = D(\mathbf{i_0})/m$ for all $j$, the corresponding price function is most sensitive to parallel yield curve shifts since then $\mathbf{N_0} = (1, 1, \ldots, 1)$. Next, Proposition 11 shows that given $D(\mathbf{i_0})$, the range of price sensitivity displayed in (3.49) is minimized for this case.

__Proposition 11__ Let $\mathbf{D}(\mathbf{i_0})$ be a total duration vector with associated duration $D(\mathbf{i_0})$. Then:

(3.51) $$|\mathbf{D}(\mathbf{i_0})| \geq |D(\mathbf{i_0})|/\sqrt{m},$$

where $m$ is the dimension of $\mathbf{D}(\mathbf{i_0})$. Further, the lower bound in (3.51) is achieved if and only if $D_j(\mathbf{i_0}) = D(\mathbf{i_0})/m$, for all $j$.

__Proof__ Although this is a familiar calculus result, a simple non-calculus proof is possible. Changing notation, let $A$ be the vector with $a_j = D(\mathbf{i_0})/m$, for all $j$, and let $B$ also have the property that $\Sigma b_j = D(\mathbf{i_0})$. Then $C = B - A$ satisfies $\Sigma c_i = 0$, so $|B|^2 = |A|^2 + |C|^2$. Hence, since $|C|^2 \geq 0$, $|B|^2$ is minimized when $C = 0$. ||

__Proposition 12__ Let $P_1(\mathbf{i})$ and $P_2(\mathbf{i})$ be price functions with corresponding total duration vectors $\mathbf{D_1}(\mathbf{i})$, $\mathbf{D_2}(\mathbf{i})$, and total convexity matrices $\mathbf{C_1}(\mathbf{i})$ and $\mathbf{C_2}(\mathbf{i})$. Let $P(\mathbf{i}) = P_1(\mathbf{i}) + P_2(\mathbf{i})$. Then for $P(\mathbf{i}) \neq 0$,

$$D(i) = [P_1(i)D_1(i) + P_2(i)D_2(i)]/P(i), \tag{3.52}$$

$$C(i) = [P_1(i)C_1(i) + P_2(i)C_2(i)]/P(i). \tag{3.53}$$

<u>Proof</u> As is the case for the traditional values, this result follows directly from the additive property of derivatives. ||

Clearly, Propostion 12 implies that both partial values and directional values satisfy similar identities.

As a final comment, it should be noted that the conclusions noted in section 1.e. for the one variable models hold in the multivariate context as well. For example, the directional duration exponential approximations can be interpreted as the limiting case of applying the directional linear approximations to ever finer subdivisions of the segment $[i_0, i_0 + \Delta iN]$. The assumption of a constant directional duration then leads to the first order exponential formulas, while the assumption that this function is linear over the segment leads to the second order formulas. As before, use of the second order directional approximations with a constant directional convexity does not change this result. In addition, the exponential identity can be viewed as the limiting case of the corresponding first order approximations with exact directional duration values.

For the partial duration models restricted to $\Gamma(t) = i_0 + t\,\Delta i$, similar results hold. For general $\Gamma(t)$, the linear approximation converges to the exponential identity as can be shown by defining the partition $\{j/n \mid j = 0,\ldots,n\}$ on $[0,1]$, the domain of $\Gamma(t)$, and proceeding as before.

## g. Compound Duration Functions

In section 1.c., the concept of the duration of duration was defined and used to restate the second order approximations in an intuitively natural way. Here, this compound duration approach will be generalized to the multivariate models.

**Definition 3.6** Given a directional duration function $D_N(i)$, the **compound directional duration**, $D_N D_N(i)$, is defined for $D_N(i) \neq 0$ as follows:

$$D_N D_N(i) = -\frac{\partial D_N}{\partial N} \Big/ D_N(i). \quad !! \tag{3.54}$$

**Definition 3.7** Given a partial duration function, $D_k(i)$, the **compound jkth partial duration**, $D_j D_k(i)$, is defined for $D_k(i) \neq 0$ as follows:

$$D_j D_k(i) = -\frac{\partial D_k}{\partial i_j} \Big/ D_k(i). \quad !! \tag{3.55}$$

From Proposition 9:

(3.56) $$D_N D_N(i) = C_N(i)/D_N(i) - D_N(i),$$

(3.57) $$D_J D_K(i) = C_{JK}(i)/D_K(i) - D_J(i).$$

As in section 1, the first order Taylor series approximation:

(3.58) $$D_N(i_0 + tN) \approx D_N(i_0)[1 - D_N D_N(i_0)t],$$

can be substituted into the exponential identity (3.9) and integrated with respect to t to produce:

(3.59) $$P(i_0 + \Delta iN)/P(i_0) \approx \exp[-\Delta i D_N(i_0)(1 - D_N D_N(i_0)\Delta i/2)].$$

A calculation shows that (3.59) is equivalent to the second order exponential approximation in (3.12). In a similar way, the second order approximation in (3.7) can be restated as:

(3.60) $$P(i_0 + \Delta iN)/P(i_0) \approx 1 - \Delta i D_N(i_0)[1 - (D_N D_N(i_0) + D_N(i_0))\Delta i/2].$$

As was the case in section 1.c., we see that these second order approximations can be interpreted as the corresponding first order approximations with adjusted directional duration values. The adjustments again correspond to a yield change of $\Delta i/2$.

In a similar fashion, the approximation:

(3.61) $$D_k(i_0 + t\,\Delta i) \approx D_k(i_0)[1 - t\sum_j D_j D_k(i_0)\,\Delta i_j],$$

can be substituted into the exponential identity (3.33), with $\Gamma(t) = i_0 + t\,\Delta i$, and integrated to obtain:

(3.62) $$P(i_0 + \Delta i)/P(i_0) \approx \exp[-\Sigma\,\Delta i_k D_k(i_0)[1 - \sum_j D_j D_k(i_0)\,\Delta i_j/2]].$$

This exponential approximation is equivalent to (3.36) with $\Gamma(t) = i_0 + t\,\Delta i$. Finally, the second order approximation of (3.22) can be restated:

(3.63) $$P(i_0 + \Delta i)/P(i_0) \approx 1 - \Sigma\,\Delta i_k D_k(i_0)[1 - \sum_j (D_j D_k(i_0) + D_j(i_0))\,\Delta i_j/2].$$

## a. Partial Duration and Convexity Estimates

In general, one can only apply the various derivative based definitions directly when cash flows are fixed and independent of interest rates. For example, when financial options do not exist which make cash flows "interest sensitive."

For example, given a fixed vector of annual cash flows, $K = (c_1, \ldots, c_m)$, and a corresponding spot rate vector, $\mathbf{i} = (i_1, \ldots, i_m)$, the price function is given by:

(4.1) $$P(\mathbf{i}) = \Sigma c_j v_j^{\,j},$$

where $v_j = (1 + i_j)^{-1}$. A simple calculation produces:

(4.2) $$D_j(\mathbf{i}) = \frac{j c_j v_j^{\,j+1}}{P(\mathbf{i})},$$

(4.3) $$C_{jj}(\mathbf{i}) = \frac{j(j+1) c_j v_j^{\,j+2}}{P(\mathbf{i})}, \qquad C_{jk}(\mathbf{i}) = 0, \; j \neq k.$$

In this context, it is obvious that these partial durations sum to duration, and similarly for the partial convexities. In addition, because $C(\mathbf{i})$ is a diagonal matrix, the second order formulas simplify. For example,

$$P(i + \Delta i)/P(i) = 1 - \Sigma D_j(i)\,\Delta i_j + \tfrac{1}{2}\Sigma C_{jj}(i)(\Delta i_j)^2. \tag{4.4}$$

In the real world, however, many financial instruments contain options. Assets can be pre-paid (i.e. "called") at the option of the borrower for a fixed price. Liability streams associated with guaranteed interest contracts (GICs), single premium deferred annuities (SPDAs), savings accounts, etc., usually contain withdrawal (i.e. "put") options which benefit the contractholder. Also, contractholder call options are common, whereby the contractholder can invest more in the original contract.

For such cash flow streams, the formal derivatives of the price function involve both derivatives of the interest factors, as in this paper's examples, and derivatives of the cash flow stream itself. Typically, cash flow sensitivity cannot be modelled directly in closed mathematical form which lends itself to differentiation. Rather, this sensitivity is modelled discretely via interest rate projections and "if-then" algorithms.

