V

Multivariate Duration Analysis

Robert R. Reitano

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, changes in which each yield point moves by the same amount. More recently, 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 happens 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 that is equally applicable in virtually all yield curve environments. In addition, this model is practical and relatively easy to apply.

To motivate the multivariate approach, simple examples are presented that demonstrate the limitations of the traditional model when yield curve shifts are not parallel. Multivariate models are then developed in detail and shown to readily overcome these limitations. Examples are utilized throughout to make the theory more accessible. The last section focuses on applications of these models as well as on a variety of practical considerations.

1. Introduction

The concept of duration has generated a great deal of interest and research activity during its relatively short history. Bierwag, Kaufman and Khang [3] and Ingersoll, Skelton and Weil [13] 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 are only highlighted here.

The notion of duration was independently discovered by at least four authors. The earliest source is Macaulay [16], 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, and his measure reflected a weighted average of the times to maturity. At about the same time, Hicks [10] developed the same duration formula, calling it the "average period," analyzing the price sensitivity of an income stream to changes in the underlying interest rate. Specifically, the Macaulay duration equalled the elasticity of the price of a bond with respect to $v = (1 + i)^{-1}$.

A number of years later, Redington [17] and Samuelson [25] discovered a very similar formula analyzing questions in what has come to be known as immunization theory. Redington sought to "immunize" a liability stream with an asset stream. This meant that the value of each was to be equally responsive to changes in the underlying interest rate. This was accomplished by equating first derivatives of the associated price or present value functions, thereby introducing the approach to duration that was later generalized in the development of what 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 formulations, the price function and the corresponding duration measure were defined in terms of "the interest rate," which was typically taken as the yield to maturity. This approach was also followed in Vanderhoof [27], [28], which adapted the Redington model and became, to many actuaries, an introduction to this field of thought. Fisher and Weil [9] later generalized the notion of duration so that the price function could reflect a complete yield curve. In this context, a change in yields was modeled in terms of a parallel yield curve shift, whereby each yield rate is changed by the same amount. This duration measure has sometimes been referred to as D_2, to distinguish it from the Macaulay duration, denoted D_1. Corresponding to other models of yield curve shifts, other duration measures have been defined (see [1]–[4], [14], and [15], for example). In [4], it is also shown that losses associated with choosing the wrong model can be substantial.

More recently, Stock and Simonson [26] have analyzed after-tax adjustments to price sensitivity, while Chambers, Carleton and McEnally [6] have explored the notion of a duration vector in immunizing default-free bond portfolios. In this latter paper, the various components of the duration vector correspond to cash-flow-weighted moments of the adjusted times to maturity. The first component is similar to D_2, while the second reflects a measure of the average time squared, then average time cubed, and so on. 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 that is applicable to virtually any model of yield curve movements. Of course, multivariate models have been used elsewhere ([1] and [12], for example). The purpose here is to explore the general mathematical theory and its applications in some detail. In particular, two general multivariate approaches are analyzed that are relatively easy to apply, yet provide a clearer understanding of the yield curve risks inherent in the portfolio being analyzed.

Common to both approaches is a discrete representation of a yield curve. Although this curve is usually visualized as a continuous function, in practice it is typically generated by yield values at well-defined pivotal points. These "yield curve drivers" usually correspond to semiannual yields at the actively traded commercial paper, note, and bond maturities. For example, one might base a yield curve on observed market yields at maturities of 0.25, 0.5, 1, 2, 3, 4, 5, 7, 10, 20, and 30 years. Given these observed yields, the remainder of the yield curve is then generated by interpolation. Consequently, these other yields are functionally dependent on the observed values. That is, the yield curve continuum is in practice equivalent to an m-point "vector" of observed variables. Naturally, other discretizations are possible in theory, and many are common in practice.

Price functions can therefore be modeled in terms of these m external variables. The actual units of these observed yields are irrelevant for our purposes, as is their basis. Semiannual bond yields are as usable as effective spot rates. All that is assumed for these models is that the price function of the portfolio can be evaluated based on the yield variables used. Whether this price calculation is performed directly, such as by taking the present value of fixed cash flows, or with an option-pricing or other model is again not important for our purposes.

Given this m-point representation, two duration approaches are developed. The "directional duration" approach models yield curve shifts in terms of an arbitrary direction vector $\mathbf{N}$. That is, the initial yield curve vector, $\mathbf{i}_0$, is modeled as moving Δi units in the direction of $\mathbf{N}$. The price function, $P(\mathbf{i}_0 + \Delta i\mathbf{N})$, viewed as a function of Δi, then reflects the price sensitivity in this direction. Of course, when $\mathbf{N} = (1, 1, \ldots, 1)$, the parallel shift vector, this directional duration analysis reduces to the classical modified duration model.

A closely related model is also developed using a "partial duration" calculus. Here, the yield curve shift, Δi, is explicitly modeled as multivariate, and the price function $P(\mathbf{i}_0 + \Delta i)$ is analyzed in terms of its partial derivatives.

To motivate the use of these multivariate models, a simple example is analyzed using the traditional one-variable approach. This example reflects positive and negative cash flows, as is usually the case for the surplus or net worth portfolio. For example, a duration-matching program that uses a "barbell" or "reverse barbell" strategy (that is, intermediate liabilities funded by long and short assets, or the reverse) always produces a net worth position with "long" and "short" net positions at various points of the yield curve. In such a case, the traditional modified duration measure provides useful information about parallel yield curve shifts, as expected. However, nonparallel shifts produce price changes that are orders of magnitude larger and/ or of an opposite sign compared with the price changes the modified duration measure would suggest.

The multivariate duration approaches are then developed, and this example is revisited and shown to behave quite understandably by using these more general models. Section 5 then explores practical considerations and two applications to yield curve slope sensitivity.

This paper has been written at a level that assumes some familiarity with traditional duration analysis theory and applications. However, the examples used throughout have been kept simple and intuitive in an attempt to make the general theory accessible to even beginning practitioners. The reader is referred to Reitano [18] for a more introductory approach to the models developed here. In particular, the one-variable model and its properties are more fully developed and exemplified.

For a variety of applications of the multivariate models developed in this paper, see Reitano [19]-[24].

2. The One-Variable Model and Its Limitations

a. Definitions

Let $P(i)$ denote the price function that assigns to each interest rate $i \geq 0$, the value of a given portfolio of future cash flows. The actual rate i can be defined within any system of units—annual, semiannual, continuous, and so on—and generally follows from the context of the problem. The future cash flows can be positive or negative, fixed or dependent on i. We assume that $P(i)$ is twice differentiable and has a continuous second derivative.

Definition 2.1:

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

$$D(i) = -\frac{dP}{di}/P(i). \square \qquad (2.1)$$

Using the standard first-order Taylor series approximation, we have:

$$P(i)/P(i_0) \approx 1 - D(i_0)\,\Delta i, \qquad (2.2)$$

where $\Delta i = i - i_0$.

Definition 2.2:

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

$$C(i) = \frac{d^2P}{di^2}/P(i). \square \qquad (2.3)$$

Using the second-order Taylor series approximation:

$$P(i)/P(i_0) \approx 1 - D(i_0)\,\Delta i + 1/2C(i_0)\,(\Delta i)^2. \qquad (2.4)$$

In applications, there are two common approaches to using this model. With the yield-to-maturity approach, i_0 is taken as the (not necessarily unique) value such that $P(i_0)$ equals the given initial price. Equivalently, the yield curve is assumed to be flat with value i_0. $P(i_0 + \Delta i)$ then reflects the price when the yield to maturity is changed by Δi. The parallel-shift approach allows cash flows to be initially valued on the actual yield curve, producing the value $P(0)$. Then $P(\Delta i)$ represents the price when the yield curve is changed "in parallel" by amount Δi, that is, when each yield point is changed by this common amount. Unfortunately, the use of one-variable models is not without its limitations, as the following example demonstrates.

Assume a simple portfolio of three fixed cash flows equal to 20, –20, and 11, at time 0, 1, and 2 years, respectively. Also, assume that the one-year spot rate is 0.105 and the two-year spot rate is 0.10. For simplicity, such a spot rate curve will be denoted (0.105, 0.10). At these rates, the current price is easily calculated to be 10.99136.

b. Yield-to-Maturity Approach

Using the *yield-to-maturity (YTM) approach*, the price function $P(i)$ is modeled:

$$P(i) = 20 - 20v + 11v^2, \quad v = (1 + i)^{-1}. \qquad (2.5)$$

The equation $P(i) = 10.99136$ has two solutions: 0.00445 and 0.21565. Choosing the smaller YTM of 0.00445, the duration of $P(i)$ is calculated to be 0.172, and the convexity equals 2.308.

Using the linear approximation in (2.2):

$$P(i)/P(0.00445) \approx 1 - 0.172(i - 0.00445). \qquad (2.6)$$

If the yield curve increases uniformly by 0.01 to (0.115, 0.11), the use of 0.01445 = 0.00445 + 0.01 for i in (2.6) would yield a very poor approximation. The actual portfolio decrease in this case is 0.0067%, while this linear approximation and i value would predict a decrease of 0.17%. Making the adjustment for the convexity value of 2.308 improves the approximation slightly to a predicted decrease of 0.16%, still orders of magnitude from the correct answer.

The problem here is one of units: yield curve units versus YTM units. The proper value to use for i in (2.6) is not 0.01445, but the YTM corresponding to the yield curve (0.115, 0.11). A calculation shows this value to be 0.00485. That is, the 0.01 change in the yield curve corresponds to only a 0.0004 change in YTM, so it is obvious why the above initial approximation was so poor. Using the new YTM in (2.6) produces a predicted decrease of 0.0069%, which compares quite favorably to the actual decrease of 0.0067%. Here, the convexity adjustment is 0 to four decimal places (in percentage units).

If the larger YTM value of 0.21565 had been chosen, its negative duration of –0.117 can also be interpreted as a problem of units. That is, an increase in spot yields corresponds to a decrease in YTMs, thereby correcting for both the wrong sign and the wrong order of magnitude. Specifically, the yield curve increase of 0.01 corresponds to a YTM change of –0.0006.

Consequently, one could correct for the "units" problem inherent with the YTM approach if an appropriate conversion formula can be developed (Section 3c). However, the YTM approach also has the uncorrectable problem of nonexistence of solutions. For example, the yield curve (0.109, 0.110) produces a price for the above cash flows of 10.8936, which is below the minimum value in (2.5) of 10.909. Hence, no YTM exists, nor does an estimable Δi.

c. Parallel Shift Approach

Using the *parallel shift approach*, the price function for the above cash flows is:

$$P(\Delta i) = 20 - 20v + 11w^2, \qquad v = (1.105 + \Delta i)^{-1}, \qquad w = (1.10 + \Delta i)^{-1}. \qquad (2.7)$$

The equation $P(\Delta i) = 10.99136$ now has the obvious solution of $\Delta i = 0$. A calculation produces $D(0) = 0.0136$ and $C(0) = 1.404$. Using (2.2), $P(\Delta i)$ is linearly approximated by:

$$P(\Delta i)/P(0) \approx 1 - 0.0136\,\Delta i. \qquad (2.8)$$

For a parallel yield curve increase of 0.01 to (0.115, 0.11), the approximation in (2.8) predicts a portfolio decrease of 0.0136%, which overstates the actual decrease of 0.0067%. The convexity adjustment improves the approximation from 0.0136% to 0.0066%.

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 (0.105, 0.10) to (0.1075, 0.1075), that is, an increase of 25 basis points in the one-year spot rate and 75 basis points in the two-year value. Because the duration of the portfolio is positive at 0.0136, one expects 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 0.745%.

However, this decrease would not have been predicted from the first- or second-order approximations for $P(\Delta i)/P(0)$, choosing Δi to be equal to 25 or 75 basis points. The best of the four approximations would predict a portfolio decrease of only 0.010%, a very poor estimate. It appears that for this nonparallel yield curve change, the portfolio is far more sensitive than the duration and convexity values imply. This problem has little to do with the size of the yield curve shift.

For example, assume that the yield curve had increased only slightly from (0.105, 0.10) to (0.1052, 0.1001). This shift is positive and nearly parallel, so again a portfolio decrease is expected. However, the portfolio value actually increases in this case by 0.015%. Both linear and quadratic approximations predict decreases at both 1 and 2 basis points. The best of these approximations calls for a decrease of 0.0001%. As before, the sensitivity of the portfolio to this nonparallel shift appears much greater than $D(0)$ and $C(0)$ imply. Unlike before, not even the sign of the sensitivity is accurately predicted.

