

Chapter 5 *Convexity, Time Value and Immunization*

- 5.1 Fixed Income Portfolio Management
 - Institutional considerations
 - Classical immunization
 - Portfolio management and international bond portfolios
- 5.2 Mathematics for Advanced Fixed Income Analysis*
 - Taylor series solutions*
 - Stochastic processes and Ito's lemma*
 - Binomial models and other discrete processes*
- 5.3 The Time Value–Convexity Tradeoff*
 - The contribution from convexity*
 - The Taylor series representation of the time value–convexity tradeoff*
 - A continuous time solution for the time value–convexity tradeoff*
- 5.4 Immunization with Non-parallel Yield Curve Shifts*
 - Partial durations and convexities*
 - Duration and convexity bounds*
 - Some basic calculations and examples*

Questions

Notes

* Indicates sections and subsections that can be considered as advanced material

chapter 5

Convexity, Time Value and Immunization

5.1 FIXED INCOME PORTFOLIO MANAGEMENT

Institutional considerations

The conventional starting point for an examination of fixed income portfolio management is the *classical immunization rules*. The development of immunization rules can be traced to Redington (1952) which provides theoretical conditions to address the general problem of ensuring that funds accumulating in a fixed income portfolio are sufficient to discharge a portfolio of future liabilities. Though Redington considered the case of a life insurance company, the rules apply to a wide range of situations. In most applications, the liabilities are well defined and the primary source of risk is the impact of future changes in interest rates on the return from a portfolio of fixed income securities. It is possible to generalize the asset portfolio to include other securities than just fixed income. More generally, immunization rules can be viewed as strategies for managing the balance sheet to control the impact of changes in interest rates.

A discussion of immunization strategies typically begins with the properties of fixed income securities, e.g., Bierwag (1987). To motivate the discussion, it is convenient to consider the potential interactions between the underlying business and the selection of securities required to address the immunization problem. There are numerous important applications of this problem, including *management of pension fund portfolios*, *life insurance funds* and *financial institution portfolios*. The structure of the portfolio management decision will differ depending on the specifics of the underlying business. For example, a property and casualty company may face the possibility of a large stochastic liability payout that can occur at the single date in the future due to, say, a large hurricane hitting the Florida coast. In contrast, a life insurance company will typically have a more predictable liability structure due to the connection with the more predictable payouts associated with the life table for the pool of policy holders.

Immunization is not the only possible portfolio management strategy that could be used to manage fixed income portfolios or, more generally, the balance sheets of financial institutions. Four of the most important strategies that can be used are: *dedication* or *cash flow matching*, where the cash flow from the fixed income assets is

structured to match the requirements of a portfolio of predetermined liabilities; *immunization*, where the interest rate sensitivity of the cash flows from the fixed income assets and liabilities is matched (*contingent immunization* permits the fund manager to incorporate expectations into the immunization decision); *dollar duration matching*, which extends immunization to positive surplus situations; and, *horizon matching*, which combines dedication and immunization methods by dividing the assets and liabilities being managed into two parts, one part being managed with dedication and the other with immunization. The selection of a particular method of portfolio management will depend on the specifics of the situation. For example, if it is not possible to precisely specify the cash flow requirements associated with the liabilities, then it is not possible to use cash flow matching.

What is the appropriate strategy to use for, say, a life insurance company? In theory, the liability structure of a life company will depend on the age and health characteristics of the pool of policy holders. However, life companies offer a range of products including whole life, guaranteed insurance contracts (GIC's) and various types of bundled and unbundled fixed income and life contingent securities.¹ These products may contain various types of embedded options such as floors, caps and collars. In addition, insurance companies operate in a competitive business environment for the products being sold requiring premiums and rates of return to be offered that are competitive with other types of financial intermediaries. Insurance companies are subject to a range of specific and general regulations, e.g., state insurance laws that impose actuarial restrictions on the composition and rate of return of the asset portfolio. Though it may be possible for insurance companies to employ some cash flow matching, immunization strategies provide greater potential for achieving competitive advantages from asset management.

In the US, insurance companies are subject to state regulation though there have been recent efforts to develop a federal presence. The National Association of Insurance Commissioners (NAIC) serves the role of bringing state regulators and regulations into collective agreement facilitating, for example, the reduction in the regulatory burden on insurance companies operating in multiple state jurisdictions and acting as a vehicle for the implementation of a collective accreditation program, e.g., Vaughan (2002). An important element of the NAIC initiatives is the introduction of a risk based capital system aimed to establish uniform minimum capitalization standards across states, e.g., Barth (2001). Introduced in 1993, the risk based capital (RBC) requirement met considerable resistance from the industry due to concerns about onerous capital requirements being imposed. The evolution of the implementation of RBC requirements reveals that these concerns were largely unfounded.

A number of useful points can be gleaned from the RBC implementation by insurance companies. For example, the asset portfolios of most insurance companies, both life and property/casualty contain a significant amount of non-fixed income securities, such as equities. By implication this would tend to reduce the applicability of a pure immunization model. Yet, despite some qualifications, Barth (2001, p. 240) is able to conclude:

One of the most important outcomes of the RBC system is that the insurance industry itself now has a better understanding of financial risk. Insurers have long been recognized

experts on insurable risk, but the new emphasis on financial risk and capitalization have improved not only the safety but also the efficiency of the insurance industry. New private initiatives into risk measurement and modeling such as dynamic financial analysis and value-at-risk are becoming more prevalent throughout the industry.

Recognizing the duration and convexity can be interpreted as value-at-risk measures for interest rate changes, e.g., Jorion (2001), the relevance of immunization theory for the management of insurance companies is apparent.

Portfolio management for an insurance company is considerably more complicated than, say, for a private company pension fund where investment decisions can be more focused on the sole objective of maximizing the benefits accruing to fund beneficiaries, subject to the actuarial and legal requirements imposed on pension fund managers, e.g., FAS 87 governing the reporting of changes in pension surplus (deficit) in the firm's financial statements. Following Jorion (2001, ch. 17), value-at-risk modeling can be of some value in pension fund management, particularly for large funds that are involved in active trading. As such, classical immunization theory has value for the management of interest rate risk. However, much like insurance companies, the asset portfolios of most pension funds are invested in a range of interest sensitive and interest insensitive assets. As such, the connection may not be as clean as desired. Arguably, it is institutions such as commercial banks that provide the most important application for immunization strategies.