So-called "option pricing" models are common today ([5],[7],[8],[11]). With them, P(i) and P(i) are not defined directly in terms of discounted cash flows, but rather, are defined indirectly in a manner which reflects the effect of options on the cash flow stream. These models are stochastic, in that a variety of future projections are encompassed and summarized, rather than deterministic, whereby the future is treated as known. Naturally, such option pricing models produce a price which is very much a function of the yield curve assumed, so in particular, the price function can be discretely estimated.

As common as such models are today, so it is common to use discrete definitions of duration and convexity. For example, one can estimate $D(i)$ and $C(i)$ by the following central difference formulas:

(4.5) $$D^{\epsilon}(i) = -[P(i + \epsilon) - P(i - \epsilon)] / 2\epsilon P(i),$$

(4.6) $$C^{\epsilon}(i) = [P(i + \epsilon) - 2P(i) + P(i - \epsilon)] / \epsilon^2 P(i).$$

Forward difference formulas are also common, even though they can often be "biased." That is, they better reflect sensitivity to an increase in interest rates, rather than sensitivity to change in general. Of course, formulas (4.5) and (4.6) readily generalize to directional duration and convexity estimates. For this purpose, $P(i)$ is interpreted as $P(\mathbf{i_0})$, and $P(i + \epsilon)$ interpreted as $P(\mathbf{i_0} + \epsilon\mathbf{N})$, where $\mathbf{N}$ is the direction vector. In the special case where $\mathbf{N} = (1,\ldots,1)$, the parallel shift vector, the formulas above provide estimates for the parallel shift approach discussed in section 2.b.

As for the proper value of $\epsilon$, one commonly uses judgement and some trial and error. Theoretically, one can estimate the error in the duration and convexity estimates in (4.5) and (4.6) by expanding $P(i + \epsilon)$ and $P(i - \epsilon)$ as Taylor series in $\epsilon$ and substituting into the respective formulas. This produces:

(4.7) $$D^{\epsilon}(i) - D(i) = -P^{(3)}(i)\epsilon^2/6P(i) + \mathcal{O}(\epsilon^4),$$

(4.8) $$C^{\epsilon}(i) - C(i) = P^{(4)}(i)\epsilon^2/12P(i) + \mathcal{O}(\epsilon^4).$$

As can be seen from these formulas, the duration and convexity estimates improve quickly as $\epsilon$ decreases. However, the third and fourth derivatives of $P(i)$ are generally not known, so the direct application of (4.7) and (4.8) to select an $\epsilon$ with a given error tolerance first requires their estimation. Logically, this formal approach is iterative in that an $\epsilon$ is chosen, higher order derivatives estimated, and the approximate error evaluated via (4.7) and (4.8). If necessary, the process is repeated. Error estimates so derived are only approximate since the estimated higher order derivatives will also contain errors depending on yet higher order derivatives. In practice, however, good results can often be obtained with $\epsilon$ equal to 1 to 5 basis points.

To calculate the various directional derivatives and convexities using Proposition 8, it is sufficient to estimate only the partial duration and convexity values. The above formulas generalize in the natural way to:

(4.9) $$D_j^{\epsilon}(i) = -[P(i + \epsilon_j) - P(i - \epsilon_j)] / 2\epsilon_j P(i),$$

$$(4.10) \qquad C_{jk}^{\epsilon}(i) = [P(i + \epsilon_j + \epsilon_k) - P(i - \epsilon_j + \epsilon_k) - P(i + \epsilon_j - \epsilon_k) + P(i - \epsilon_j - \epsilon_k)] \,/\, 4\epsilon_j\epsilon_k P(i).$$

Here, $\epsilon_j = \epsilon_j(0,\ldots,1,\ldots0)$, where $\epsilon_j$ is the jth coordinate, and $\epsilon = (\epsilon_1,\ldots,\epsilon_m)$. As was true for the one variable model, judgement and trial and error are needed to determine an appropriate set of values for $\epsilon_j$, which could be chosen to be equal for simplicity. Error estimation formulas generalizing (4.7) and (4.8) can again be developed using multivariate Taylor series expansions, to produce:

$$(4.11) \qquad D_j^{\epsilon}(i) - D_j(i) = -P_j^{(3)}(i)\epsilon_j^2/6P(i) + \mathcal{O}(\epsilon_j^4)$$

$$(4.12) \qquad C_{jk}^{\epsilon}(i) - C_{jk}(i) = [\epsilon_j^2 P_{jk}^{(3,1)}(i) + \epsilon_k^2 P_{jk}^{(1,3)}(i)]/6P(i) + \mathcal{O}(\epsilon_j, \epsilon_k)^4.$$