As was the case for the YTM approach, the problem here is again a problem of units. The above approximation formulas for $P(\Delta i)$ reflect the sensitivity of price to parallel shifts of the yield curve of Δi. This parallel shift is really a vector shift of $\Delta\mathbf{i}$, where $\Delta\mathbf{i} \equiv (\Delta i, \Delta i)$ represents a yield change vector that moves the yield curve from $\mathbf{i}_0 = (i_1, i_2)$, to $\mathbf{i}_0 + \Delta\mathbf{i} = (i_1 + \Delta i, i_2 + \Delta i)$. Looked at this way, the shift vector $\Delta\mathbf{i}$ encompasses a "magnitude," Δi, and a "direction," $\mathbf{N} = (1,1)$:

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

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

Decomposing the various shifts exemplified above, we obtain:

$$(0.01, 0.01) = 0.01\ (1,1) \qquad (2.10a)$$

$$(0.0025, 0.0075) = 0.0025\ (1,3) \qquad (2.10b)$$

$$(0.0002, 0.0001) = 0.0001\ (2,1). \qquad (2.10c)$$

Of course, these decompositions are not uniquely defined. The approximation formulas worked well for shift (2.10a) because the direction of change was **N** = (1,1), the direction explicitly assumed in the derivation of these formulas. Nonparallel shifts (2.10b and c) caused poor estimates because their direction vectors were not equivalent to (1,1), and for the cash flows underlying $P(\Delta i)$, this difference in directions was very important.

For notational convenience here, let $D_{(1,1)}$ denote the duration as defined in (2.2), with the underlying direction vector **N** = (1,1) explicitly displayed. For the example above, we have $D_{(1,1)} = 0.0136$. In the next section, duration and convexity are formally defined with respect to directions other than (1,1). With those definitions, one can calculate:

$$D_{(1,1)} = 0.0136 \quad C_{(1,1)} = 1.404 \qquad (2.11a)$$

$$D_{(1,3)} = 3.0212 \quad C_{(1,3)} = 34.214 \qquad (2.11b)$$

$$D_{(2,1)} = -1.4767 \quad C_{(2,1)} = -6.688 \qquad (2.11c)$$

These duration and convexity values reflect the price sensitivity to yield curve shifts in various directions. They are seen to differ greatly.

Once such *directional durations and convexities* have been defined and calculated, one can develop the corresponding approximation formulas, such as the counterpart to (2.4):

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx 1 - D_N(\mathbf{i}_0)\Delta i + 1/2C_N(\mathbf{i}_0)(\Delta i)^2. \qquad (2.12)$$

Utilizing (2.12) and the directional values in (2.11), the following improved estimates are obtained:

Shift	First Order	Second Order	Exact Value	
(0.01, 0.01)	−0.0136%	−0.0066%	−0.0067%	
(0.0025, 0.0075)	−0.7533%	−0.7446%	−0.7447%	(2.13)
(0.0002, 0.0001)	+0.0148%	+0.0148%	+0.0148%	

This multivariate approach to duration and convexity is explored in detail in Section 3.

3. Multivariate Models

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. For example, the components of this yield vector could correspond to yield curve pivotal points, such as yields for terms: 0.25, 0.5, 1, 2, 3, 4, 5, 7, 10, 20, and 30 years. These yield curve drivers are then the defining variables of the price function, since other yield values are typically interpolated and therefore dependent on these values. Also, let $\mathbf{N} = (n_1, \ldots, n_m)$ be a direction vector, $\mathbf{N} \neq \mathbf{0}$, and $|\mathbf{N}| = (\sum n_i^2)^{1/2}$ denote its length. In general, vectors will be identified with column matrices when used in matrix calculations, with the exception of the total duration vector (Section 3c), which will be identified with a low matrix.

Consider $P(t)=P(\mathbf{i}_0+t\mathbf{N})$, 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 t units in the direction of **N**, that is, where $\mathbf{i}_{01}$ is shifted tn_1 units, $\mathbf{i}_{02}$ *is* shifted tn_2 units, and so on. Using a Taylor series expansion, $P(t)$ can be approximated to first and second order in t as follows:

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

$$P(t) \approx P(0) + P'(0)t + 1/2P''(0)t^2. \qquad (3.1b)$$

In order to calculate the derivatives of $P(t)$ needed in (3.1), let $P_j(\mathbf{i})$ denote the j-th partial derivative of $P(\mathbf{i})$, and $P_{jk}(\mathbf{i})$ denote the corresponding mixed second-order partial derivative. We then obtain:

$$P'(t) = \sum n_j P_j\,(\mathbf{i}_0 + t\mathbf{N}), \qquad (3.2a)$$

$$P''(t) = \sum\sum n_j n_k P_{jk}\,(\mathbf{i}_0 + t\mathbf{N}). \qquad (3.2b)$$

Evaluated at t=0, the expressions in (3.2) are seen to be the first- and second-order directional derivatives of the price function $P(\mathbf{i})$ evaluated at $\mathbf{i}_0$; that is,

$$P'(0) \equiv \frac{\partial P}{\partial \mathbf{N}}\bigg|\mathbf{i}_0 = \sum n_j P_j(\mathbf{i}_0), \qquad (3.3a)$$

$$P''(0) \equiv \frac{\partial^2 P}{\partial \mathbf{N}^2}\bigg|\mathbf{i}_0 = \sum\sum n_j n_k P_{jk}(\mathbf{i}_0). \qquad (3.3b)$$

In anticipation of combining (3.1) and (3.3), the following definitions are motivated:

Definition 3.1:

Let $P(\mathbf{i})$ be a multivariate price function and $\mathbf{N} \neq \mathbf{0}$ a direction vector. The *directional duration function* in the direction of **N**, $D_N(\mathbf{i})$, is defined for $P(\mathbf{i}) \neq 0$ as follows:

$$D_N(\mathbf{i}) = -\frac{\partial P}{\partial \mathbf{N}}/P(\mathbf{i}). \square \tag{3.4}$$

Definition 3.2:

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

$$C_N(\mathbf{i}) = \frac{\partial^2 P}{\partial \mathbf{N}^2}/P(\mathbf{i}). \square \tag{3.5}$$

Substituting (3.3) into (3.1), the following counterparts to (2.2) and (2.4) are produced:

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx 1 - D_N(\mathbf{i}_0)\,\Delta i, \tag{3.6}$$

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx 1 - D_N(\mathbf{i}_0)\,\Delta i + 1/2C_N(\mathbf{i}_0)\,(\Delta i)^2. \tag{3.7}$$

As an example, consider the price function in (2.7) explicitly expressed as a multivariate function:

$$P(i_1,i_2) = 20 - 20v + 11w^2, \tag{3.8}$$

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:

$$P_1(i_1,i_2) = 20v^2;\ P_2(i_1,i_2) = -22w^3 \tag{3.9a}$$

$$P_{11}(i_1,i_2) = -40v^3;\ P_{22}(i_1,i_2) = 66w^4;\ P_{12} = P_{21} \equiv 0. \tag{3.9b}$$

Evaluating these derivatives on $\mathbf{i}_0 = (0.105, 0.10)$ and performing the necessary weighted summations in (3.3), the directional durations and convexities displayed in (2.11) can be readily verified.

Before continuing, note 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 $C_N(\mathbf{i}_0) = C(0)$, where these latter values are calculated utilizing the parallel shift approach. Below, these traditional values will also be denoted $D(\mathbf{i}_0)$ and $C(\mathbf{i}_0)$.

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

To be uniquely defined, one can normalize the model by requiring the direction vector **N** to satisfy $|\mathbf{N}| = 1$. The magnitude variable, Δi, is then uniquely defined as the length of the shift vector $\Delta i\mathbf{N}$. However, regardless of whether **N** is normalized, consistent estimates are produced.

(3) A variety of the duration measures developed in the past and referenced in the introduction are special cases of directional durations, because they reflect explicit models of assumed yield curve shifts.

In addition, "key rate" durations of Ho [12] are also directional durations. In this model, the yield curve components in $\mathbf{i}_0$ are spot rates, often on a monthly basis. A collection of "pyramid" direction vectors, $\mathbf{N}_j$, are then defined, such as:

$$\mathbf{N}_j = (0, \ldots, 0, 1/2, 1, 2/3, 1/3, 0, 0 \ldots).$$

The actual spot rate corresponding to the component 1 in $\mathbf{N}_j$ is the "key rate," and the various key rate durations are equivalent to the directional durations $D_N(\mathbf{i}_0)$.

The collection of pyramid direction vectors used in the Ho model form a "partition" of the parallel shift vector:

$$\Sigma\, \mathbf{N}_j = (1, 1, .., 1).$$

In Section 4a, this property will be seen to have an important corollary.

Proposition 1:

Let $P(\mathbf{i})$ be a multivariate price function and **N** a direction vector with $P(\mathbf{i}_0 + \Delta i\mathbf{N}) \neq 0$ for $|\Delta i| \leq K$. Then

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) = \exp\left[-\int_0^{\Delta i} D_N(\mathbf{i}_0 + t\mathbf{N})dt\right], \tag{3.10}$$

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

Proof: Define $f(t) = \ln|P(\mathbf{i}_0 + t\mathbf{N})|$. Then $-f'(t) = D_N(\mathbf{i}_0 + t\mathbf{N})$, which can be integrated and exponentiated to produce (3.10). $\square$

From (3.10), the following first-order exponential approximation is transparent:

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx \exp[-D_N(\mathbf{i}_0)\,\Delta i]. \tag{3.11}$$

To develop a second-order exponential formula, we must expand the exponent function in (3.10) as a Taylor series in Δi. To do this, let:

$$f(\Delta i) = \int_0^{\Delta i} D_N(\mathbf{i}_0 + t\mathbf{N})dt. \quad (3.12)$$

We then have:

$$f'(\Delta i) = D_N(\mathbf{i}_0 + \Delta i\mathbf{N}), \quad (3.13)$$
$$f''(\Delta i) = D_N^2(\mathbf{i}_0 + \Delta i\mathbf{N}) - C_N(\mathbf{i}_0 + \Delta i\mathbf{N}).$$

The second-derivative formula is readily verified by taking directional derivatives of the identity, $\partial P/\partial \mathbf{N} = -D_N P$.

Approximating $f(\Delta i)$ by a second-order Taylor series about $\Delta i = 0$ and substituting into (3.10), we obtain:

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx \exp\{-D_N(\mathbf{i}_0)\,\Delta i + 1/2\,[C_N(\mathbf{i}_0) - D_N^2(\mathbf{i}_0)](\Delta i)^2\}. \quad (3.14)$$

b. Properties of the Directional Duration Approximations

In this section, properties of the various approximations above are explored. We begin with an error analysis of the first-order estimates.

Proposition 2:

Let $P(\mathbf{i})$ be a price function which is nonzero at $\mathbf{i}_0$. Then for Δi sufficiently small:

$$\exp[-D_N(\mathbf{i}_0)\,\Delta i] < P(\mathbf{i})/P(\mathbf{i}_0) \qquad C > D^2$$
$$1 - D_N(\mathbf{i}_0)\,\Delta i < P(\mathbf{i})/P(\mathbf{i}_0) < \exp[-D_N(\mathbf{i}_0)\,\Delta i] \qquad 0 < C < D^2 \quad (3.15)$$
$$P(\mathbf{i})/P(\mathbf{i}_0) < 1 - D_N(\mathbf{i}_0)\,\Delta i \qquad C < 0$$

where $\mathbf{i} = \mathbf{i}_0 + \Delta i\mathbf{N}$, $D = D_N(\mathbf{i}_0)$, and $C = C_N(\mathbf{i}_0)$.

Proof: The bounds in (3.15) correspond to the linear and first-order exponential approximations in (3.6) and (3.11). For small Δi, the sign of the error in these first-order approximations equals the sign of the second-order terms in the respective expansions in (3.7) and (3.14). For the linear approximation, this term has the sign of $C_N(\mathbf{i}_0)$, while for the exponential approximation, this term has the sign of $C_N(\mathbf{i}_0) - D_N^2(\mathbf{i}_0)$. The bounds in (3.15) follow from this and the observation that $1 + x \le e^x$ for all x. □

Next, we investigate the conditions under which the various approximations for $P(\mathbf{i})/P(\mathbf{i}_0)$ are exact. Using the identity in Proposition 1, it is natural to expect that such exactness is related to the behavior of $D(\mathbf{i})$ near $\mathbf{i}_0$.