Many banking institutions fall within the scope of the risk based capital requirements that evolved from the 1988 Basle Accord and the subsequent revisions of 1996 and 1999. These accords institutionalize the use of risk management techniques such as value-at-risk for the assessment of capital requirements. This development lends considerable support to the importance of immunization methods when it is recognized that banking institutions in the US are subject to a range of regulations that restrict the balance sheet to largely interest sensitive securities. For example, there are restrictions on the ability of commercial banks to invest directly in equity securities. In addition, the large number of stochastic inflows and outflows make it impractical, if not impossible, for banking institutions to use cash flow matching.

Classical immunization

For pedagogical purposes, it is convenient to develop theoretical conditions for immunization under the assumptions of perfect markets. Particular assumptions can then be relaxed to allow the immunization conditions to be adapted to a specific situation. A key parameter governing the immunization problem is the date at which an obligation is to be discharged, i.e., the *planning* or *investment horizon*. The selection of the investment horizon is a problem for funds that have *multiple* planning periods making it difficult to administer using techniques that treat each individual liability separately by creating dedicated portfolios for each liability, i.e., using *cash flow matching*. The advantage of the immunization approach to portfolio management is that it is possible to treat all the different funds associated with the different possible planning

periods and manage these funds as a single investment fund with a single investment horizon.

The initial statement and solution of the classical immunization problem is due to F.M. Redington (1952), a British actuary. Redington posed the following problem: What allocation of assets and liabilities would minimize a life insurance company's possibility of losses from unexpected changes in market rates of interest? If the initial difference between the market value of assets and liabilities is greater than or equal to zero, Redington proposed that if fund managers adhere to two rules there would be sufficient assets in the portfolio to discharge the liabilities. *Redington's rules*, known as the *classical immunization conditions*, are:

- 1 *Duration matching*: duration of cash inflows equals the duration of the outflows.
- 2 *Higher convexity of assets*: when there is more than one planning period for the fund to satisfy, the value of the cash inflows should be more "dispersed" around the duration than the value of the cash outflows.²

Together, (1) and (2) form the basis of the classical immunization approach to fixed income portfolio analysis.

Redington (1952, p. 290) derived these rules by specifying the "present value of the liability outgo" (V_L) and the "present value of the asset-proceeds" (V_A), both discounted at the same rate of interest. These two present values are then assumed to be equal with "any excess being 'free' funds to be separately invested." This means that Redington was not assuming a zero surplus fund, per se, but was modeling the immunization problem by creating a segmented problem with the surplus considered separately. Redington then evaluates the impact of a change in the interest rate using a Taylor series expansion (see section 5.2). Using the ϵ form of the expansion (see endnote 4), Redington states the expansion as:

$$V'_A - V'_L = (V_A - V_L) + \epsilon \frac{d(V_A - V_L)}{dy} + \frac{\epsilon^2}{2!} \frac{d^2(V_A - V_L)}{dy^2} + \dots$$

Redington then observes that the first term will be zero by assumption and: "In practice the first derivative is the most important for small changes of the rate of interest and I shall define a fund as immunized if the assets are invested so that $[d(V_A - V_L)/dy]$ is zero."

Redington deals with the convexity term by observing that if $[d^2(V_A - V_L)/dy^2] > 0$ then any change in interest "will result in a profit to the fund so long as the change is not so large that the higher terms in the expansion begin to take effect." Redington felt that higher convexity for assets relative to liabilities was "desirable" but used illustrations that showed "the point is not of great importance." Redington then states the theoretical conditions for "a satisfactory immunization policy" by specifying that the first derivative in the Taylor series be zero, i.e., equate the duration of assets and liabilities, and that the second derivative be positive, i.e., the convexity of assets be greater than the convexity of the liabilities. Given this result, the bulk of Redington (1952) is concerned with developing the implications of these immunization rules. For example, Redington

(p. 291) recognizes that the perpetuity represents a limit to the possible asset duration, i.e., the possibility of a duration gap is identified. There are so many insights contained in Redington (1952) that the contribution can still be considered required reading for those seeking to understand the application of immunization theory.

Introductory presentations of the classical immunization conditions assume there is one planning period and a liability with a single payout at a specific date in the future. In this case only the duration matching condition need be considered as the convexity condition is automatically satisfied. This leads to the result: when the initial market value of assets and liabilities is zero (there is no fund surplus or the surplus is treated separately) then immunization occurs when the Macaulay duration of assets equals the Macaulay duration of liabilities. The basic idea behind duration matching in this case requires assuming one planning period and that the liability is in the form of a single cash flow that has to be paid at T , the end of the investment horizon, and that there are only coupon bonds available for investment. If only one bond is purchased to fund the liability, duration matching requires that the liability be covered by buying a coupon bond with a longer term to maturity than the investment horizon. Term to maturity matching of the liability and the asset will fail.

Redington summarizes the logic of the immunization rule for a life insurance company as:

A sceptical reader may ask how it is possible to immunize existing business when, even if no new business is written, the funds will continue to grow for some years, and require the investment of that growth at unknown rates of interest. The verbal answer to this question is that, if the rate of interest falls, there will be a shortfall in the yield on the future investments which have to be made, but there will be an exactly balancing excess in the appreciation of existing investments which have been invested for longer terms than the liabilities they have to meet.

In other words, there are two risks from holding the bond: (1) coupon reinvestment risk; and, (2) risk of capital gain or loss on the bond when it is sold at the end of the investment horizon. Duration matching implies that the time required for investment accumulation associated with coupon reinvestment to offset any capital gain or loss from changing yields is exactly equal to the initial duration. Taking T to be the length of the investment horizon (the payout date on the single liability in the simple case), if $D_A > T$ then the capital gain or loss will dominate the coupon reinvestment return. If $D_A < T$ the coupon reinvestment return will exceed the capital gain or loss.

As Redington explicitly recognized, immunization requires that the portfolio be rebalanced as time goes on, e.g., after each coupon payment date, in order to ensure that the durations of assets and liabilities are equal. "The equations define the position at a moment of time. Their solutions change continuously." In the case of a life insurance company, these changes could be due to "changes in the constitution of the business" or, "for a fixed block of business," due to the diminishing term to maturity of the assets and liabilities with the passage of time. Redington recognizes that cash flow matching ("absolute" matching) is a specific case of immunization where all the terms in the Taylor series are zero and changes in the rate of interest will have no impact "however violent." In addition, Redington recognizes: "The theory would be difficult to

interpret in practice because of the existence of such assets as equity shares, properties, mortgages on an open basis, and so on. Either the income, or the term, or both, of many assets are indeterminate."