In (4.11), $P_j^{(3)}$ denotes the third partial derivative with respect to $i_j$, while in (4.12), the (3,1) and (1,3) notation denotes the corresponding mixed fourth order partial derivatives with respect to j and k. The second term on the right in (4.12) denotes a homogeneous fourth order polynomial in $\epsilon_j$ and $\epsilon_k$, which for $\epsilon_j = \epsilon_k$ becomes $\mathcal{O}(\epsilon^4)$. In practice, 1 to 5 basis points will often suffice.

As a final comment, it should be noted that partial duration and convexity estimates should be "normalized" to satisfy Proposition 9. That is, these values should be scaled so that they sum to the estimated duration or convexity values, respectively.

b. Price Sensitivity - Direct Yield Curve Approach

Once the partial durations have been calculated, the first important exercise is one of observation. Since duration equals the sum of the partial durations, one can observe to what extent parallel price sensitivity as measured by $D(i_0)$ decomposes along the yield curve. In general, price sensitivity to nonparallel shifts will be greater if the partial durations are large, with some positive and others negative, rather than relatively uniform of size $D(i_0)/m$.

For example, the duration of the price function defined in (2.3) equalled .0136, implying relatively little interest sensitivity. However, this value was seen to decompose into partial durations of $D_1(i_0) = -1.4902$ and $D_2(i_0) = 1.5038$, which had the effect of "leveraging" some nonparallel shifts into a great deal of price sensitivity. By "leveraging" is meant that the change in price observed could be very large or of the opposite orientation relative to what would have been estimated based on $D(i)$ and the actual values of $\Delta i_j$.

In those examples, had both partial durations been equal to .0068, this leveraging would not have occurred. That is, the actual change in price would have been estimable by the duration, $D(i)$, and a yield change value within the range of the $\Delta i_j$ values.

Specifically, for $\Delta i$ equal to the simple average of the $\Delta i_j$. On the other hand, had the total duration vector been given by $\mathbf{D}$ = (-10.4902, 10.5038), more leveraging would have been observed for nonparallel yield curve shifts.

As an example, assume that $\boldsymbol{\Delta i}$ = (.0025, .0075). Using (3.40), we see that for the original total duration vector, $\mathbf{D}$ = (-1.4902, 1.5038), the equivalent parallel shift would have been $\Delta i$ = .5554. For the uniform vector, $\mathbf{D}$ = (.0068, .0068), the equivalent parallel shift is $\Delta i$ = .005 as expected. Finally, for the vector $\mathbf{D}$ = (-10.4902, 10.5038), the equivalent parallel shift is calculated to be $\Delta i$ = 3.8642.

Beyond this informal exercise of observation, one can formally calculate price sensitivity a number of ways. By definition, the duration value, $D(\mathbf{i_0})$, reflects sensitivity to parallel yield curve shifts, while the various *partial durations*, $D_j(\mathbf{i_0})$, *reflect* sensitivity to changes in the yield curve point by point. Similarly, for a given direction vector, $\mathbf{N}$, one can calculate the directional duration $D_N(\mathbf{i_0})$ from (3.41). This value then reflects price sensitivity to yield curve shifts which are proportional to $\mathbf{N}$.

One direction vector of note is $\mathbf{N_0}$ as defined in (3.50). As demonstrated in Proposition 10, this vector represents the yield curve shift which produces the maximum value of $D_N(\mathbf{i_0})$, and consequently, the greatest relative change in the price function, given $|N| = 1$. Similarly, yield curve shifts proportional to $N_0$ also provide extreme values of $D_N(i_0)$, and hence, represent yield curve directions of maximal relative price sensitivity. By Proposition 10, the length of the total duration vector, $|D(i_0)|$, quantifies the amount of this maximal relative price sensitivity.