Proposition 3:

The various approximations for $P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0)$ will be exact if and only if $D_N(\mathbf{i})$ assumes one of the following functional forms:

Exponential Approximation	Model for $D_N(\mathbf{i})$
(3.11) 1st Order	D
(3.14) 2nd Order	$D + [D^2 - C]\,\Delta i$
Polynomial Approximation	**Model for $D_N(\mathbf{i})$**
(3.6) 1st Order	$D/(1 - D\Delta i)$
(3.7) 2nd Order	$(D - C\Delta i)/(1 - D\Delta i + 1/2C(\Delta i)^2)$

(3.16)

where $\mathbf{i} = \mathbf{i}_0 + \Delta i\mathbf{N}$, $D = D_N(\mathbf{i}_0)$, and $C = C_N(\mathbf{i}_0)$.

Proof. The models for $D_N(\mathbf{i})$ in (3.16) can be derived by equating the exact value of $P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0)$ as given in (3.10) to the respective approximations, and solving for $D_N(\mathbf{i})$. Although integral equations are encountered, these are easily solved by first taking logarithms, then differentiating with respect to Δi. □

Note that the underlying model for $D(\mathbf{i})$ in (3.6) can be counter-intuitive. A calculation shows that this function is an increasing function of Δi, while $D_N(\mathbf{i})$ is an increasing function locally only when it has a positive directional derivative. Based on (3.13), this occurs only when $D_N^2(\mathbf{i}_0)$ exceeds $C_N(\mathbf{i}_0)$. While somewhat more complicated, the model for $D_N(\mathbf{i})$ underlying (3.7) does not have this potential problem, in that it too will be an increasing function locally only when $D_N^2(\mathbf{i}_0)$ exceeds $C_N(\mathbf{i}_0)$.

As a final investigation, it is next shown that each of the exponential relationships in (3.10), (3.11), and (3.14) equals the limiting case of applying the linear approximation in (3.6) to ever finer subdivisions of the segment from $\mathbf{i}_0$ to $\mathbf{i}$. The formula that results depends on the assumption made about the values of $D_N(\mathbf{i})$ in this approximation.

To this end, let $\mathbf{i}_0$ and $\mathbf{i} = \mathbf{i}_0 + \Delta i\mathbf{N}$ be given and define a subdivision of the corresponding segment by:

$$\mathbf{i}_j = \mathbf{i}_0 + \frac{j}{n}\Delta i\mathbf{N}, \quad j = 0, \ldots, n. \quad (3.17)$$

Clearly, we have that:

$$\frac{P(\mathbf{i})}{P(\mathbf{i}_0)} = \prod_{j=1}^{n} \frac{P(\mathbf{i}_j)}{P(\mathbf{i}_{j-1})}. \quad (3.18)$$

Applying the linear approximation in (3.6) to each term in this product, let:

$$K_n = \prod_{j=1}^{n}[1 - D_N(\mathbf{i}_{j-1})(\Delta i/n)]. \quad (3.19)$$

Proposition 4:

Let K_n be defined as in (3.19) above. Then:

$$\lim(K_n) = \exp\left[-\int_0^{\Delta i} D_N(\mathbf{i}_0 + t\mathbf{N})dt\right], \quad (3.20)$$

as $n \to \infty$.

Proof: Because $P(\mathbf{i})$ is twice continuously differentiable by assumption, $D_N(\mathbf{i})$ is bounded on the segment $[\mathbf{i}_0, \mathbf{i}]$. 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.

Now,

$$\ln(K_n) = \sum_{j=1}^{n} \ln[1 - D_N(\mathbf{i}_{j-1})(\Delta i/n)]$$
$$= -\sum_{j=1}^{n} D_N(\mathbf{i}_{j-1})(\Delta i/n) + \mathcal{O}(1/n) \quad (3.21)$$

Taking limits in (3.21), we see that the summation converges to the Riemann integral of $D_N(\mathbf{i})$ as in (3.20). □

As is easily seen, if $D_N(\mathbf{i}_{j-1})$ in (3.19) is set equal to $D_N(\mathbf{i}_0)$, or approximated linearly by $D_N(\mathbf{i}_0) + [D_N^2(\mathbf{i}_0) - C_N(\mathbf{i}_0)](j-1)\,\Delta i/n$, the corresponding limits are equal to the approximations in (3.11) and (3.14), respectively.

c. *Partial Durations and Convexities*

As shown in Section 3a, the classical duration and convexity analysis of Section 2 can be readily generalized to include yield curve shifts that are not parallel. An alternative model would be one that more explicitly recognizes the multivariate nature of yield curve changes, that is, a model that 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:

$$P(\mathbf{i}_0 + \Delta\mathbf{i}) \approx P(\mathbf{i}_0) + \Sigma P_j(\mathbf{i}_0)\,\Delta i_j, \quad (3.22a)$$

$$P(\mathbf{i}_0 + \Delta\mathbf{i}) \approx P(\mathbf{i}_0) + \Sigma P_j(\mathbf{i}_0)\,\Delta i_j + 1/2\Sigma\Sigma P_{jk}(\mathbf{i}_0)\,\Delta i_j\,\Delta i_k. \quad (3.22b)$$

These approximations naturally motivate the following definitions:

Definition 3.3:

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

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

Definition 3.4:

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

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

Definition 3.5:

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

$$\mathbf{D}(\mathbf{i}) = (D_1(\mathbf{i}), \ldots, D_m(\mathbf{i})), \quad (3.25)$$

$$\mathbf{C}(\mathbf{i}) = \begin{pmatrix} C_{11}(\mathbf{i}) & \ldots & C_{1m}(\mathbf{i}) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ C_{m1}(\mathbf{i}) & \ldots & C_{mm}(\mathbf{i}) \end{pmatrix}. \ \square \quad (3.26)$$

Utilizing these definitions in (3.22), the following generalizations of (2.2) and (2.4) are produced:

$$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0) \approx 1 - \mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i} \quad (3.27)$$

$$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0) \approx 1 - \mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i} + 1/2\Delta\mathbf{i}^T\mathbf{C}(\mathbf{i}_0)\,\Delta\mathbf{i}. \quad (3.28)$$

To simplify notation, (3.27) utilizes the well known *dot product* or *inner product* notation, whereby if $\mathbf{x}$ and $\mathbf{y}$ are m-vectors, $\mathbf{x} \cdot \mathbf{y}$ is defined:

$$\mathbf{x} \cdot \mathbf{y} = \Sigma x_j y_j. \quad (3.29)$$

Equivalently, this is the matrix product of the $1 \times m$ row matrix $\mathbf{D}(\mathbf{i}_0)$, and the $m \times 1$ column matrix $\Delta\mathbf{i}$. Also, the last term in (3.28) is expressed in *matrix product* *notation*, or more specifically, as a *quadratic form* in $\Delta\mathbf{i}$. By the above convention for $\Delta\mathbf{i}$, $\Delta\mathbf{i}^T$ is the corresponding row matrix, or *transpose* of $\Delta\mathbf{i}$. Standard matrix calculations then produce:

$$\mathbf{x}^T\mathbf{C}\mathbf{x} = \sum\sum c_{jk}x_jx_k. \qquad (3.30)$$

Note that for the smooth price functions assumed here:

$$C_{jk}(\mathbf{i}) = C_{kj}(\mathbf{i}),$$

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

$$\mathbf{C}(\mathbf{i}) = \mathbf{C}(\mathbf{i})^T. \qquad (3.32)$$

Again returning to the example in (3.8) with $\mathbf{i}_0 = (0.105, 0.10)$, the partial derivatives in (3.9) imply:

$$D_1(\mathbf{i}_0) = -1.4902,\ D_2(\mathbf{i}_0) = 1.5038, \qquad (3.33a)$$

$$C_{11}(\mathbf{i}_0) = -2.697,\ C_{22}(\mathbf{i}_0) = 4.101,\ C_{12} = C_{21} = 0. \qquad (3.33b)$$

Hence, the first-order approximation in (3.27) becomes:

$$P(\mathbf{i}_0 + \Delta\mathbf{i}) \approx 10.99136(1 + 1.4902\Delta i_1 - 1.5038\Delta i_2). \qquad (3.34)$$

Noting the functional form of (3.34), it is little wonder that for nonparallel yield curve shifts, $\Delta i_1 \neq \Delta i_2$, this price function changed in ways not anticipated by the traditional approximation (2.8). Namely, this price function is relatively sensitive to movements in Δi_1 and Δi_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 0.0136. Similarly, the traditional convexity value of 1.404 disguises the greater sensitivities implied by the partial convexities in (3.33b).

In this example, the partial durations are seen to sum to the modified duration, while the partial convexities sum to the traditional convexity value. The following proposition formalizes this result:

Proposition 5:

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 using the "parallel shift" approach. Then:

$$D(\mathbf{i}_0) = \sum D_j(\mathbf{i}_0), \qquad (3.35)$$

$$C(\mathbf{i}_0) = \sum\sum C_{jk}(\mathbf{i}_0). \qquad (3.36)$$

Proof: Let $\mathbf{M} = (1, \ldots, 1)$, the parallel shift direction vector and define the price function $P(i) = P(\mathbf{i}_0 + i\mathbf{M})$. Then:

$$P'(i) = \sum P_j(\mathbf{i}_0 + i\mathbf{M}), \qquad (3.37a)$$

$$P''(i) = \sum\sum P_{jk}(\mathbf{i}_0 + i\mathbf{M}). \qquad (3.37b)$$

Evaluating (3.37) at $i = 0$ and dividing by $P(0) = P(\mathbf{i}_0)$ completes the proof. □

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

Proposition 6:

Let $\mathbf{r}(t)$ be a smooth parametrization of yield curve vectors defined on [0, 1] so that $\mathbf{r}(0) = \mathbf{i}_0$, $\mathbf{r}(1) = \mathbf{i}_0 + \Delta\mathbf{i}$. Also, assume that $P[\mathbf{r}(t)] \neq 0$ for $0 \le t \le 1$. Then:

$$P(\mathbf{i}_0 + \Delta i)/P(\mathbf{i}_0) = \exp\left\{-\int_0^1 \mathbf{D}[\mathbf{r}(t)]\cdot\mathbf{r}'(t)dt\right\}, \qquad (3.38)$$

where $\mathbf{r}'(t)$ denotes the ordinary derivative of this vector valued function.

Proof. Define $f(t) = \ln|P[\mathbf{r}(t)]|$. A calculation shows that $f'(t) = -\mathbf{D}[\mathbf{r}(t)]\cdot\mathbf{r}'(t)$, which can be integrated and exponentiated to complete the proof. □

In the special case in which $\mathbf{r}(t)$ is linear, $\mathbf{r}(t) = \mathbf{i}_0 + t\Delta\mathbf{i}$, the more general formula in (3.38) is easily seen to reduce to the directional derivative counterpart in (3.10), with $\Delta\mathbf{i}$ here corresponding to $\Delta i\mathbf{N}$ above.

From Proposition 6, the following approximation results:

$$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0) \approx \exp[-\mathbf{D}(\mathbf{i}_0)\cdot\mathbf{r}'(0)]. \qquad (3.39)$$

To develop the second-order exponential approximation, partial derivatives of the various partial durations are required. Analogous to (3.13), we have:

$$\frac{\partial D_j}{\partial i_k} = D_kD_j - C_{jk}, \qquad (3.40)$$

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 (3.38) 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:

$$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0) \approx \exp\{-\mathbf{D}(\mathbf{i}_0)\cdot\mathbf{r}'(0) + 1/2\,[\mathbf{r}'(0)^T [\mathbf{C}(\mathbf{i}_0) - \mathbf{D}(\mathbf{i}_0)^T\mathbf{D}(\mathbf{i}_0)]\mathbf{r}'(0) - \mathbf{D}(\mathbf{i}_0)\cdot\mathbf{r}''(0)]\}. \qquad (3.41)$$

In the special case in which $\mathbf{r}(t)$ is linear, $\mathbf{r}'(t) = \Delta\mathbf{i}$ and $\mathbf{r}''(0) = \mathbf{0}$. Consequently, (3.39) and (3.41) reduce to the directional derivative counterparts in (3.11) and (3.14), respectively.

d. YTM Approach Revisited

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,

$$P(\mathbf{i}_0 + \Delta\mathbf{i}) \approx P(\mathbf{i}_0)\,[1 - \mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i}], \tag{3.42a}$$

$$P(I_0 + \Delta I) \approx P(I_0)\,[1 - D(I_0) \cdot \Delta I]. \tag{3.42b}$$

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

$$\Delta I \approx \frac{\mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i}}{D(I_0)}. \tag{3.43}$$

When $\Delta\mathbf{i}$ is a parallel shift, the numerator of (3.43) reduces to $D(\mathbf{i}_0)\Delta i$ since $D(\mathbf{i}_0) = \sum D_j(\mathbf{i}_0)$ by Proposition 5.