Classical immunization theory is, at heart, a hedging theory. As Redington observes: "It has to be remembered that apart from the minor second derivative profits the immunization is against profit as well as loss." Duration measures the $\% \Delta P_B$ when yields change instantaneously, it is not a measure of the return on a bond. For example, consider the case of a zero coupon bond. If the change in yields is zero the return is obtained as the price of the bond increases over time. Similarly, if yields do not change, then the price of par bonds will not change but the return will be obtained by the receipt of coupon income. Duration does not provide any guidance about return in these situations. However, it is possible to use immunization methods to design profitable strategies such as those aimed at "riding the yield curve." Similarly, contingent immunization theory can be used to model key problems in the fund management, such as surplus immunization, and in the management of financial institutions seeking duration gap management.

Portfolio management of international bond portfolios

Classical immunization proposes rules involving the duration and convexity of assets and liabilities. In this approach to fixed income portfolio management, duration is interpreted as the elasticity of the bond price with respect to $(1 + y)$. Even if the problem of currency translation is ignored, the extension of domestic duration to foreign bond prices still poses a problem:

$$\begin{aligned} -\frac{1+y}{P_{Bf}} \frac{dP_{Bf}}{d(1+y)} &= -\frac{1+y}{1+y_f} \frac{1+y_f}{P_{Bf}} \frac{dP_{Bf}}{d(1+y_f)} \frac{d(1+y_f)}{d(1+y)} \\ &= -\frac{1+y_f}{P_{Bf}} \frac{dP_{Bf}}{d(1+y_f)} \frac{1+y}{1+y_f} \frac{d(1+y_f)}{d(1+y)} = D_f^* \eta_{y_f, y} \end{aligned}$$

where P_{Bf} is the price of the foreign bond denominated in foreign currency terms, y_f is the yield on the foreign bond, and D_f^* is the Macaulay duration (foreign duration) measured as the elasticity of the foreign bond price with respect to the foreign yield. This result shows that the elasticity of the foreign bond price with respect to a change in domestic interest rates is the product of two elasticities: the foreign duration, i.e., the elasticity of the foreign bond price with respect to $(1 + y_f)$; and, the elasticity of the foreign interest rate with respect to the domestic interest rate.

Given this, it follows that unless changes in foreign and domestic interest rates are highly correlated, then domestic duration will not be a useful measure of interest rate risk for a foreign bond. To get some idea of the size of the relationship between foreign and domestic interest rates, consider the correlation coefficients for changes in monthly 10 year government bond interest rates across major industrialized countries observed from 2/84 to 09/90 given in table 5.1.³ (By construction, these bond interest rates are

Table 5.1 Correlation coefficients for monthly changes in 10-year government bond interest rates, February 1984–October 1990

	Canada	France	Italy	Japan	Netherlands	Switzerland	UK	W. Germany	US	US *
Canada	1	0.4400	0.1340	0.4865	0.3271	0.2243	0.3944	0.5657	0.8339	(0.70)
France	0.4400	1	0.2000	0.2868	0.5712	0.1364	0.3302	0.4727	0.4711	(0.65)
Italy	0.1340	0.2000	1	0.0923	0.1760	0.2251	-0.0522	0.1579	0.2155	
Japan	0.4865	0.2868	0.0923	1	0.3503	0.3368	0.3456	0.5486	0.4517	(0.30)
Netherlands	0.3271	0.5712	0.1760	0.3503	1	0.2956	0.4881	0.6785	0.3894	(0.65)
Switzerland	0.2243	0.1364	0.2251	0.3368	0.2956	1	0.1889	0.2249	0.2780	
UK	0.3944	0.3302	-0.0522	0.3456	0.4881	0.1889	1	0.5261	0.3081	(0.55)
W. Germany	0.5657	0.4727	0.1579	0.5486	0.6785	0.2249	0.5261	1	0.5123	(0.55)
United States	0.8339	0.4711	0.2155	0.4517	0.3894	0.2780	0.3081	0.5123	1	

Sources: Government bond interest rates from Bloomberg Financial Markets, adapted from Dym (1992). Values in (US*) column are from Thomas and Willner (1997) for a monthly sample over 1977–95

calculated in terms of the currency of issue.) Though there is a high correlation of .8339 between the US/Canada, correlations are typically much lower with Italy and the UK, two members of the European Monetary System (EMS), even exhibiting negative correlations. (The EMS was a precursor of the European Monetary Union). As indicated by the bracketed correlation coefficients in table 5.1, empirical evidence from other sources for different sample periods indicates that there is some variation when different samples are used. On balance, these empirical results confirm that duration for foreign bonds (measured in terms of domestic interest rate changes) will tend to be inaccurate indicators of the price elasticity with respect to domestic interest rate changes.

Following Dym (1992) and Thomas and Willner (1997), it is possible to extend the duration measure to produce a risk measure for the change in a bond price, measured in terms of the issue currency. Recalling that the total derivative of $\ln[x]$, $d\{\ln[x]\} = (1/x) dx = dx/x$, manipulation of the duration formula gives:

$$d\{\ln[P_{Bf}]\} = \frac{dP_{Bf}}{P_{Bf}} = -D^* \cdot d\{\ln[1 + y_f]\}$$

This form directly illustrates the role played by duration in translating yield changes into bond price changes. A risk measure associated with domestic returns now follows from evaluating the standard deviation of dP/P :

$$RM = Risk = std[dP_{Bf}/P_{Bf}] = (D^*) std[d\{\ln[1 + y_f]\}]$$

where $std [\cdot]$ is the standard deviation operator. The resulting decomposition of the right hand side of the equation, together with the calculated risk measure, is provided in table 5.2. The large difference in the local yield volatility is apparent. Contrary to the

Table 5.2 Government bond interest rate risk measures, February 1984–October 1990^a

Country	Macaulay duration	Yield volatility (×100)	Average yield (%)	RM (risk measure)	β (beta)
Canada	6.34	0.3867	10.55	0.0222	1.0838
France	6.48	0.3695	10.29	0.0217	0.6871
Italy	6.07	0.7696	12.19	0.0416	0.6052
Japan	7.33	0.3286	5.52	0.0228	0.8226
Netherlands	6.84	0.2366	7.16	0.0151	0.4767
Switzerland	7.35	0.1876	7.87	0.0128	0.2514
UK	6.22	0.4439	10.49	0.0250	0.7528
W. Germany	6.86	0.2503	7.01	0.0160	0.6018
US	6.92	0.3878	9.25	0.0246	1.1874

^a β measures the change in the foreign government bond yield relative to changes in the Merrill Lynch Global Bond index measure of global interest rates.

prediction of the simple duration model, the riskiest bonds do not have the highest durations. Quite the opposite, the highest duration bond (Switzerland) has the lowest risk and the lowest duration bond (Italy) has the highest risk.