Clearly, the value of $|D(i_0)|$ provides a more rigorous basis for the "leveraging" effect discussed above. For the three total duration vectors considered above with $D(i_0) = .0136$, the corresponding values of $|D(i_0)|$ are:

(4.13a) $$|(-1.4902, 1.5038)| = 2.1171,$$

(4.13b) $$|(.0068, .0068)| = .0096,$$

(4.13c) $$|(-10.4902, 10.5038)| = 14.8450.$$

From Proposition 11, it is clear that of all two-dimensional total duration vectors with $D(i_0) = .0136$, the vector in (4.13b) is of minimal length. Naturally, there is no corresponding duration vector of maximal length given $D(i_0)$, so any amount of leveraging is possible at least in theory.

To formalize the notion of leveraging exemplified above, we seek a relationship between a yield curve shift, $\boldsymbol{\Delta}\mathbf{i}$, and the equivalent parallel shift value, $\Delta i$, so that the change in price due to $\boldsymbol{\Delta}\mathbf{i}$ is estimable with $D(i_0)$ and $\Delta i$. By (3.40), for $D(i_0) \neq 0$ the parallel shift equivalent, $\Delta i$, of the vector $\boldsymbol{\Delta}\mathbf{i}$, is given by:

$$(4.14) \qquad \Delta i = \frac{\mathbf{D}(i_0) \cdot \boldsymbol{\Delta i}}{D(i_0)} = \frac{D_{\Delta i}(i_0)}{D(i_0)}.$$

Consequently, by Proposition 10, we have

$$(4.15) \qquad |\Delta i| \le \frac{|\mathbf{D}(i_0)|}{|D(i_0)|} \cdot |\boldsymbol{\Delta i}|,$$

and the upper bound in (4.15) is achieved for $\boldsymbol{\Delta i}$ proportional to $\mathbf{D}(i_0)$.

This analysis motivates the following definition:

__Definition 4.1__ Let $P(i)$ be a price function. The __durational leverage__ of $P(i)$ at $i_0$ is defined when $P(i_0)$, $D(i_0) \neq 0$ as follows:

$$(4.16) \qquad L(i_0) = |\mathbf{D}(i_0)|/|D(i_0)|. \quad \|$$

From (4.15) we see that given $\boldsymbol{\Delta i}$, the corresponding parallel shift value can be as large as $L(i_0)$ times $|\boldsymbol{\Delta i}|$. In addition, this maximum value is attained for shifts proportional to $\mathbf{D}(i_0)$. The durational leverage values corresponding to the examples in (4.13) are easily calculated to be 155.67, .71, and 1091.54, respectively. By Proposition 11, it is clear that:

$$(4.17) \qquad L(i_0) \ge 1/\sqrt{m},$$

with equality if and only if $D_j(i_0) = D(i_0)/m$ for all $j$. As was the case for $|D(i_0)|$, $L(i_0)$ has no upper bound in theory.

### c. Price Sensitivity - Yield Curve Slope Approach

One relatively common generalization today of the "parallel shift" model is the "linear shift" model. That is, where the direction vector, $L = (l_1, \ldots, l_m)$ is defined by:

$$l_j = am_j + b, \tag{4.18}$$

where $m_j$ denotes the time value of the pivotal yield curve point, $i_j$. For example, one might have $m_1 = .25$, $m_2 = .5$, $m_3 = 1$, etc.

For such yield curve shifts, the associated directional duration and convexity functions are readily calculated by Proposition 8. For example, the directional duration is given by:

$$D_L(i_0) = a\Sigma m_j D_j(i_0) + bD(i_0). \tag{4.19}$$

That is, the directional duration naturally splits into two first order components. The first component, $\Sigma m_j D_j(i_0)$, reflects price sensitivity to yield slope changes, while the second component, $D(i_0)$, reflects price sensitivity to parallel yield changes as expected.

Similarly, the directional convexity is calculated to be:

$$C_L(\mathbf{I_0}) = a^2\Sigma\Sigma m_j m_k C_{jk}(\mathbf{i_0}) + 2ab\Sigma\Sigma m_j C_{jk}(\mathbf{i_0}) + b^2 C(\mathbf{i_0}). \tag{4.20}$$

Here we have used the symmetry of $C(\mathbf{i_0})$; that is, $C_{jk} = C_{kj}$. Unlike duration, the directional convexity splits into three components, reflecting quadratic sensitivities to slope and level changes, as well as a mixed slope/level sensitivity term. Analogous to (4.19), the pure parallel shift component is simply convexity, while the slope terms reflect weighted sums of partial convexities.