As an example, recall the price function (2.5) of Section 2b, where the initial yield curve, $\mathbf{i}_0 = (0.\,105, 0.10)$, was seen to be equivalent to the yield to maturity, $I_0 = 0.00445$; that is, both produced an initial price of 10.99136. Consider first the parallel yield curve shift of 0.01 exemplified there. To apply (3.43), recall that $D(I_0) = 0.172$ from (2.6), while $D(\mathbf{i}_0) = 0.0136$ from (2.8). We then obtain $\Delta I \approx 0.0008$, compared with the exact value of 0.0004. Consider next the small nonparallel shift, $\Delta\mathbf{i} = (0.0005, 0.001)$. Using (3.43) and the partial durations in (3.33), one approximates the associated change in the yield to maturity, $\Delta I \approx 0.00442$. Estimating ΔI directly proves this result to be a little understated, in that $\Delta I = 0.00455$.

By expanding the Taylor series in (3.42) to include second-order terms, ΔI can be estimated using the quadratic formula:

$$\Delta I \approx \{D - \sqrt{[D^2 - 2C\mathbf{D} \cdot \Delta\mathbf{i} + C\Delta\mathbf{i}^T C\Delta\mathbf{i}]}\}/C, \tag{3.44}$$

where $D=D(I_0)$, $C=C(I_0)$, $\mathbf{D}=\mathbf{D}(\mathbf{i}_0)$, and $\mathbf{C}=\mathbf{C}(\mathbf{i}_0)$. This formula simplifies greatly for parallel shifts since $\mathbf{D}\cdot\Delta\mathbf{i} = D(\mathbf{i}_0)\Delta i$, and $\Delta\mathbf{i}^T\mathbf{C}\Delta\mathbf{i} = C(\mathbf{i}_0)(\Delta i)^2$. In (3.44), the negative square root is chosen to satisfy the initial condition that $\Delta I=0$ when $\Delta\mathbf{i}=\mathbf{0}$.

Using (3.44), the parallel shift of 0.01 is seen to be equivalent to a YTM shift of 0.0004, which is exact to four decimal places. For the nonparallel shift, $\Delta\mathbf{i} = (0.0005, 0.001)$, the estimate for ΔI is also improved compared with the linear estimate, reproducing the exact value of $\Delta I = 0.00455$ to five decimal places. Note, however, that it is possible to obtain a negative quantity within the square root in (3.44), for example, the shift $\Delta\mathbf{i} = (0.005, 0.01)$. In such a case, there is no real number, ΔI, for which the one-variable second-order Taylor series equals the multivariate series reflecting $\Delta\mathbf{i}$, $\mathbf{D}(\mathbf{i})$, and $\mathbf{C}(\mathbf{i})$.

e. Parallel Shift Approach Revisited

Consider next the parallel shift analysis of Section 2c. Recall that it was shown that nonparallel shifts could be accommodated by redefining duration and convexity to reflect these nonparallel yield curve directions. Another interpretation is possible whereby nonparallel shifts are first translated to "equivalent parallel shifts," and the traditional Section 2a formulas are then applied. This notion is more fully explored in Section 4b and seen to provide an intuitive basis for new yield curve risk exposure measures.

To this end, the first-order expansion of $P(\mathbf{i}_0 + \Delta\mathbf{i})$ in (3.42a) must be used twice, once for the general $\Delta\mathbf{i}$ and once for the parallel shift vector, $\Delta\mathbf{i}=\Delta i\mathbf{M}$, where $\mathbf{M}=(1, \ldots, 1)$. Equating these approximations, we can solve for Δi when $D(\mathbf{i}_0) \neq 0$, obtaining:

$$\Delta i \approx \frac{\mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i}}{D(\mathbf{i}_0)}. \tag{3.45}$$

Unlike the YTM counterpart formula in (3.43), here Δi is seen to be a weighted average of the various component Δi_j values since $\sum D_j(\mathbf{i}_0)=D(\mathbf{i}_0)$.

Using the partial durations in (3.33a), we can apply (3.45) to the nonparallel shifts in (2.10), to obtain:

$\Delta\mathbf{i}$	"Equivalent" Δi
(0.0025, 0.0075)	0.5554
(0.0002, 0.0001)	–0.0109

Interpreted this way, we see that the traditional formulas can provide poor estimates for nonparallel shifts because the units of the equivalent parallel shift, Δi, can be orders of magnitude larger, and/or of a different sign, than may be inferred from the various nonparallel shift values of Δi_j. This cannot happen if all $D_j(\mathbf{i}_0)$ values have the same sign. In such a case, the equivalent Δi will be within the range of Δi_j values (Proposition 13).

A second-order counterpart to (3.45) can also be developed. A calculation shows it to be identical to (3.44), only with $D=D(\mathbf{i}_0)$ and $C=C(\mathbf{i}_0)$.

4. Additional Properties of Multivariate Models

a. Duration and Convexity Relationships

In this section, relationships between the various duration and convexity measures defined in the previous sections are investigated.

Proposition 7:

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 $\mathbf{P}(\mathbf{i}_0) \neq 0$,

$$\mathbf{D}(\mathbf{i}_0) = [P_1(\mathbf{i}_0)\mathbf{D}_1(\mathbf{i}_0) + P_2(\mathbf{i}_0)\mathbf{D}_2(\mathbf{i}_0)]/\mathbf{P}(\mathbf{i}_0), \tag{4.1}$$

$$\mathbf{C}(\mathbf{i}_0) = [P_1(\mathbf{i}_0)\mathbf{C}_1(\mathbf{i}_0) + P_2(\mathbf{i}_0)\mathbf{C}_2(\mathbf{i}_0)]/\mathbf{P}(\mathbf{i}_0). \tag{4.2}$$

Proof: As is the case for the traditional values, this result follows directly from the additive property of derivatives. □

Proposition 8:

Let $\mathbf{N} \neq \mathbf{0}$ be a direction vector. Then:

$$D_N(\mathbf{i}_0) = \mathbf{N} \cdot \mathbf{D}(\mathbf{i}_0), \tag{4.3}$$

$$C_N(\mathbf{i}_0) = \mathbf{N}^T\mathbf{C}(\mathbf{i}_0)\mathbf{N}. \tag{4.4}$$

Proof: Both formulas follow directly from (3.3) and the definitions of the various duration and convexity values. □

A simple corollary to Proposition 8 is possible concerning the "key rate" durations of Ho [12]. As noted in Section 3a, the collection of direction vectors, $\mathbf{N}_j$, form a partition of the parallel shift vector, (1, 1, ..., 1). Consequently, key rate durations sum to the traditional duration measure since:

$$\begin{aligned}\Sigma D_N(\mathbf{i}_0) &= \Sigma \mathbf{N}_j \cdot \mathbf{D}(\mathbf{i}_0)\\ &= (1, \ldots, 1) \cdot \mathbf{D}(\mathbf{i}_0)\\ &= D(\mathbf{i}_0),\end{aligned}$$

by Proposition 5.

This result has been independently derived by Ho.

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

Proposition 9:

Let $\mathbf{N} \neq \mathbf{0}$ be a direction vector. Then:

$$\frac{d}{di}D(i_0) = D^2(i_0) - C(i_0), \tag{4.5}$$

$$\frac{\partial}{\partial \mathbf{N}}D_N(\mathbf{i}_0) = D_N^2(\mathbf{i}_0) - C_N(\mathbf{i}_0), \tag{4.6}$$

$$\frac{\partial}{\partial i_j}D_k(\mathbf{i}_0) = D_j(\mathbf{i}_0)D_k(\mathbf{i}_0) - C_{jk}(\mathbf{i}_0), \tag{4.7}$$

$$\frac{\partial}{\partial i_j}D(\mathbf{i}_0) = D(\mathbf{i}_0)D_j(\mathbf{i}_0) - \sum_k C_{jk}(\mathbf{i}_0). \tag{4.8}$$

Proof: Relationship (4.5) is derived by differentiating the identity, $P'(i) = -P(i)D(i)$, solving for $D'(i)$ and substituting $i = i_0$. Similarly, (4.6) 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 that are taken.

Differentiating the identity, $P_k(\mathbf{i}) = -P(\mathbf{i})D_k(\mathbf{i})$ with respect to i_j leads to (4.7), while summing this result with respect to k and using (3.35) produces (4.8). □

Turning next 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 on $\mathbf{i}_0$. Then for all direction vectors, $\mathbf{N}$,

$$-|\mathbf{D}(\mathbf{i}_0)||\mathbf{N}| \leq D_N(\mathbf{i}_0) \leq |\mathbf{D}(\mathbf{i}_0)||\mathbf{N}|, \tag{4.9}$$

where | | denotes the length of the given vectors. Further, the upper bound in (4.9) is achieved for all positive multiples of the unit vector:

$$\mathbf{N}_0 = \mathbf{D}(\mathbf{i}_0)/|\mathbf{D}(\mathbf{i}_0)|, \tag{4.10}$$

while the lower bound is achieved for all negative multiples.

Proof: This proposition is an immediate consequence of the Cauchy-Schwarz inequality, since by Proposition 8, $D_N(\mathbf{i}_0)$ is an inner product. Specifically, the absolute value of an inner product is less than or equal to the product of the vectors' lengths, with equality if and only if the vectors are parallel. □

Note 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 $\mathbf{N}_0 = (1, 1, \ldots, 1)$.

Proposition 11 shows that given $D(\mathbf{i}_0)$, the range of price sensitivity displayed in (4.9) is minimized in this case.

Proposition 11:

Let $\mathbf{D}(\mathbf{i}_0)$ be a total duration vector with associated duration $D(\mathbf{i}_0)$. Then:

$$|\mathbf{D}(\mathbf{i}_0)| \geq |D(\mathbf{i}_0)|/\sqrt{m}, \tag{4.11}$$

where m is the dimension of $\mathbf{D}(\mathbf{i}_0)$. Further, the lower bound in (4.11) 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 noncalculus proof is possible. Changing notation, let $\mathbf{A}$ be the vector with $a_j = D(\mathbf{i}_0)/m$, for all j, and let $\mathbf{B}$ also have the property that $\sum b_j = D(\mathbf{i}_0)$. Then $\mathbf{C} = \mathbf{B} - \mathbf{A}$ satisfies $\sum c_i = 0$, so $|\mathbf{B}|^2 = |\mathbf{A}|^2 + |\mathbf{C}|^2$. Hence, since $|\mathbf{C}|^2 \geq 0$, $|\mathbf{B}|^2$ is minimized when $\mathbf{C} = \mathbf{0}$. □

Bounds for directional convexities are considered next. While the following result and proof reflect known extremal properties of quadratic forms and use well-known techniques, they are included here for completeness.

Proposition 12:

Let $P(\mathbf{i})$ be a price function and $\mathbf{C}(\mathbf{i}_0)$ its total convexity matrix evaluated on $\mathbf{i}_0$. Then:

$$\lambda_1|\mathbf{N}|^2 \leq C_N(\mathbf{i}_0) \leq \lambda_m|\mathbf{N}|^2, \tag{4.12}$$

where λ_1 and λ_m are the smallest and largest eigenvalues of $\mathbf{C}(\mathbf{i}_0)$, respectively. Further, the bounds in (4.12) are achieved for all multiples of the associated eigenvectors, $\mathbf{N}_1$ and $\mathbf{N}_m$.

Proof: From (4.4), it is clear that:

$$C_{aN}(\mathbf{i}_0) = a^2C_N(\mathbf{i}_0), \tag{4.13}$$

and hence (4.12) need only be established for $|\mathbf{N}| = 1$. By (3.32), $\mathbf{C}(\mathbf{i}_0)$ is a symmetric matrix, so all eigenvalues are real numbers. In addition, $\mathbf{C}(\mathbf{i}_0)$ must have m independent unit eigenvectors, $\mathbf{N}_1, \ldots, \mathbf{N}_m$, which are mutually orthogonal and in which basis $\mathbf{C}(\mathbf{i}_0)$ is a diagonal matrix.