Ultimately, the domestic (US) currency return on a foreign bond requires incorporating exchange risk. In addition, it is useful to decompose a given bond interest volatility into the influence of local and global interest rate movements, where the global component is measured by a market size weighted average of long-term bond rates of industrialized countries as measured by the Merrill-Lynch global bond market index. In this case, a CAPM-type model for bond pricing can be introduced:

$$\Delta(1 + y_f) = \Delta y = \beta \Delta I + \Delta u$$

where I is the global interest rate and u is the random error term representing the country specific (local) component. Estimates for β are provided in table 5.2. Assuming independence of u and I gives:

$$\text{var}[\Delta(1 + y_f)] = \beta^2 \text{var}[\Delta(1 + I)] + \text{var}[\Delta(1 + u)]$$

where $\text{var}[\cdot]$ is the variance operator. This result can be used to decompose the domestic currency denominated interest rate risk of a foreign bond into local and global components.

The results for bond returns denominated in the currency of issue confirm, to a certain extent, the importance of the composition of the global portfolio. This is reflected in the size of the global component in the individual country measures for interest rate risk. Given this, correlations among changes in the local factors (not reported) are typically small for the non-EMS currencies. Correlations between EMS currencies and the currencies that compose the bulk of the global market portfolio, i.e., Japan and the US, are typically negative. Given this information on the risk structure of the domestic currency denominated bond returns, it is now possible to extend the analysis to permit the return of each bond to be valued in US dollars. In this case the bond's US dollar value at time t , $V(t)$, will equal $P_{Bf}(t)/S(t)$, where $S(t)$ is measured in foreign (not US) direct terms. Taking logs and first differences gives:

$$\Delta \ln V = -D^* \cdot \Delta \ln [1 + y_f] - \Delta \ln e$$

where $e = dS/S$. Given this it is now possible to determine a domestic currency denominated risk measure that accounts for the risks of local and global interest rates and the exchange rate.

The procedure for developing the measure extends the method used for table 5.2. Consider first:

$$\text{Std} [\Delta \ln[V]] = \sqrt{(D^*)^2 \text{var}[\Delta \ln[1 + y_f]] + \text{var}[\Delta \ln[e]] + 2D^* \text{cov}[\Delta \ln[1 + y_f], \Delta \ln[e]]}$$

where $\text{cov}[\cdot]$ is the covariance operator. Following Dym (1992, pp. 87-8), it is now possible to substitute the CAPM-type model of interest rates for $\Delta(1 + y_f)$ to provide

Table 5.3 Components of the risk of the domestic currency return for foreign government bonds divided by total foreign bond risk, February 1984–October 1990^a

Country	Global interest rate	Local interest rate	Exchange rate	Global interest/exchange rate covariance	Local interest/exchange rate covariance
Canada	0.4226	0.2485	0.1834	0.0780	0.0672
France	0.0860	0.2255	0.6304	0.1005	-0.0425
Italy	0.0313	0.6013	0.3210	0.0501	-0.0039
Japan	0.1254	0.1233	0.5285	0.1533	0.0692
Netherlands	0.0469	0.0967	0.6951	0.0845	0.0766
Switzerland	0.0154	0.0908	0.7623	0.0516	0.0797
UK	0.0575	0.1914	0.4639	0.0518	0.2352
W. Germany	0.0746	0.0858	0.6561	0.1111	0.0722
US	0.7421	0.2579	—	—	—

^a All values are expressed in terms of shares of $std[\Delta \ln[V]]$ where std is the standard deviation and V is the domestic (US) currency value of the foreign government bond. Values may not sum to one due to rounding error.

Source: Adapted from Dym (1992)

a decomposition of foreign bond risk into the relevant components associated with global interest rate risk, local interest rate risk and exchange rate risk. Recognizing that the local and global interest rate components have no covariance by construction, there will also be covariance terms for the exchange rate with the global and local interest rate components. Dividing these risk components by $std[\Delta \ln[V]]$ reexpresses the values in terms of percentage contribution to the total risk. These results are given in table 5.3.

Examination of table 5.3 reveals that, for a US investor, the addition that currency risk makes to total risk is similar to the results reported in table 1.14. More precisely, though there is significant variation in the exchange rate component across countries, the general result is that the bulk of the risk associated with unhedged holding of foreign government bonds is associated with the exchange rate. Excluding Canada and Italy, the impact of global and local interest rate changes is not much different from the covariance parts. In the Canadian case, the relatively stable US\$/C\$ exchange rate results in the contribution of currency risk to total risk for Canadian bonds being well below 20%. Another exception is Italy where more than 60% of the total risk was due to the local interest rate component. A number of countries exhibited high positive covariance terms, either with the global or local interest rate component. The variation in the local interest rate/exchange rate covariance shares could be partially due to the use of domestic interest rate policy to control the spot exchange rate.

5.2 MATHEMATICS FOR ADVANCED FIXED INCOME ANALYSIS*

*Taylor series solutions**

The Taylor series expansion is one of the most practical and widely used tools of real analysis, e.g., Rudin (1964). The technique is a special type of power series expansion that applies to *analytic functions*. The basic idea is a specific solution to the more general problem of approximation of functions, a topic that has occupied mathematicians for centuries. Functions appear in a wide range of situations. Real analysis is concerned with functions defined using real variables, i.e., defined over a subset of the real line ($\mathbb{R}^1$) for functions of one variable and $\mathbb{R}^1 \times \mathbb{R}^1 \times \dots$ for functions of n variables. The restriction that the function be "analytic" imposes further conditions on the function. For a function of one variable, these conditions are that the function be continuous and differentiable in the interval of convergence of the power series. In fixed income security analysis, the main function of interest is the value or price function. For default and option free bonds, the price function is convex in the yield to maturity, a property that easily satisfies the conditions required for a Taylor series to be used. With appropriate restrictions, Taylor series expansions can be used to extend the analysis to bonds with option features or the possibility of default.