An alternative "slope" model involves a reparametrization of the yield curve. That is, rather than interpret the yield curve as a vector, $\mathbf{i} = (i_1,...i_m)$, a yield slope vector, $\mathbf{s} = (s_1,...,s_m)$ is defined as follows:

$$s_1 = i_1; \quad s_j = i_j - i_{j-1}, \; j = 2,...,m. \tag{4.21}$$

Clearly, $s_j$ reflects the increase (or decrease) in the yield curve between the (j-1)st and the jth rate. This change is often referred to as the "slope" between the respective yield points.

From (4.21) we have that $\mathbf{s} = \mathbf{Ai}$, where $\mathbf{A}$ is a linear transformation. Here we again follow the notational convention that $\mathbf{s}$ and $\mathbf{i}$ are interpreted as column vectors. This transformation is given by:

$$
A = \begin{vmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & & & & & \cdot & \cdot \\ \cdot & & & & & \cdot & \cdot \\ \cdot & & & & & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{vmatrix}. \tag{4.22}
$$

That is, $A = (a_{jk})$, where

$$
a_{jk} = \begin{cases} 1 & j = k, \\ -1 & j = k + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}
$$

Because $A$ is linear, shifts in the yield rate vector readily translate into shifts in the yield slope vector. That is,

$$
\Delta s = A \, \Delta i \tag{4.23}
$$

It is easy to see that $A$ is invertible, with:

$$
A^{-1} = \begin{vmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & & & & \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{vmatrix}. \tag{4.24}
$$

That is, $A^{-1} = B$ where:

$$
b_{jk} = \begin{cases} 1 & j \geq k \\ 0 & \text{otherwise.} \end{cases} \tag{4.25}
$$

Based on this transformation, it is possible to convert the various approximation formulas in section 3 from functions of $\Delta i$ to functions of $\Delta s$.

For example, we have from (3.22):

$$P(i_0 + \Delta i)/P(i_0) \approx 1 - D(i_0)\Delta i + \tfrac{1}{2}\Delta i^T C(i_0)\Delta i. \tag{4.26}$$

Here, the duration term is rewritten in matrix form rather than as a dot product, with $D(i_0)$ treated as a row matrix. Substituting $\Delta i^T = [A^{-1}\Delta s]^T$, and using the property of transpose that $(XY)^T = Y^T X^T$, we get:

$$P(i_0 + \Delta i)/P(i_0) \approx 1 - D_s(i_0)\Delta s + \tfrac{1}{2}\Delta s^T C_s(i_0)\Delta s, \tag{4.27}$$

where $\Delta s$ is given by (4.23) and:

$$D_s(i_0) = D(i_0)A^{-1}, \tag{4.28}$$

$$C_s(i_0) = (A^{-1})^T C(i_0)A^{-1}. \tag{4.29}$$

Here, $D_s(i_0)$ and $C_s(i_0)$ are the total duration vector and total convexity matrix, respectively, defined in the context of the yield slope vectors.

A calculation shows that the total duration vector is given by:

$$D_s(i_0) = \left(\sum_1^m D_j(i_0), \sum_2^m D_j(i_0), \ldots, D_m(i_0)\right). \tag{4.30}$$

That is, the relative sensitivity of the price function to the $j$th slope, $\Delta s_j$, is the sum of the partial durations from the $j$th to the $m$th value. Not surprisingly, the sensitivity of the price function to $\Delta s_1$ equals the duration $D(i_0)$, since $\Delta s_1 = \Delta i_1$, and for this yield curve parametrization, $\Delta i_1$ determines the change in the "level" of the yield curve.

Analogously, the total convexity matrix reflects sums of partial convexities as follows:

$$(4.31)\qquad (C_s(i_0))_{jk} = \sum_{a=j}^{m} \sum_{b=k}^{m} C_{ab}(i_0),$$

where the jkth term quantifies the sensitivity of the price function to the product of the jth and kth slopes, i.e. $\Delta s_j \Delta s_k$. Again not surprisingly, the sensitivity to $(\Delta s_1)^2$ is the convexity $C(i_0)$.