Let $\mathbf{P}$ be the change of basis matrix with the $\mathbf{N}_j$ as column vectors. For convenience, we enumerate the eigenvectors so that $\mathbf{N}_1$ is associated with the smallest eigenvalue, and $\mathbf{N}_m$ the largest. Because the columns of $\mathbf{P}$ are mutually orthogonal, $\mathbf{P}^{-1} = \mathbf{P}^T$, where $\mathbf{P}^T$ is the transpose of $\mathbf{P}$.

Changing coordinates, let $\mathbf{N} = \mathbf{Px}$, so the components of $\mathbf{x}$ equal the coordinates of $\mathbf{N}$ in the $\{\mathbf{N}_j\}$ basis. From (4.4), we obtain by substitution, recalling that $(\mathbf{Px})^T = \mathbf{x}^T\mathbf{P}^T$:

$$C_N(\mathbf{i}_0) = \mathbf{x}^T\mathbf{P}^T\mathbf{C}(\mathbf{i}_0)\mathbf{Px} = \sum_{i=1}^{m} \lambda_i x_i^2, \tag{4.14}$$

since $\mathbf{P}^T\mathbf{CP}$ is diagonal as noted above. In addition, expressing $|\mathbf{N}|^2$ as $\mathbf{N}^T\mathbf{N}$, the constraint $|\mathbf{N}|^2 - 1 = 0$ becomes:

$$|\mathbf{N}|^2 - 1 = \mathbf{x}^T\mathbf{P}^T\mathbf{Px} - 1 = \sum_{i=1}^{m} x_i^2 - 1 = 0. \tag{4.15}$$

Substituting $x_1^2 = 1 - \sum_{i=2}^{m} x_i^2$ into (4.14), we obtain:

$$C_N(\mathbf{i}_0) = \lambda_1 + \sum_{i=2}^{m} (\lambda_i - \lambda_1)x_i^2. \tag{4.16}$$

Because the summation in (4.16) is non-negative, the minimum $C_N(\mathbf{i}_0)$ is obtained when $x_i = 0$ for $i \geq 2$, and $x_1 = 1$. That is, $C_N(\mathbf{i}_0)$ has minimum value λ_1, when $\mathbf{x} = (1, 0, \ldots, 0)$, and hence $\mathbf{N} = \mathbf{Px} = \mathbf{N}_1$.

Substituting $x_m^2 = 1 - \sum_{i=1}^{m-1} x_i^2$, an identical argument completes the proof. □

From Proposition 12, it is clear that the directional convexities of a price function need not have the same sign. In particular, all $C_N(\mathbf{i}_0)$ will be positive only when all λ_j are positive, that is, only when $\mathbf{C}(\mathbf{i}_0)$ is a positive definite matrix. Similarly, all $C_N(\mathbf{i}_0)$ will be negative only when $\mathbf{C}(\mathbf{i}_0)$ is a negative definite matrix. In general, $C_N(\mathbf{i}_0)$ will take on both signs for different values of $\mathbf{N}$.

The simple example in (3.8) has directional convexities of both signs. By (3.33), $\mathbf{C}(\mathbf{i}_0)$ is a diagonal matrix. Consequently, its eigenvalues equal the respective diagonal elements, and we have by (4.12):

$$-2.697|\mathbf{N}|^2 \leq C_N(\mathbf{i}_0) \leq 4.101|\mathbf{N}|^2, \tag{4.17}$$

with corresponding unit eigenvectors: $\mathbf{N}_1 = (1,0)$ and $\mathbf{N}_m = \mathbf{N}_2 = (0,1)$.

This observation concerning the sign of $C_N(\mathbf{i}_0)$ is important because it is often tacitly assumed that "positive convexity," or $C_N(\mathbf{i}_0) > 0$ when $\mathbf{N} = (1, \ldots, 1)$, is always good, and more is always better. See Reitano [22] for a more detailed analysis of this issue.

A fast way to estimate the potential size of the interval in (4.12) is to calculate the "norm" of the total convexity matrix, $|\mathbf{C}(\mathbf{i}_0)|$, using any submultiplicative norm. This is because $|\lambda_j| \leq |\mathbf{C}(\mathbf{i}_0)|$ for all eigenvalues λ_j. Consequently, (4.12) can be rewritten:

$$|C_N(\mathbf{i}_0)| \leq |\mathbf{C}(\mathbf{i}_0)||\mathbf{N}|^2 . \tag{4.12$'$}$$

Though not a sharp estimate like that produced by the interval in (4.12), the above interval is easily calculated. For example, one possible norm is:

$$|\mathbf{C}(\mathbf{i}_0)| = \max_j \sum_i |C_{ij}(\mathbf{i}_0)| .$$

For the above example, we see from (3.33) that $|\mathbf{C}(\mathbf{i}_0)| = 4.101$ using this norm, and (4.12)′ simply symmetrizes the interval in (4.17). In general, however, the estimates may differ significantly, especially when (4.12) is highly asymmetric.

b. Durational Leverage and the Durational Multiplier

In Section 3e above, the notion of an equivalent parallel shift was introduced in (3.45). Here, we formalize this concept and investigate its properties.

Definition 4.1:

Let $P(\mathbf{i})$ be a price function and $\mathbf{i}_0$ a yield curve vector so that $D(\mathbf{i}_0) \neq 0$. For a yield curve shift $\Delta\mathbf{i}$, the *equivalent parallel shift*, $\Delta\mathbf{i}^E$, is defined:

$$\Delta i^E = \frac{\mathbf{D}(\mathbf{i}_0) \cdot \Delta\mathbf{i}}{D(\mathbf{i}_0)} . \square \tag{4.18}$$

Clearly, Δi^E is a function of both $\mathbf{i}_0$ and $\Delta\mathbf{i}$, though for notational convenience, this dependence will usually be suppressed. The relationship between Δi^E and the length of $\Delta\mathbf{i}$ is of immediate importance. As noted in Section 3e, we have the following:

Proposition 13:

Assume $D(\mathbf{i}_0) \neq 0$ and all $D_j(\mathbf{i}_0)$ have the same sign. Then for all $\Delta\mathbf{i}$:

$$\min(\Delta i_j) \leq \Delta i^E \leq \max(\Delta i_j). \tag{4.19}$$

Proof: By (4.18), $\Delta i^E = \sum\lambda_j \Delta i_j$ where $\sum\lambda_j = 1$. By assumption, all λ_j satisfy $0 \leq \lambda_j \leq 1$, implying (4.19). □

In the more general case, the relationship between Δi^E and $\Delta\mathbf{i}$ is somewhat more complicated. To this end, we have:

Deftnition 4.2:

Given $\mathbf{i}_0$ and $\Delta\mathbf{i}$, the *directional leverage of* $P(\mathbf{i})$ in the direction of $\Delta\mathbf{i}$, denoted $L(\Delta\mathbf{i})$, is defined:

$$L(\Delta\mathbf{i}) = \frac{\Delta i^E}{|\Delta\mathbf{i}|} . \tag{4.20}$$

The *durational leverage of* $P(\mathbf{i})$ at $\mathbf{i}_0$, denoted $L(\mathbf{i}_0)$, is defined:

$$L(\mathbf{i}_0) = \max L(\Delta\mathbf{i}). \square \tag{4.21}$$

As for Δi^E, the dependence of $L(\Delta\mathbf{i})$ on $\mathbf{i}_0$ will usually be suppressed. From Definition 4.1, we see that $L(\Delta\mathbf{i})$ is truly a function of direction alone, since for any $\lambda>0$, $L(\lambda\Delta\mathbf{i})=L(\Delta\mathbf{i})$. Consequently, $L(\Delta\mathbf{i})$ achieves all its values on the unit sphere, $|\Delta\mathbf{i}| = 1$. Since $L(\Delta\mathbf{i})$ is clearly a continuous function, it attains a maximum on this sphere and $L(\mathbf{i}_0)$ is consequently well defined. Because $L(\Delta\mathbf{i})$ is an odd function, that is, $L(-\Delta\mathbf{i}) = -L(\Delta\mathbf{i})$, we have that:

$$-L(\mathbf{i}_0)|\Delta\mathbf{i}| \leq \Delta i^E \leq L(\mathbf{i}_0)|\Delta\mathbf{i}| . \tag{4.22}$$

Proposition 14:

Given the definition above, we have:

$$-\left|\frac{\mathbf{D}(\mathbf{i}_0)}{D(\mathbf{i}_0)}\right| \leq L(\Delta\mathbf{i}) \leq \left|\frac{\mathbf{D}(\mathbf{i}_0)}{D(\mathbf{i}_0)}\right| . \tag{4.23}$$

Further, the upper bound in (4.23) is achieved if and only if $\Delta\mathbf{i} = c\mathbf{D}(\mathbf{i}_0)$, where sign c = sign $D(\mathbf{i}_0)$.

Proof: This result is an immediate consequence of (4.18) and Proposition 10, since $\mathbf{D}(\mathbf{i}_0) \times \Delta\mathbf{i} = D_{\Delta i}(\mathbf{i}_0)$ by Proposition 8. □

Corollary:

$$L(\mathbf{i}_0) = \left|\frac{\mathbf{D}(\mathbf{i}_0)}{D(\mathbf{i}_0)}\right| . \square$$

From the above analysis we see that the total duration vector $\mathbf{D}(\mathbf{i}_0)$ provides the direction in which $L(\Delta\mathbf{i})$ is maximized. Further, its length, in units of $D(\mathbf{i}_0)$, quantifies the relationship between Δi^E and $|\Delta\mathbf{i}|$. Consequently, if $|\mathbf{D}(\mathbf{i}_0)|$ is large relative to $|D(\mathbf{i}_0)|$; that is, if $L(\mathbf{i}_0)$ is large, even small nonparallel shifts have the potential to produce large equivalent parallel shifts and hence large changes in price.

Proposition 15:

For any price function, $P(\mathbf{i})$,

$$L(\mathbf{i}_0) \geq 1/\sqrt{m}, \tag{4.24}$$

with equality if and only if $D_j(\mathbf{i}_0) = D(\mathbf{i}_0)/m$ for all j.

Further, if all $D_j(\mathbf{i}_0)$ have the same sign,

$$L(\mathbf{i}_0) \leq 1. \tag{4.25}$$

Proof: Inequality (4.24) follows from the above corollary and Proposition 11. For (4.25), note that:

$$L(\mathbf{i}_0)^2 = \sum D_j^2/(\sum D_j)^2$$
$$= \sum D_j^2/\left(\sum D_j^2 + 2\sum_{i<j} D_i D_j\right),$$

which is clearly less than or equal to 1 if all D_j have the same sign. □

For the example in (3.8), we have from (3.33a) that $L(\mathbf{i}_0) = 155.7$. That is, given any restriction on $|\Delta\mathbf{i}|$, one can find yield curve shifts of that length so that $\Delta i^E = \pm 155.7|\Delta\mathbf{i}|$. By Proposition 14, all such critical shifts are proportional to $\mathbf{D}(\mathbf{i}_0) = (-1.4902, 1.5038)$. For example, the shift $\Delta\mathbf{i} = (-0.00070, 0.00071)$ has a length equal to about 10 bp, with $\Delta i^E = 0.155$. Changing the signs in $\Delta\mathbf{i}$ produces $\Delta i^E = -0.155$.

The leverage concept above has intuitive appeal, because it provides a method of relating the sizes of nonparallel shifts with those of the corresponding equivalent parallel shifts. The basis of this correspondence is that the durational effect in (2.4) and (3.27) is the same for each shift. Note, however, that the units used to measure the shifts are different. For $\Delta\mathbf{i}$, the unit basis is vector length, $|\Delta\mathbf{i}|$, while for Δi^E, the unit basis equals the amount of the parallel displacement. In particular, if $\Delta\mathbf{i}^E$ is the parallel shift vector corresponding to Δi^E, we have $|\Delta\mathbf{i}^E| = \sqrt{m}|\Delta i^E|$. This difference in units causes the value of $L(\mathbf{i}_0)$ and the inequalities in (4.22) to disguise somewhat the potential for yield curve risk.

We proceed to quantify yield curve risk in a manner that overcomes this difference in units. Given a yield curve shift $\Delta\mathbf{i}$, we seek a relationship between its durational effect and that produced by a parallel shift of the same length and orientation. By "orientation," we mean as given by the sign of Δi^E. So if $\Delta i^E>0$, we compare the durational effect of $\Delta\mathbf{i}$ to that of a positive parallel shift of the same length, and conversely.