The Taylor series technique can be illustrated by making a comparison between the discrete and continuous forms of discounting operators (where t^* is the fraction of the year remaining to maturity). Observe that for $t^* \leq 1$ with the continuous discounting operator $\ln[\exp\{-rt^*\}] = -rt^*$, the corresponding $\ln[(1 + rt^*)^{-1}]$ for the discrete compounding case does not provide a similar simple solution but can be evaluated by taking a *Taylor series expansion*. For a function of one variable, this expansion takes the general form:⁴

$$f[x] = f[a] + \frac{df[a]}{dx}(x - a) + \frac{1}{2!} \frac{d^2f[a]}{dx^2}(x - a)^2 + \frac{1}{3!} \frac{d^3f[a]}{dx^3}(x - a)^3 + \dots$$

In other words, any function $f[x]$ that has continuous derivatives over some interval $[b, c]$ can be equivalently expressed as the converging infinite sum provided by the Taylor series. This expression expands the univariate function $f[x]$ about the fixed point a , where $b \leq a \leq c$. Each of the derivatives in the expansion are evaluated by setting $x = a$. For this expansion to be valid, the function $f[x]$ must have derivatives of all orders over $[b, c]$ (some of which can be zero).

Term by term inspection of the Taylor series reveals how the function $f[x]$ is approximated. The first term $f[a]$ is a point, the value of the function evaluated at the point $x = a$. The sum of the first term and the second term is the linear approximation to the function about the point a . Assuming that $f[x]$ is the price function for a default and option free bond, the linear approximation is illustrated in figure 5.1. Taking into account the third term in the Taylor series involves using a quadratic approximation. The implications of this are illustrated in figure 5.2. As illustrated, even though the assumed function is convex, it does not follow that a quadratic approximation will

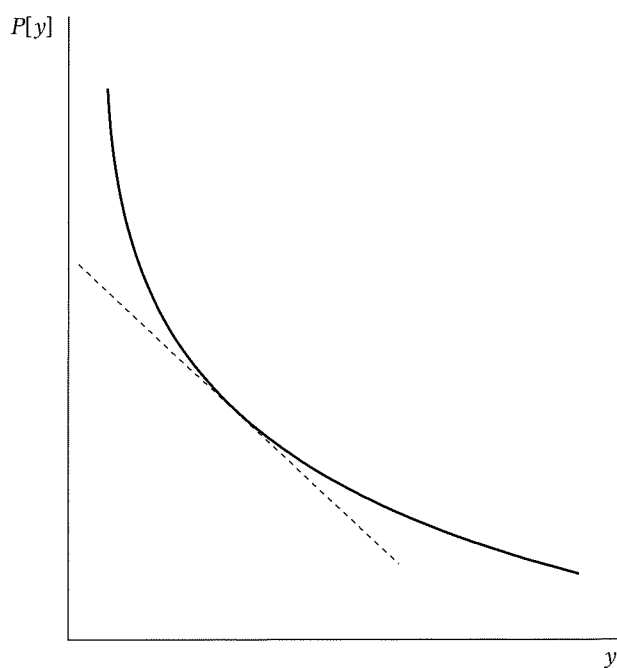


Figure 5.1 Linear approximation to a convex function

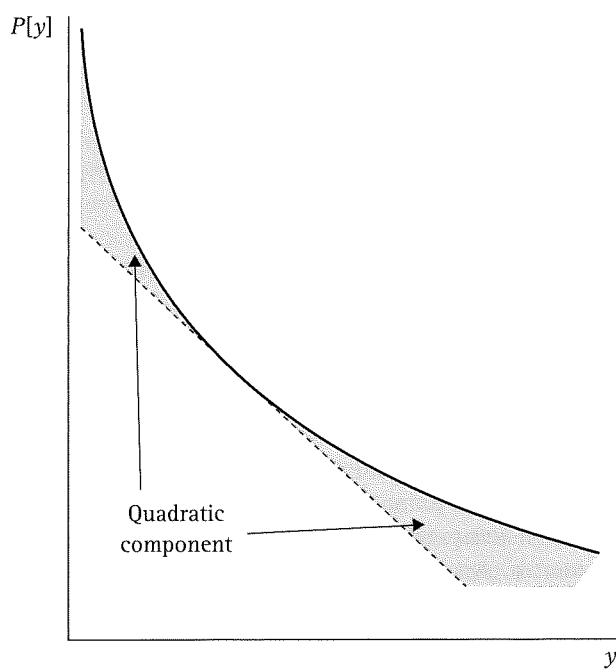


Figure 5.2 Quadratic approximation to a convex function

provide an exact fit to the function. In addition, because a Taylor series only possesses the property of uniform convergence, it does not follow that taking an additional term into account will necessarily produce a closer approximation. All that the Cauchy criterion for uniform convergence ensures is that if enough terms are taken into account then the approximation will be within ϵ of the true function, for all $\epsilon > 0$.

A Taylor series is a specific form of the more general mathematical concept of a power series. Other types of power series can also be used to represent an admissible function as an expansion expressed as an infinite converging sum. Examples of such expansions occur frequently, particularly in mathematical statistics. Some of the more important are the Edgeworth expansion, Cornish-Fisher expansion and the Gram-Charlier expansion. The various types of power series are distinguished primarily in the method used for determining the coefficients for each of the terms in the expansion. The selection of one method of expansion over another depends on the specifics of the function being examined and the use required for the expansion. The Taylor series uses derivatives of the function $f[x]$ in order to determine the coefficients. An important special case of the Taylor series occurs where zero is used as the fixed point in the expansion. This form of the Taylor series is sometimes referred to as a Maclaurin series.