Although perhaps not readily apparent, the total duration vector and convexity matrix defined in (4.30) and (4.31) could have been calculated directly from Definition 3.5 by defining the price function directly in terms of $\mathbf{s}$. In particular, given $P(\mathbf{i})$, let the price function $R(\mathbf{s})$ be defined by:

$$(4.32)\qquad R(\mathbf{s}) = P(A^{-1}\mathbf{s}).$$

Then $\mathbf{D_s}(i_0)$ as defined in (4.30) is just the total duration vector of $R(\mathbf{s})$ evaluated at $\mathbf{s}_0 = A\mathbf{i}_0$. Similarly, $\mathbf{C_s}(i_0)$ is the total convexity matrix of $R(\mathbf{s})$.

# APPENDIX

**Proposition** Let $P(i)$ be a smooth price function and let $\{i_j\}$ define a partition of the interval $[i_0, i]$,

(A.1) $$i_j = i_0 + (j/n)\Delta i, \qquad j = 0, 1, \ldots, n$$

where $\Delta i = i - i_0$. Further, let $K_n$ be defined as the approximation to $P(i)/P(i_0)$ obtained by applying (1.12) to the terms in (1.29):

(A.2) $$K_n = \prod_{j=1}^{n} (1 - D(i_{j-1})\, \Delta i/n + \tfrac{1}{2}C(i_{j-1})(\Delta i/n)^2).$$

Then, if $D(i_{j-1}) = D(i_0)$ and $C(i_{j-1}) = C(i_0)$:

(A.3) $$\lim_{n \to \infty} K_n = \exp\,[-D(i_0)\, \Delta i].$$

Further, if $D(i_{j-1}) = D(i_0) + [D^2(i_0) - C(i_0)](j - 1)\, \Delta i/n$ and $C(i_{j-1}) = C(i_0)$:

(A.4) $$\lim_{n \to \infty} K_n = \exp\,[-D(i_0)\, \Delta i + \tfrac{1}{2}[C(i_0) - D^2(i_0)](\Delta i)^2].$$

Finally, for exact values of $D(i_{j-1})$ and $C(i_{j-1})$:

(A.5) $$\lim_{n\to\infty} K_n = \exp\left[-\int_{i_0}^{i} D(y)\,dy\right].$$

For all three limits above, the conclusions are the same if $K_n$ is defined with respect to the linear approximation in (1.5) rather than the quadratic estimate (1.12).

<u>Proof</u> Because $P''(i)$ is continuous, $C(i)$ and $D(i)$ are bounded on $(i_0, i)$. Hence, an initial value of $n_0$ can be chosen so that for $n \geq n_0$, $K_n$ equals the product of positive factors. For such an $n$, $\ln(K_n)$ is therefore well defined. Because $\ln x$ is a continuous function, as is its inverse $e^x$, $K_n$ will converge if and only if $\ln(K_n)$ converges.

Assume that $D(i_{j-1}) = D_0$ and $C(i_{j-1}) = C_0$. Then:

(A.6) $$\ln(K_n) = \sum_{j=1}^{n} \ln\left[1 - D_0\,\Delta i/n + \tfrac{1}{2}C_0(\Delta i)^2/n^2\right].$$

Using the Taylor series expansion,

(A.7) $$\ln(1 + x) = x + \mathcal{O}(x^2),$$

which is allowable because the arguments in (A.6) are uniformly bounded for $n \geq n_0$, we get:

(A.8) $$\ln(K_n) = \sum_{j=1}^{n} [-D_0\,\Delta i/n + ½C_0(\Delta i)^2/n^2 + \theta(1/n^2)]$$

$$= -D_0\,\Delta i + ½C(\Delta i)^2/n + \theta(1/n).$$

From (A.8), limits are readily taken to prove (A.3).

Using a similar argument, assume that $D(i_{j-1}) = D_0 + E_0(j-1)\Delta i/n$, where $E_0 = D_0^2 - C_0$, and $C(i_{j-1}) = C_0$. Then for n sufficiently large:

(A.9) $$\ln(K_n) = \sum_{j=1}^{n} \ln\left(1 - D_0\,\Delta i/n - E_0(j-1)(\Delta i)^2/n^2 + ½C_0(\Delta i)^2/n^2\right).$$

Again using (A.7), and $\sum_{j=1}^{n}(j-1) = n(n-1)/2$, we get:

(A.10) $$\ln(K_n) = -D_0\,\Delta i - ½E_0(\Delta i)^2(n-1)/n + ½C_0(\Delta i)^2/n + \theta(1/n).$$

Taking limits in (A.10) demonstrates (A.4).