To this end, the durational effect of $\Delta\mathbf{i}$ is $\mathbf{D}(\mathbf{i}_0)\cdot\Delta\mathbf{i}$, while the durational effect of the parallel shift of the same length and orientation is $\pm D(\mathbf{i}_0)|\Delta\mathbf{i}|/\sqrt{m}$. Here, we choose the sign consistent with the sign of Δi^E. The "directional multiplier" is defined as the ratio of these durational effects. By the above orientation convention, this ratio is always positive, so absolute values are used to simplify notation.

Definition 4.3:

Let $P(\mathbf{i})$ be a price function and $\mathbf{i}_0$ a yield vector so that $D(\mathbf{i}_0)\neq 0$. For a yield curve shift $\Delta\mathbf{i}$, the *directional multiplier of* $P(\mathbf{i})$ *in the direction of* $\Delta\mathbf{i}$, denoted $M(\Delta\mathbf{i})$, is defined:

$$M(\Delta\mathbf{i}) = \frac{\sqrt{m}|\mathbf{D}(\mathbf{i}_0)\cdot\Delta\mathbf{i}|}{|\mathbf{D}(\mathbf{i}_0)||\Delta\mathbf{i}|}. \tag{4.26}$$

The *durational multiplier*, denoted $M(\mathbf{i}_0)$, is defined:

$$M(\mathbf{i}_0) = \max M(\Delta\mathbf{i}). \;\square$$

As was the case for $L(\Delta\mathbf{i})$, $M(\Delta\mathbf{i})$ is a function of direction alone since $M(\lambda\Delta\mathbf{i})=M(\Delta\mathbf{i})$ for $\lambda>0$. Moreover, $M(\Delta\mathbf{i})$ is an even function in that $M(-\Delta\mathbf{i})=M(\Delta\mathbf{i})$. Consequently, $M(\mathbf{i}_0)$ is well defined, though this maximum is achieved at two points. In addition, note that $M(\Delta\mathbf{i}) = \sqrt{m}|L(\Delta\mathbf{i})|$, and so $M(\mathbf{i}_0) = \sqrt{m}L(\mathbf{i}_0)$. Consequently, the above propositions apply immediately to $M(\Delta\mathbf{i})$.

Also, note that:

$$M(\Delta\mathbf{i}) = |\Delta\mathbf{i}^E|/|\Delta\mathbf{i}|, \tag{4.28}$$

where $\Delta\mathbf{i}^E$ is the vector corresponding to Δi^E.

For the example in (3.8), we have $M(\mathbf{i}_0) = 220.2$. That is, the durational effect of a yield curve shift can be 220 times greater than the effect of a parallel shift of the same length and orientation. By Proposition 14, this multiplier is realized when $\Delta\mathbf{i}$ equals any multiple by $\mathbf{D}(\mathbf{i}_0)$.

In addition to providing intuitive measures of yield curve exposure, $L(\Delta\mathbf{i})$ and $M(\Delta\mathbf{i})$ can be used to quantify an effective duration measure. To this end, let $\Delta\mathbf{i}$ be given, and let Δi equal the value of the parallel shift of the same length and orientation. As noted above:

$$\Delta i = \text{sign}\,(\Delta i^E)\;|\Delta\mathbf{i}|/\sqrt{m}. \tag{4.29}$$

From (3.27), we have:

$$P(\mathbf{i}_0+\Delta\mathbf{i})/P(\mathbf{i}_0) \approx 1 - L(\Delta\mathbf{i})\, D(\mathbf{i}_0)\, |\Delta\mathbf{i}|\,. \qquad (4.30)$$

Consequently, $L(\Delta\mathbf{i})D(\mathbf{i}_0)$ quantifies an effective duration measure in units of $|\Delta\mathbf{i}|$, while $L(\mathbf{i}_0)D(\mathbf{i}_0)$ equals the maximum effective duration in these units. Equivalently,

$$P(\mathbf{i}_0+\Delta\mathbf{i})/P(\mathbf{i}_0) \approx 1 - M(\Delta\mathbf{i})\, D(\mathbf{i}_0)\Delta i, \qquad (4.31)$$

where Δi is given by (4.29). $M(\Delta\mathbf{i})D(\mathbf{i}_0)$ quantifies an effective duration measure in units of parallel shifts Δi, while $M(\mathbf{i}_0)D(\mathbf{i}_0)$ equals its maximum value.

In practice, (4.31) is easier and more intuitive to use because it is a straightforward generalization of (2.2). This is because $M(\Delta\mathbf{i})=1$ for parallel shifts by (4.28). Also, because $M(\Delta\mathbf{i})>0$ by definition, this effective duration measure has the same sign as $D(\mathbf{i}_0)$, reflecting only the muliplier effect of nonparallel shifts of the same length and orientation as $\Delta\mathbf{i}$. In this light, $M(\mathbf{i}_0)$ is indeed a durational multiplier in that, in units of parallel shifts Δi, the effective duration can be as great as $M(\mathbf{i}_0)D(\mathbf{i}_0)$. Consequently, $M(\mathbf{i}_0)D(\mathbf{i}_0)$ can be viewed as a proxy for potential yield curve risk.

c. Compound Duration Functions

In this section, the concept of the duration of duration is defined and used to restate the second-order approximations in an intuitively natural way.

Definition 4.4:

Given a directional duration function $D_N(\mathbf{i})$, the *compound directional duration*, $D_ND_N(\mathbf{i})$, is defined for $D_N(\mathbf{i}) \neq 0$ as follows:

$$D_ND_N(\mathbf{i}) = \frac{\partial D_N}{\partial \mathbf{N}} / D_N(\mathbf{i})\,. \qquad (4.32)$$

When $\mathbf{N} = (1, 1, \ldots, 1)$, the parallel shift vector, this compound duration is called the *duration of duration* and denoted $DD(\mathbf{i})$. □

Definition 4.5:

Given a partial duration function, $D_k(\mathbf{i})$, the *compound jk-th partial duration*, $D_jD_k(\mathbf{i})$, is defined for $D_k(\mathbf{i}) \neq 0$ as follows:

$$D_jD_k(\mathbf{i}) = -\frac{\partial D_k}{\partial i_j} / D_k(\mathbf{i})\,. \;\square \qquad (4.33)$$

From Proposition 9:

$$DD(\mathbf{i}) = C(\mathbf{i})/D(\mathbf{i}) - D(\mathbf{i}), \qquad (4.34)$$

$$D_ND_N(\mathbf{i}) = C_N(\mathbf{i})/D_N(\mathbf{i}) - D_N(\mathbf{i}), \qquad (4.35)$$

$$D_jD_k(\mathbf{i}) = C_{jk}(\mathbf{i})/D_k(\mathbf{i}) - D_j(\mathbf{i}). \qquad (4.36)$$

Substituting the first-order Taylor series approximation:

$$D_N(\mathbf{i}_0 + t\mathbf{N}) \approx D_N(\mathbf{i}_0)\, [1 - D_ND_N(\mathbf{i}_0)t] \qquad (4.37)$$

into the exponential identity (3.10) and integrating with respect to t produces:

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx \exp\{-\Delta i D_N(\mathbf{i}_0)\, [1 - D_ND_N(\mathbf{i}_0)\, \Delta i/2]\}. \qquad (4.38)$$

A simple calculation shows that (4.38) is equivalent to the second-order exponential approximation in (3.14). Note, however, that this approximation can be interpreted as the corresponding first-order approximation in (3.11) with an adjusted directional duration value. The adjustment corresponds to a yield change of $\Delta i/2$ and resembles the classical linear duration approximation (2.2), using $D_ND_N(\mathbf{i}_0)$. In particular, from (4.37) this adjusted directional duration equals an approximation for $D_N(\mathbf{i}_0 + \mathbf{N}\Delta i/2)$.

For example, consider the price function in (2.7) and the parallel shift of 0.01 in (2.10a). Letting $\mathbf{N} = (1,1)$, we have from (2.11a) that $D_N(\mathbf{i}_0)=0.0136$, and $D_ND_N(\mathbf{i}_0)=103.2$. For $\Delta i = 0.01$, the adjusted duration equals 0.0066, which when used in (4.38) reproduces the second-order estimate in (2.13). For the nonparallel shifts, $\mathbf{N}=(1,3)$ and $\mathbf{N}=(2,1)$, the corresponding values of $D_ND_N(\mathbf{i}_0)$ are easily calculated to be 8.3 and 6.0, respectively.

By definition, the second-order approximation in (3.7) can also be restated:

$$P(\mathbf{i}_0 + \Delta i\mathbf{N})/P(\mathbf{i}_0) \approx 1 - \Delta i D_N(\mathbf{i}_0) \times \{1 - [D_ND_N(\mathbf{i}_0) + D_N(\mathbf{i}_0)]\, \Delta i/2\} \qquad (4.39)$$

Again, this approximation utilizes an adjusted duration value, where the adjustment reflects (2.2). Here, however, $D_ND_N(\mathbf{i}_0) + D_N(\mathbf{i}_0)$ or $C_N(\mathbf{i}_0)/D_N(\mathbf{i}_0)$ is the adjusting factor.

For the partial duration counterparts, the approximation:

$$D_k(\mathbf{i}_0 + t\Delta\mathbf{i}) \approx D_k(\mathbf{i}_0)\, [1 - t \sum_j D_jD_k(\mathbf{i}_0)\, \Delta i_j], \qquad (4.40)$$

can be substituted into the exponential identity (3.38), with $\mathbf{r}(t)=\mathbf{i}_0+t\Delta\mathbf{i}$, and integrated to obtain:

$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0)$

$$\approx \exp\left\{-\sum_k \Delta i_k D_k(\mathbf{i}_0)\left[1 - \sum_j D_j D_k(\mathbf{i}_0)\Delta i_j/2\right]\right\}. \quad (4.41)$$

This exponential approximation is equivalent to (3.40) with $\mathbf{r}(t)=\mathbf{i}_0+t\Delta\mathbf{i}$. By definition, the second-order approximation in (3.28) can also be restated:

$$P(\mathbf{i}_0 + \Delta\mathbf{i})/P(\mathbf{i}_0) \approx 1 - \sum_k \Delta i_k\ D_k(\mathbf{i}_0) \times \left\{1 - \sum_j [D_j D_k(\mathbf{i}_0) + D_j(\mathbf{i}_0)]\Delta i_j/2\right\}. \quad (4.42)$$

5. Applications

a. Partial Duration and Convexity Estimates

In general, the various derivative-based definitions can be applied directly only when cash flows are fixed and independent of interest rates, and when the yield vector used reflects the corresponding spot rates. For example, assume a fixed vector of annual cash flows, $\mathbf{K} = (c_1, \ldots, c_m)$, and the associated spot rate vector, $\mathbf{i} = (i_1, \ldots, i_m)$. Naturally, the price function is given by:

$$P(\mathbf{i}) = \sum c_j v_j^j, \quad (5.1)$$

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

$$D_j(\mathbf{i}) = \frac{j c_j v_j^{j+1}}{P(\mathbf{i})}, \quad (5.2)$$

$$C_{jj}(\mathbf{i}) = \frac{j(j+1)c_j v_j^{j+2}}{P(\mathbf{i})}, \quad C_{jk}(\mathbf{i}) = 0, \quad j \neq k. \quad (5.3)$$

These partial durations clearly sum to the modified duration, and the partial convexities sum to the traditional convexity value. In addition, because $\mathbf{C}(\mathbf{i})$ is a diagonal matrix, the second-order formulas simplify. For example, (3.28) reduces to:

$$P(\mathbf{i} + \Delta\mathbf{i})/P(\mathbf{i}) = 1 - \Sigma D_j(\mathbf{i})\Delta i_j + 1/2\Sigma C_{jj}(\mathbf{i})(\Delta i_j)^2. \quad (5.4)$$

In the real world, however, many financial models contain options that make cash flows interest-sensitive. Assets can be prepaid (that is, "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, and so on often contain put options (that is, for withdrawal) and call options (that is, for additional investment). In addition, complex portfolios typically reflect hundreds of spot rates, potentially requiring hundreds of partial durations and convexities. The total duration vectors therefore are quite large, contain generally very small values, and provide little insight on the portfolio's yield curve sensitivities.

For interest-sensitive 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 modeled directly in closed mathematical form, precluding differentiation. Rather, "option pricing" models are commonly used ([5], [7], [8], [11]). With them, $P(i)$ and $P(\mathbf{i})$ are not defined directly in terms of discounted cash flows, but are defined indirectly in a manner that reflects the effect of options on the value of the cash-flow stream. Such option-pricing models produce a price that is very much a function of the yield curve assumed, and the price function can therefore be discretely estimated.