Applied to the function $f[r] = \ln[(1 + rt^*)^{-1}]$ evaluated around the origin ($a = 0$) gives:

$$\begin{aligned} \ln \left[\frac{1}{1 + rt^*} \right] &= -\ln[1 + rt^*] = - \left\{ \ln[1 + (0)t^*] + \frac{t^*}{1 + (0)t^*} (r - 0) \right. \\ &\quad \left. - \frac{1}{2!} \frac{(t^*)}{(1 + (0)t^*)^2} (r - 0)^2 + \frac{1}{3!} \frac{(t^*)}{2(1 + (0)t^*)^3} - \dots \right\} \\ &= - \{ 0 + rt^* - r^2 t^{*2} + r^3 t^{*3} - \dots \} \end{aligned}$$

For small r , say r is .1, the terms involving squares, cubes and so on will be relatively small and the expansion reduces to $\ln[1/(1 + rt^*)]$ being approximately equal to $-rt^*$, which is the same as the exact value given when the log of the continuous discounting operator was evaluated. Using this approach, the difference between the continuous and discrete interest rates can be represented by the higher order terms which are being ignored in the Taylor series expansion. Application of the Taylor series to the geometric discounting operator produces: $\ln[(1 + r)^{-t^*}] = -t^* \ln[(1 + r)]$ which is also approximately $-rt^*$. However, unlike the linear form where the approximation to the log was taken for $(1 + rt^*)$, this approximation is $-t^*$ times $\ln[1 + r]$ which will, generally, produce a different error of approximation.

This discussion can also be used to clarify an approximation that is commonly used in financial calculation: $\ln[1 + x]$ is approximately equal to x when x is small. The question this operation raises is: how small can x be in order to ensure that substitution of x for $\ln[1 + x]$ does not produce sizeable errors of approximation? Examining the Taylor series for $\ln[1 + x]$ reveals that the square of x must be small relative to x . For example, if x is .1 then $x^2 = .01$ which, for many applications, is sufficiently small to permit the approximation. Similarly, if $x = .05$ then $x^2 = .0025$, and the approximation is again admissible in many situations. However, for $x = .5$, $x^2 = .25$, a value which is

not small relative to x . In practice, the admissibility of the approximation will depend on the specific type of problem in which the approximation is used. It is always possible to improve the approximation by taking account of higher order terms in the expansion.

For present purposes, the application involves applying the Taylor series expansion to the bond price function, $P[y]$. This application goes back at least to Redington (1952). While it is possible to define the bond price function in terms of yield and time, i.e., $P[y, t]$, the conventional textbook explanation for the relationship between duration and convexity, e.g., Fabozzi (1993, ch. 4), is to treat the bond price as a univariate function of yield. Applying a Taylor series expansion to this function gives:

$$P_B[y] = P_B[y_0] + \frac{dP}{dy}(y - y_0) + \frac{1}{2} \frac{d^2P}{dy^2}(y - y_0)^2 + \dots H.O.T.$$

$$\frac{P_B[y] - P_B[y_0]}{P_B[y_0]} \cong -DUR(y - y_0) + \frac{1}{2} CON(y - y_0)^2$$

where $DUR = D$ and CON are the modified duration and convexity of the bond. This result can be used to show that for two different fixed income portfolios, with equal initial yield and duration, that the higher convexity portfolio will have a more favorable percentage change in price, whether yields go up or down (see section 5.3 below).

As demonstrated by Christensen and Sorensen (1994) and others, this analysis is incomplete because the bond price function can be modeled as a function of two variables, yield and time. This expansion involves the application of a multivariate Taylor series:

$$\begin{aligned} P_B[y, t] &= \sum_{t=1}^T \frac{C}{(1+y)^t} + \frac{M}{(1+y)^T} \\ &= P_B[y_0, t_0] + \frac{\partial P}{\partial y}(y - y_0) + \frac{\partial P}{\partial t}(t - t_0) \\ &\quad + \frac{1}{2} \left\{ \frac{\partial^2 P}{\partial y^2}(y - y_0)^2 + \frac{\partial^2 P}{\partial t^2}(t - t_0)^2 \right\} + \frac{\partial^2 P}{\partial y \partial t}(y - y_0)(t - t_0) + \dots H.O.T. \end{aligned}$$

Dividing through by $P_B[y_0, t_0]$ will produce the familiar expansion in terms of modified duration, convexity, theta and the cross product terms, e.g., Jordan and Chance (1995), Kang and Chen (2002). Though this expansion does clarify certain points, there are also some limitations in arriving at conclusions based on this form of the expansion (see section 5.3 below).

*Stochastic processes and Ito's lemma**

BASIC BACKGROUND

As discussed in section 1.3, the theory of stochastic processes describes the behavior of random variables, the $X(t)$'s, over time, $t \in \mathfrak{T}$. A random variable is a function which

maps from a prespecified domain, or sample space, to some portion of the real line, $\mathbb{R}^1$. In certain financial applications, e.g., where X refers to a security price, X takes values only on the positive, half line. In this case as well as when the X values are allowed to assume any value along the real line, it is conventional to assume that there is a zero probability of X being equal to plus or minus infinity. When t is fixed at a given point, $X(t)$ has the conventional interpretation of a random variable, with associated (one-dimensional) probability density function. Specification of the stochastic process for X requires further specification of the joint density functions that relate X 's at different points in time: the joint densities provide a probabilistic specification of how X evolves over time. This potentially complicated mapping can involve various combinations of discrete or continuous observations on X and t . While in financial applications X is usually continuous, time can be either. The terms *discrete stochastic process* and *continuous stochastic process* are used to refer to the time intervals at which $X(t)$ is observed. Applications of stochastic processes have contributed significantly in numerous fields, e.g., engineering, physics, biology and statistics.

A number of important notions from stochastic processes have received considerable attention in modern Finance. While martingales and random walks are possibly the most familiar,⁵ most of the literature on option pricing is developed using diffusion processes. These concepts are all closely related. *Diffusions* are (strong) Markov processes, continuous in both X and t . In turn, diffusions are constructed from Wiener processes which are the continuous time representation of a discrete time random walk. Both types of processes obey the martingale property. The essential concept of a *Markov* process is most readily illustrated in discrete time. Using the conditional probability distribution $Pr[\cdot]$, the Markov property can be stated:

$$\begin{aligned} Pr[X(t+1) = j \mid X(0) = a, X(1) = b, X(2) = c \dots, X(t) = s] \\ = Pr[X(t+1) = j \mid X(t) = s] \end{aligned}$$

In effect, the Markov process is "memoryless," the past and future are statistically independent when the present is known.