Using exact values for $D(i_{j-1})$ and $C(i_{j-1})$ and (A.7):

(A.11) $$\ln(K_n) = \sum_{j=1}^{n} \ln\left(1 - D(i_{j-1})\Delta i/n + ½C(i_{j-1})(\Delta i)^2/n^2\right)$$

$$= -\sum_{j=1}^{n} D(i_{j-1})\Delta i/n + (\Delta i/2n)\sum_{j=1}^{n} C(i_{j-1})\Delta i/n + \theta(1/n).$$

Taking limits in (A.11), we see that the first summation converges to the Riemann integral of D(y). The second term converges to zero because the summation converges to the integral of $C(y)$, while its coeffecient converges to 0. Hence, (A.5) is demonstrated.

Finally, had the first order approximation been used in the definition of $K_n$, the same limits would have resulted. This is due to the fact that in each case above, the convexity adjustment was seen to be $\mathcal{O}(1/n)$, and consequently added nothing in the limit. ||

## REFERENCES

[1] Bierwag, G.O. Duration Analysis: Managing Interest Rate Risk. Cambridge, MA: Ballinger Publishing Company (1987).

[2] Bierwag, G.O.. "Immunization, Duration and the Term Structure of Interest Rates," Journal of Financial and Quantitative Analysis, 12 (December, 1977), 725-742.

[3] Bierwag, G.O.; G.C. Kaufman; and C. Khang. "Duration and Bond Portfolio Analysis: An Overview," JFQA, 13 (November, 1978), 671-685.

[4] Bierwag, G.O.; G.C. Kaufman; and A. Teovs. "Bond Portfolio Immunization and Stochastic Process Risk," Journal of Bank Research, (Winter, 1983).

[5] Black, F.; and M. Scholes. "The Pricing of Options and Corporate Liabilities," The Journal of Political Economy, 3 (May -June 1973), 637-654.

[6] Chambers, D.R.; W.T. Carleton; and R.W. McEnally. "Immunizing Default-Free Bond Portfolios with a Duration Vector," JFQA, 23 (March, 1988), 89-104.

[7] Clancy, R.P. "Options on Bonds and Applications to Product Pricing," Transactions of the Society of Actuaries, XXXVII (1985), 97-130.

[8] Cox, J.C.; S.A.Ross; and M. Rubenstein. "Option Pricing: A Simplified Approach," Journal of Financial Economics, 7 (1979), 229-263.

[9] Fisher, L; and R.L. Weil. "Coping with the Risk of Interest Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies," Journal of Business (October, 1971), 408-431.

[10] Hicks, J.R. Value and Capital. Oxford University Press, 1939.

[11] Ho, T.S.Y.; and S. Lee. "Term Structure Movements and Pricing Interest Rate Contingent Claims," Journal of Finance, 41 (1986), 1011-1029.

[12] Ingersoll, J; J. Skelton; and R. Weil. "Duration Forty Years Later," JFQA, 13 (November, 1978), 627-650.

[13] Jacob, D.; G. Lord; and J. Tilley. "Price, Duration and Convexity of a Stream of Interest-Sensitive Cash Flows," Morgan Stanley & Co., (April, 1986)

[14] Jacob, D.; G. Lord; and J. Tilley. "A Generalized Framework for Pricing Contingent Cash Flows," Financial Management, (Autumn, 1987)

[15] Macaulay, F.R. Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the U.S. Since 1856. New York: National Bureau of Economic Research, 1938.

[16] Redington, F.M. "Review of the Principle of Life Office Valuations," Journal of the Institute of Actuaries, vol. 18 (1952), 286-340.

[17] Samuelson, P.A. "The Effect of Interest Rate Increases on the Banking System," American Economic Review (March, 1945), 16-27.

[18] Stock, D.; and D.G. Simonson. "Tax Adjusted Duration for Amortizing Debt Instrument," JFQA, 23 (September, 1988), 313-327.

[19] Vanderhoof, I.T. "The Interest Rate Assumptions and the Maturity Structure of the Assets of a Life Insurance Company," Transactions of the Society of Actuaries, XXIV (1972), 157-192.

[20] Vanderhoof, I.T. "Interest Rate Assumptions and the Relationship Between Asset and Liability Structure," Part 8 Study Notes, Society of Actuaries (8-201-79).