While the spot rate basis is workable, it often produces large numbers of very small partial duration and convexity estimates. A preferable approach is to "group" yield curve sensitivity into a smaller number of yield points, producing more meaningful estimates. A natural basis for this is the observed yield curve drivers on a typical bond yield curve. Such a curve may reflect yields at maturities 0.25, 0.5, 1, 2, 3, 4, 5, 7, 10, 20, and 30 years, for example. From these yields, other values are interpolated before this yield curve is transformed into the corresponding spot rate curve, which is then used as input to an option-pricing model or used directly for discounting fixed cash flows. Consequently, all yield curve sensitivities emanate from these basic ten or so variables, and this is the basis recommended for use as the yield curve vector.

By using such a yield curve basis to model $P(\mathbf{i})$ and an option-pricing model or direct calculation, $D_N(\mathbf{i}_0)$ and $C_N(\mathbf{i}_0)$ can be estimated discretely by central difference formulas:

$$D_N^{\varepsilon}(\mathbf{i}_0) = -[P(\mathbf{i}_0 + \varepsilon\mathbf{N}) - P(\mathbf{i}_0 - \varepsilon\mathbf{N})]/2\varepsilon P(\mathbf{i}_0), \quad (5.5)$$

$$C_N^{\varepsilon}(\mathbf{i}_0) = [P(\mathbf{i}_0 + \varepsilon\mathbf{N}) - 2P(\mathbf{i}_0) + P(\mathbf{i}_0 - \varepsilon\mathbf{N})]/\varepsilon^2 P(\mathbf{i}_0). \quad (5.6)$$

Forward difference formulas are also common, though they tend to be "biased" in that they better reflect sensitivity to an increase in interest rates.

To estimate ε, one commonly uses judgment and some trial and error. Theoretically, the error in these estimates can be displayed by expanding $P(\mathbf{i}_0 + \varepsilon\mathbf{N})$ and $P(\mathbf{i}_0 - \varepsilon\mathbf{N})$ into Taylor series in ε and substituting into the respective formulas. This produces:

$$D_N^\varepsilon(\mathbf{i}_0) - D_N(\mathbf{i}_0) = -P_N^{(3)}(\mathbf{i}_0)\,\varepsilon^2/6P(\mathbf{i}_0) + O(\varepsilon^4), \tag{5.7}$$

$$C_N^\varepsilon(\mathbf{i}_0) - C_N(\mathbf{i}_0) = P_N^{(4)}(\mathbf{i}_0)\,\varepsilon^2/12P(\mathbf{i}_0) + O(\varepsilon^4). \tag{5.8}$$

As can be seen from these formulas, the duration and convexity estimates improve quickly as ε decreases. However, the third and fourth directional derivatives of $P(\mathbf{i}_0)$ are generally not known, so the direct application of (5.7) and (5.8) to select an ε with a given error tolerance is not practical. Logically, an ε is desired that makes $D_N^\varepsilon(\mathbf{i})$ close to $D_N(\mathbf{i})$ in the sense that using $\varepsilon/2$, say, improves the estimate little. In practice, good results can often be obtained with ε equal to 5 to 10 basis points, when $|\mathbf{N}|$ equals the length of the parallel shift vector $(1, \ldots, 1)$.

Alternatively, to calculate the various directional derivatives and convexities, it is sufficient to estimate only the partial duration and convexity values by Proposition 8. The above formulas generalize to:

$$D_j^\varepsilon(\mathbf{i}_0) = -[P(\mathbf{i}_0 + \varepsilon_j) - P(\mathbf{i}_0 - \varepsilon_j)]/2\varepsilon_j P(\mathbf{i}_0), \tag{5.9}$$

$$C_{jk}^\varepsilon(\mathbf{i}_0) = [P(\mathbf{i}_0 + \varepsilon_j + \varepsilon_k) - P(\mathbf{i}_0 - \varepsilon_j + \varepsilon_k) - P(\mathbf{i}_0 + \varepsilon_j - \varepsilon_k) + P(\mathbf{i}_0 - \varepsilon_j - \varepsilon_k)]/4\varepsilon_j\varepsilon_k P(\mathbf{i}_0). \tag{5.10}$$

Here, $\varepsilon_j = \varepsilon_j(0, \ldots, 1, \ldots, 0)$, where ε_j is the j-th coordinate, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$. As was true for the one-variable model, judgment and trial and error are needed to determine an appropriate set of values for ε_j, which could be chosen to be equal for simplicity. Error estimation formulas generalizing (5.7) and (5.8) can again be developed by using multivariate Taylor series expansions, to produce:

$$D_j(\mathbf{i}_0) - D_j(\mathbf{i}_0) = -\,P_j^{(3)}(\mathbf{i}_0)\,\varepsilon_j^2/6P(\mathbf{i}_0) + O(\varepsilon_j^4) \tag{5.11}$$

$$C_{jk}^\varepsilon(\mathbf{i}_0) - C_{jk}(\mathbf{i}_0) = [\varepsilon_j^2 P_{jk}^{(3,1)}(\mathbf{i}_0) + \varepsilon_k^2 P_{jk}^{(1,3)}(\mathbf{i}_0)]/6P(\mathbf{i}_0) + O(\varepsilon_j, \varepsilon_k)^4. \tag{5.12}$$

In (5.11), $P_j^{(3)}$ denotes the third partial derivative with respect to i_j, while in (5.12), the (3, 1) and (1, 3) notation denotes the corresponding mixed fourth-order partial derivatives with respect to j and k. In practice, 5 to 10 basis points will often suffice.

Given m yield points, $2m+1$ price calculations are required for the partial durations in (5.9), and $2m(m-1)$ additional calculations are needed for the partial convexities in (5.10), totalling $2m^2+1$ price calculations in all. Here we assume that $C_{jj}(\mathbf{i}_0)$ in (5.10) is estimated with $\varepsilon_j/2$ when ε_j is used for (5.9).

If desired, the total number of calculations can be reduced by almost half, to m^2+m+1, in the following way. Let $\mathbf{N}_j=\varepsilon_j$ above and $\mathbf{N}_{jk}=\varepsilon_{jk}(0, \ldots, 1, \ldots, 0, 1, \ldots, 0)$, with $j<k$ and $\mathbf{N}_{jk}$ non-zero in the j-th and k-th components. Using the $\mathbf{N}_j$ vectors, $D_j(\mathbf{i}_0)$ and $C_{jj}(\mathbf{i}_0)$ can be estimated as in (5.5) and (5.6) with $\varepsilon=1$ and a total of $2m+1$ price calculations. This is equivalent to the above estimates with (5.9) and (5.10). With an additional $m(m-1)$ calculations and (5.6), $C_N(\mathbf{i}_0)$ can be estimated for each $\mathbf{N}_{jk}$. Using (4.4), we then obtain:

$$C_{jk}(\mathbf{i}_0) = 1/2[C_N(\mathbf{i}_0) - C_{jj}(\mathbf{i}_0) - C_{kk}(\mathbf{i}_0)], \tag{5.13}$$

where $\mathbf{N}=\mathbf{N}_{jk}$. Also, by (3.31), $C_{kj}(\mathbf{i}_0)=C_{jk}(\mathbf{i}_0)$. Consequently, the total number of price calculations needed is m^2+m+1.

As a final comment, note that the partial duration and convexity estimates above should be "normalized" to satisfy Proposition 5. That is, these values should be scaled so that they sum to the estimated duration or convexity values, respectively. In practice, relative discrepancies are typically well under 1 percent before scaling.

b. Price Sensitivity—Direct Yield Curve Approach

Once the partial durations have been calculated, the first exercise is one of observation. Because modified duration equals the sum of the partial durations, one can observe to what extent parallel price sensitivity as measured by $D(\mathbf{i}_0)$ decomposes along the yield curve. In general, price sensitivity to nonparallel shifts is greater if the partial durations are large, with some positive and others negative, rather than relatively uniform of size $D(\mathbf{i}_0)/m$.

Beyond this informal exercise of observation, price sensitivity can be calculated 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, $\mathbf{N}$, the directional duration $D_N(\mathbf{i}_0)$ can be calculated from (4.3). This value then reflects price sensitivity to yield curve shifts that are proportional to $\mathbf{N}$.

One direction vector of note is $\mathbf{N}_0$ as defined in (4.10). Recall that $\mathbf{N}_0$ was parallel to $\mathbf{D}(\mathbf{i}_0)$, only with unit length. As demonstrated in Proposition 10, this vector represents the yield curve shift that produces the maximum value of $D_N(\mathbf{i}_0)$ and, consequently, the greatest relative change in the price function given $|\mathbf{N}|=1$. Similarly, yield curve shifts proportional to $\mathbf{N}_0$ also provide extreme values of $D_N(\mathbf{i}_0)$ and hence represent yield curve directions of maximal relative price sensitivity. By Proposition 10, the length of the total duration vector, $|\mathbf{D}(\mathbf{i}_0)|$, quantifies the amount of this maximal relative price sensitivity.

Another notion of interest is the directional leverage function, $L(\Delta\mathbf{i})$, and in particular, its maximum value, $L(\mathbf{i}_0)$, the durational leverage. This latter value quantifies the maximum value of the equivalent parallel shift, Δi^E, given any restriction on $|\Delta\mathbf{i}|$, the length of the original shift. As noted in Section 4b, $L(\mathbf{i}_0)$ equals the ratio of $|\mathbf{D}(\mathbf{i}_0)|$ to $|D(\mathbf{i}_0)|$, and this maximum is achieved when $\Delta\mathbf{i}$ is parallel to $\mathbf{D}(\mathbf{i}_0)$.

A final related notion of interest is the directional multiplier function, $M(\Delta\mathbf{i})$, and in particular, its maximum value, $M(\mathbf{i}_0)$, the durational multiplier. This latter value provides a simple quantitative measure of yield curve risk. In particular, the durational effect of a nonparallel yield curve shift can be $M(\mathbf{i}_0)$ times greater than for a parallel shift of the same length and orientation. That is, the effective portfolio duration can be as large as $M(\mathbf{i}_0)D(\mathbf{i}_0)$. As was the case for $L(\mathbf{i}_0)$, the direction in which $M(\Delta\mathbf{i})$ is maximized is parallel to $\mathbf{D}(\mathbf{i}_0)$.

Given any of these yield curve risk measures, $|\mathbf{D}(\mathbf{i}_0)|$, $L(\mathbf{i}_0)$, or $M(\mathbf{i}_0)$, it is clear from Propositions 11 and 15 that risk will be lessened if the partial durations are of uniform size, rather than both positive and negative. In particular, all these measures are minimized if the partial durations are equal, and none can be too great if the partial durations are at least of the same sign. In this regard, "barbell" and "reverse barbell" duration matching strategies can be quite risky, because the resultant partial durations often are large, with some positive and others negative. Correspondingly, the above risk measures also tend to be large.

c. Price Sensitivity-Yield Curve Slope Approach

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

$$l_j = am_j + b, \qquad (5.14)$$

where m_j denotes the maturity value of the pivotal yield curve point i_j. For example, one might have $m_1 = 0.25$, $m_2 = 0.5$, $m_3 = 1$, and so on.

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:

$$D_L(\mathbf{i}_0) = a\Sigma m_j D_j(\mathbf{i}_0) + bD(\mathbf{i}_0). \qquad (5.15)$$

That is, the directional duration naturally splits into two first-order components. The first component, $\Sigma m_j D_j(\mathbf{i}_0)$, reflects price sensitivity to yield slope changes, while the second component, $D(\mathbf{i}_0)$, reflects price sensitivity to parallel yield changes as expected.