While there are a number of subtle technical issues associated with the precise definition of Markov process, Markov families and diffusions (e.g., Rogers and Williams 1987), for current purposes it is sufficient to define a diffusion process as a Markov process in which both time and the state variable X are continuous. When expressed in the form of a *stochastic differential equation (SDE)*, diffusions can be used to concisely specify the joint density functions for the stochastic process. To adequately motivate stochastic differential equations, it is necessary to develop the basic concept of a Wiener process as the limit of a discrete "normal" random walk, defined by a stochastic difference equation. The discrete *random walk*, without drift, has the form:

$$X(t+1) = X(t) + Z(t+1) \quad \text{where } X(0) = 0, \text{ and } t \in \{1, 2, 3 \dots\}$$

where $Z(1), Z(2), Z(3) \dots$ form a stochastic process of independent random variables with the standard normal probability distribution: $Z(t) \sim N[0, 1]$. In other words, over the unit time interval ($\Delta t = 1$) $Z(t)$ is a normal random variable with mean 0 and variance of one.

TYPES OF DIFFUSIONS

Consider what happens to a normal random walk as the time interval Δt shrinks. In this case, $Z(t + \Delta t)$ is no longer $N(0, 1)$, but rather $N[0, \Delta t]$ (e.g., Poitras 2002, pp. 545–6) where the random walk now has the form (see section 10.4):

$$X(t + \Delta t) = X(t) + Z(t + \Delta t) = X(t) + Z(t)\sqrt{\Delta t}$$

Reexpressing in difference form and using:

$$\lim_{\Delta t \rightarrow 0} \Delta t = dt$$

Gives the stochastic differential equation for the standard Wiener process:

$$dX(t) = Z(t)\sqrt{dt} \Rightarrow dX(t) = dW(t)$$

$W(t)$ is usually referred to as a *standard Wiener process* with a unit variance parameter. The terminology “Gaussian process” or “Brownian motion” is also used. As discussed in section 1.3, the *ensemble* of sample paths for $W(t)$ conform to the evolution of a random variable that is standard normal on the unit time interval.

Geometrically, the behavior of the standard Wiener process can be illustrated in two dimensions by taking $X(t)$ on the vertical axis and time on the horizontal axis. Starting from $X(0)$, which is a predetermined point, the Wiener process specifies an infinite number of possible paths originating from $X(0)$. The pattern of these paths conform to $N[0, \Delta t]$. To see this, select any time $t_1 > 0$, take a “slice” across the X paths and plot the distribution of the paths. The distribution of the paths will be a normal distribution, centered at $X(0)$, with variance t_1 . Similarly, doing the same “slicing” operation for another time $t_2 > t_1$ and evaluating the density associated with the ensemble of X time paths will again produce a normal distribution, centered at $X(0)$, but with a larger dispersion. In this fashion, it is possible to make a connection with the distribution theory, familiar from traditional statistics, which is concerned with variables at a given point in time. The SDE technique is a method of describing the evolution over time of random variables that are functions of the class of normal random variables.

The Wiener process can be immediately generalized to allow for non-zero drift and variance that differs from Δt . When a trend or “drift” term μ and standard deviation $\sigma (\neq 1)$ are admitted then the *arithmetic Gaussian* stochastic process (with constant coefficients μ and σ) is defined:

$$dX(t) = \mu dt + \sigma dW(t)$$

This constant coefficient process is also referred to as an arithmetic or absolute Brownian motion, e.g., Poitras (1998). The ensemble of time paths for this process are also normal but differ from the standard Wiener process by allowing for a different amount of variation around a constant trend. The standard Wiener process can be used to construct a wide range of SDE's, each of which is associated with a different specification

for the joint density of the stochastic process. The construction of the Weiner process requires these densities to be functionally connected to the normal.

In general, the drift and standard deviation can be functions of both the state variable and time, such that $\alpha[x, t]$ is the drift and $\beta[x, t]$ is the volatility. Consider, a simple form of state dependence, where $\alpha[x, t] = \mu X$ and $\beta[x, t] = \sigma X$, with μ and σ being constants:

$$dX(t) = \mu X dt + \sigma X dW(t) \Rightarrow dX(t)/X = \mu dt + \sigma dW(t)$$

In this case, the instantaneous rate of change (dX/X) follows a Gaussian process. For this reason, the terms *geometric Gaussian process* or *geometric Brownian motion* are sometimes used to identify this process. It can be shown that for the geometric Gaussian process the paths of $X(t)$ correspond to a process that is lognormally distributed at each point in time. This is important for cases where $X(t)$ refers to prices which, like lognormal variables, cannot be negative (for $X(0) > 0$).⁶

The geometric Brownian motion has an important position in option pricing theory. Black and Scholes (1973) use this process to describe the behavior of the stock price in deriving their European option pricing formula. The process is often encountered in a variety of option pricing situations because geometric Brownian motion usually leads to a relatively simple closed form solution. While the empirical validity of this assumption can be questioned, the process is sufficiently close to real world processes that in many cases a reasonable approximation is provided.⁷ However, in the case of fixed income securities and interest rates, geometric Brownian motion has definite limitations. For example, the price of fixed income securities such as Tbills or Tbonds will converge to the par value on the maturity date, which is inconsistent with unrestricted geometric Brownian motion. Even where the interest rate instead of the price is used as the stochastic variable, there is considerable evidence that the empirical process exhibits mean reversion, a property that is not consistent with geometric Brownian motion.

Despite the empirical limitations, it is still possible to use geometric Brownian motion as the process for stochastic modeling of interest rate behavior. For various reasons, it is useful to know the mean and variance of an assumed diffusion process. For example, a key step in the risk neutral valuation of the European option requires this result. For geometric Brownian motion:

$$E[X(t)] = X(0)e^{\alpha t} = X(0)e^{\left(\mu + \frac{\sigma^2}{2}\right)t} \quad \text{var}[X(t)] = [X(0)^2 e^{2\alpha t} [e^{\sigma^2 t} - 1]]$$

These results are derived by directly evaluating the expectations. Applications of these results appear, for example, in the risk-neutral derivation of the Black-Scholes option pricing formula, e.g., Poitras (2002, ch. 8).

An important class of processes that is exploited in some pricing of options and stochastic modeling of interest rates are the variations of the *regular Ornstein-Uhlenbeck* (OU) process:⁸

$$dX(t) = \alpha X dt + \sigma dW(t)$$

Observe that an OU process has a positive probability of having negative values. In financial applications, the OU process is often presented as a *mean reverting* process:

$$dX(t) = \alpha[\mu - X]dt + \sigma dW(t)$$

where α can be interpreted as the speed of adjustment of $X(t)$ to the steady state mean μ . Different variations of the OU process have appeared. For example, Brennan and Schwartz (1979) and many others use the process:

$$dX(t) = \alpha[\mu - X]dt + \sigma X dW(t)$$

In this case, dX/X reduces to a mean-reverting OU.