Similarly, the directional convexity is calculated to be:

$$C_L(\mathbf{i}_0) = a^2\Sigma\Sigma m_j m_k C_{jk}(\mathbf{i}_0) + 2ab\Sigma\Sigma m_j C_{jk}(\mathbf{i}_0) + b^2C(\mathbf{i}_0). \qquad (5.16)$$

Here we have used the symmetry of $\mathbf{C}(\mathbf{i}_0)$; 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 (5.15), 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. Rather than interpreting the yield curve as the vector $\mathbf{i}=(i_1, \ldots, i_m)$, a yield slope vector, $\mathbf{s}=(s_1, \ldots, s_m)$, is defined as follows:

$$s_1 = i_1;\ s_j = i_j - i_{j-1}, j = 2, \ldots, m. \qquad (5.17)$$

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

From (5.17) we have that $\mathbf{s}=\mathbf{A}\mathbf{i}$, where $\mathbf{A}$ is a linear transformation and $\mathbf{s}$ and $\mathbf{i}$ are column matrices. This transformation is given by:

$$\mathbf{A}=\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 \\ . & & & & & . & . \\ . & & & & & . & . \\ . & & & & & . & . \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}. \tag{5.18}$$

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

$$a_{jk}=\begin{cases} 1 & j=k, \\ -1 & j=k+1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.19}$$

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

$$\Delta s=\mathbf{A}\ \Delta \mathbf{i}. \tag{5.20}$$

Also, **A** is invertible, with:

$$\mathbf{A}^{-1}=\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 & 0 \\ . & . & . & & & & \\ . & . & . & & & & \\ . & . & . & & & & \\ 1 & 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix} \tag{5.21}$$

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

$$b_{jk}=\begin{cases} 1 & j\geq k \\ 0 & \text{otherwise.} \end{cases} \tag{5.22}$$

Based on this transformation, the various approximation formulas in Section 3 can be converted from functions of $\Delta\mathbf{i}$ to functions of $\Delta\mathbf{s}$.

For example, we have from (3.28):

$$P(\mathbf{i}_0+\Delta\mathbf{i})/P(\mathbf{i}_0)\approx 1-\mathbf{D}(\mathbf{i}_0)\Delta\mathbf{i}+1/2\Delta\mathbf{i}^T\mathbf{C}(\mathbf{i}_0)\ \Delta\mathbf{i}. \tag{5.23}$$

Here, the duration term is rewritten in matrix form rather than as a dot product, with $\mathbf{D}(\mathbf{i}_0)$ treated as a row matrix. Substituting $\Delta\mathbf{i}^T=[\mathbf{A}^{-1}\Delta\mathbf{s}]^T$ and using the property of transpose that $(\mathbf{XY})^T=\mathbf{Y}^T\mathbf{X}^T$, we get:

$$P(\mathbf{i}_0+\Delta\mathbf{i})/P(\mathbf{i}_0)\approx 1-\mathbf{D}_s(\mathbf{i}_0)\Delta\mathbf{s}+1/2\Delta\mathbf{s}^T\mathbf{C}_s(\mathbf{i}_0)\ \Delta\mathbf{s}, \tag{5.24}$$

where $\Delta\mathbf{s}$ is given by (5.20) and:

$$\mathbf{D}_s(\mathbf{i}_0)=\mathbf{D}(\mathbf{i}_0)\mathbf{A}^{-1}, \tag{5.25}$$

$$\mathbf{C}_s(\mathbf{i}_0)=(\mathbf{A}^{-1})^T\mathbf{C}(\mathbf{i}_0)\mathbf{A}^{-1}. \tag{5.26}$$

Here, $\mathrm{D}_s(\mathbf{i}_0)$ and $\mathbf{C}_s(\mathbf{i}_0)$ 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:

$$\mathrm{D}_s(\mathbf{i}_0)=\left(\sum_1^m D_j(\mathbf{i}_0),\sum_2^m D_j(\mathbf{i}_0),\ \dots,\ D_m(\mathbf{i}_0)\right). \tag{5.27}$$

That is, the relative sensitivity of the price function to the j-th slope, Δs_j, is the sum of the partial durations from the j-th to the m-th value. The sensitivity of the price function to Δs_1 equals the duration $D(\mathbf{i}_0)$, since $\Delta s_1 = \Delta i_1$, and for this yield curve parametrization, Δi_1 determines the change in the "level" of the yield curve.

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

$$(\mathrm{C}_s(\mathbf{i}_0))_{jk}=\sum_{a=j}^{m}\sum_{b=k}^{m}C_{ab}(\mathbf{i}_0), \tag{5.28}$$

where the jk-th term quantifies the sensitivity of the price function to the product of the j-th and k-th slopes, that is, $\Delta s_j\Delta s_k$. The sensitivity to $(\Delta s_1)^2$ is the convexity $C(\mathbf{i}_0)$.

The total duration vector and convexity matrix defined in (5.27) and (5.28) could have been calculated directly from Definition 3.5 by defining the price function directly in terms of **s**. In particular, given $P(\mathbf{i})$, let the price function $R(\mathbf{s})$ be defined by:

$$R(\mathbf{s})=P(\mathbf{A}^{-1}\mathbf{s}). \tag{5.29}$$

Then $\mathbf{D}_s(\mathbf{i}_0)$ as defined in (5.27) is just the total duration vector of $R(\mathbf{s})$ evaluated at $\mathbf{s}_0=\mathbf{A}\mathbf{i}_0$. Similarly, $\mathbf{C}_s(\mathbf{i}_0)$ is the total convexity matrix of $R(\mathbf{s})$.

6. Summary

The traditional fixed income model for price, and its summary sensitivity measures of duration and convexity, assume parallel yield curve shifts. When the yield curve moves accordingly, this model works well. For other types of shifts, this model can fail to predict the magnitude of the price change, and even its direction. Such events provide a sobering insight to classical hedging and immunization strategies, which rely on this parallel shift assumption.

As a first step toward generalizing the classical theories, this paper has developed the subject of multivariate duration analysis, whereby a model for price sensitivity to arbitrary yield curve shifts was defined and its properties were investigated.

For any fixed yield curve shift assumption, which is identified with a vector **N**, the price function is easily modeled, and familiar approximations to the change in price, ΔP, result. Instead of traditional duration and convexity, however, these approximations reflect "directional" duration and convexity measures. In addition, ΔP was seen to satisfy an exponential identity (Proposition 1) that provided alternative approximations to ΔP that could be used alone, or in combination with the more traditional approximations (Proposition 2), for additional insight to the magnitude and direction of the change in price.

This identity also provided a methodology for investigating under what conditions various approximations would be exact (Proposition 3), and provided a framework for investigating the limiting result when the traditional formulas were applied to ever finer subdivisions of a given yield curve shift (Proposition 4).

A more general model was then investigated in which **N** was not fixed and the yield curve shift, $\Delta \mathbf{i}$, was explicitly modeled as multivariate. Partial durations and convexities then provided natural first- and second-order sensitivity measures, and the traditional parallel shift measures were shown to be summations of the corresponding partial measures (Proposition 5). Also, the earlier exponential identity and associated approximations were seen to have natural extensions to this environment (Proposition 6). In this general setting, the shortcomings of the traditional model exemplified earlier were easily analyzed and understood.

The total duration vector, or vector of partial durations, and corresponding total convexity matrix are easily calculated for a portfolio (Proposition 7) from its component instruments. The resulting measures provide a natural characterization of the yield curve sensitivities developed earlier. For example, the directional duration and convexity values are readily calculated from the corresponding partial values (Proposition 8), while sharp bounds for the resulting directional values also reflected these values (Propositions 10, 11, 12). In the process, the length of the total duration vector, $|\mathbf{D}(\mathbf{i}_0)|$, was seen to provide a summary measure of potential duration risk (Proposition 10).

The concept of equivalent parallel shift, Δi^E, was then introduced as an alternative approach to understanding and investigating duration risk, while the durational leverage, $L(\mathbf{i}_0)$, provided an alternative summary measure of this risk in this context (Proposition 14). When $L(\mathbf{i}_0)$ is large, even small nonparallel shifts can be leveraged into large equivalent parallel shifts, with correspondingly large price effects. The durational multiplier, $M(\mathbf{i}_0)$, provided a technical adjustment to $L(\mathbf{i}_0)$ to correct for the inherent difference in units between nonparallel shifts and traditional parallel shifts.

Applications were pursued in Section 5. Using fixed cash flows and a spot rate yield curve for illustration, the classical duration and convexity formulas decompose in an intuitive way into the corresponding partial duration and convexity counterparts.

For interest-sensitive cash flows, where the price function is implicitly estimated using an option-pricing or other model rather than explicitly described by mathematical formula, the derivative-based formulas for duration and convexity cannot be used directly. However, finite difference approximations to the various duration and convexity measures were shown to be natural generalizations of common approximations for the traditional measures.

While any yield curve basis is workable in theory, throughout this paper the recommended basis was the collection of yield curve drivers on a typical bond yield curve, that is, yields at 0.25, 0.5, 1, 2, 3, 4, 5, 7, 10, 20, and 30 years. Other bond yields are typically interpolated from these market-based observed variables, and all spot rates correspondingly derived from this completed yield curve. Consequently, the price function can be modeled in terms of these 10 or so variables, and all observed price changes related to changes in these values.

Finally, a number of the implications of this multivariate duration analysis for portfolio yield curve sensitivity were also developed.

References

1. BIERWAG, G.O. *Duration Analysis: Managing Interest Rate Risk.* Cambridge, Mass.: 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): 725-42.

3. Bierwag, G.O., Kaufman, G.C., and Khang, C. "Duration and Bond Portfolio Analysis: An Overview," *Journal of Financial and Quantitative Analysis* 13 (November 1978): 671-85.
4. Bierwag, G.O., Kaufman, G.C., and Toevs, A. "Bond Portfolio Immunization and Stochastic Process Risk," *Journal of Bank Research* 13, no. 4 (Winter 1983): 282-91.
5. Black, F., and Scholes, M. "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy* 3 (May-June 1973): 637-54.
6. Chambers, D.R., Carleton, W.T., and McEnally, R.W. "Immunizing Default-Free Bond Portfolios with a Duration Vector," *Journal of Financial and Quantitative Analysis* 23 (March 1988): 89-104.
7. Clancy, R.P. "Options on Bonds and Applications to Product Pricing," *TSA* XXXVII (1985): 97-130.
8. Cox, J.C., Ross, S.A., and Rubenstein, M. "Option Pricing: A Simplified Approach," *Journal of Financial Economics* 7 (1979): 229-63.
9. Fisher, L., and Weil, R.L. "Coping with the Risk of Interest Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies," *Journal of Business* (October 1971): 408-31.
10. Hicks, J.R. *Value and Capital*. New York: Oxford University Press, 1939.
11. Ho, T.S.Y., and Lee, S. "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* 41 (1986): 1011-29.
12. Ho, T.S.Y. *Strategic Fixed Income Investment*. Homewood, Ill.: Dow Jones-Irwin, 1990.
13. Ingersoll, J., Skelton, J., and Weil, R. "Duration Forty Years Later," *Journal of Financial and Quantitative Analysis* 13 (November 1978): 627-50.
14. Jacob, D., Lord, G., and Tilley, J. *Price, Duration and Convexity of a Stream of Interest-Sensitive Cash Flows*. Morgan Stanley & Co., April 1986.
15. Jacob, D., Lord, G., and Tilley, J. "A Generalized Framework for Pricing Contingent Cash Flows," *Financial Management* 16, no. 3 (Autumn 1987): 5-14.
16. 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.
17. Redington, F.M. "Review of the Principle of Life Office Valuations," *Journal of the Institute of Actuaries* 18 (1952): 286-340.
18. Reitano, R.R. "A Multivariate Approach to Duration Analysis," *ARCH* 1989.2: 97-181.
19. Reitano, R.R. "Non-Parallel Yield Curve Shifts and Durational Leverage," *Journal of Portfolio Management* (Summer 1990): 62-67.
20. Reitano, R.R. "A Multivariate Approach to Immunization Theory," *ARCH* 1990.2: 261-312.
21. Reitano, R.R. "Non-Parallel Yield Curve Shifts and Spread Leverage," *Journal of Portfolio Management* (Spring 1991): 82-87.
22. Reitano, R.R. "Non-Parallel Yield Curve Shifts and Convexity," *Study Note 480-24-92*. Schaumburg, Ill.: Society of Actuaries, 1992.
23. Reitano, R.R. "Non-Parallel Yield Curve Shifts and Immunization," *Journal of Portfolio Management* 18, no. 3 (Spring 1992): 36-43.
24. Reitano, R.R. "Multivariate Immunization Theory," *TSA* XLIII (1991): 393-423.
25. Samuelson, P.A. "The Effect of Interest Rate Increases on the Banking System," *American Economic Review* (March 1945): 16-27.
26. Stock, D., and Simonson, D.G. "Tax Adjusted Duration for Amortizing Debt Instrument," *Journal of Financial and Quantitative Analysis* 23 (September 1988): 313-27.
27. Vanderhoof, I.T. "The Interest Rate Assumptions and the Maturity Structure of the Assets of a Life Insurance Company," *TSA* XXIV (1972): 157-92.
28. Vanderhoof, I.T. *Interest Rate Assumptions and the Relationship Between Asset and Liability Structure*. Study Note 8-201-79. Chicago, Ill.: Society of Actuaries, 1979.