Unlike the geometric and arithmetic Brownian motions, the $X(t)$ distribution associated with the OU process changes over time. At any time t , $X(t)$ is governed by a non-steady state distribution which is normal but with *conditional* means and variances that are time dependent, e.g., Cox and Miller (1965, pp. 225–7). For the *regular OU*:

$$E[X(t)] = X(0)e^{-\alpha t} \quad \text{var}[X(t)] = \frac{\sigma^2(1 - e^{-2\alpha t})}{2\alpha}$$

Asymptotically, as $t \rightarrow \infty$ then the OU converges to a normal steady state distribution with mean zero and variance $\sigma^2/2\alpha$. This damping behavior of the OU is different from the geometric Brownian motion which increases indefinitely as $t \rightarrow \infty$. Similarly, the *mean-reverting OU* has conditional parameters:

$$E[X(t)] = \mu + (X(0) - \mu)e^{-\alpha t} \quad \text{var}[X(t)] = \frac{\sigma^2}{2\alpha} \{1 - e^{-2\alpha t}\}$$

As $t \rightarrow \infty$ then the mean-reverting OU converges to a steady state normal distribution with mean μ and variance $\sigma^2/2\alpha$.

Cox, Ingersoll and Ross (1985) and others use a diffusion process, the square root process, in which volatility depends on the square root of X . In mean-reverting form, the SDE for this process is represented:

$$dX(t) = \alpha[\mu - X]dt + \sigma\sqrt{X} dW(t)$$

This process can be shown to be the continuous limit of a *noncentral* chi-squared distribution. In turn, the square root process is a special case of the more general *constant elasticity of variance* (CEV) process:

$$dX(t) = \alpha X dt + \sigma X^{\beta/2} dW(t)$$

where $\beta \in [0, 2)$. Unlike the other diffusions examined to this point, the CEV model describes a class of processes defined over the range of β . For $\beta = 1$, the CEV process is equivalent to a square root process. For $\beta = 2$, the process is lognormal.

The *Brownian bridge* process is a development on the Brownian motion process that incorporates a drift term which forces the process to a fixed endpoint. For a Brownian bridge process that starts at zero and ends at zero, the unit variance process takes the form:

$$dX(t) = -X(t)/(T - t)dt + dW(t)$$

The Brownian bridge process has been used in a number of papers, e.g., Ball and Torous (1983), Chiang and Okunev (1993), to model the bond price process. For this purpose, the Brownian bridge is attractive because the bond price converges to par value as the term to maturity endpoint is approached. Cheng (1991) explores problems that assuming a Brownian bridge has for arbitrage free pricing.

Unfortunately, the number of diffusion processes for which readily interpretable conditional probability densities can be derived is limited. While the density function for the general CEV process has been derived, the formula is complicated. In addition to the processes already described and immediate extensions, such as the Brownian bridge, the only other class of processes that have manageable density functions are Bessel processes which, in squared form, have the representation: $dX(t) = \alpha dt + 2\sqrt{X}dW(t)$. This form of the Bessel process is typically used because the associated densities are stable under convolutions, a useful property not shared by geometric Brownian motion. For example, Geman and Yor (1993) use this process to model the pricing of perpetuities and Asian options. In general, while not always the best empirical fit to the distribution for financial random variables, the analytical advantages of diffusion processes have played important roles in the application of stochastic processes to problems in finance and economics.⁹

With relatively little analytical complication, the general SDE model can be extended to include *jump processes*. One of the properties of the diffusion is the continuity of its sample paths. This property can be generalized to permit certain types of jumps in the state variable to take place, usually modeled as a Poisson event process. In other words, the sample paths of the diffusion are continuous except at a countable number of discontinuity points, where the jumps are generated by a Poisson process. In this case, the SDE can be written:

$$dX(t) = \alpha[x, t]dt + \sigma[x, t]dW(t) + \nu[x, t]dQ(t)$$

where $\nu[x, t]$ obeys the same technical conditions imposed on α and σ , and $dQ(t)$ is a Poisson process assumed to be distributed independently of $W(t)$.¹⁰ For this type of SDE, there is a generalized form of Ito's lemma, e.g., Malliaris and Brock (1982, p. 122). The usefulness of this result in applications is that it, theoretically, permits the valuation of jumps in the underlying process, e.g., Merton (1976), Cox and Ross (1976).

EMPIRICAL EVIDENCE ON STOCHASTIC PROCESSES FOR INTEREST RATES

Given all the possible diffusion processes that could be selected, it is natural to ask the question: what is the appropriate SDE to use for modelling an interest rate process?¹¹

The answer to this question depends on the problem that the SDE is being used to analyse. In many applications, e.g., where the problem involves a one-factor model of the term structure, the objective is to select the SDE that has the best empirical fit to the short term interest rate that, in turn, may be used to determine the term structure of interest rates. Such applications arise in pricing interest rate contingent claims, including interest rate derivatives, bonds with embedded options and swaptions. The importance of determining the appropriate SDE (or non-linear discrete process) for the interest rate process has not escaped the proponents of modern Finance. Since the early 1990s, an impressive number of articles have been dedicated to this topic. Yet, Chapman and Pearson (2001, p. 91) conclude: "After all this research, we know a few important facts about the evolution of interest rates. Unfortunately, the list of things we do not know is as long as the list of things we do know."

The connection of this statement with the philosophical discussion in part I is difficult to ignore. However, it is still instructive to consider what "few important facts" have been identified. Chapman and Pearson (2001, p. 78) start with three "stylized facts" about short term interest rates:

Short rate fact #1: The short-rate series is a "persistent" time series; that is, it spends long consecutive periods above or below the (sample estimate of the) unconditional, or long-run, mean.

Short rate fact #2: In the 1979–82 period, the average level and volatility of the short rate was substantially higher than for other years in the 1971–2000 period.

Short rate fact #3: The volatility of the short rate level appears to be both time varying and persistent.

Chapman and Pearson also observe that although the 5 year constant maturity Treasury yields (1962–2000) have "lower mean and volatility" than the short rate measured using the one-month Eurodollar yield, "the overall movements in levels and volatility are qualitatively similar." Hence, short term interest rate processes can be characterized by "persistence in level," structural breaks and "time-varying and persistent" volatility. "Strong contemporaneous correlations between rates of different maturities" is also observed leading Chapman and Pearson to conclude that "a limited number of common factors" determine "virtually all of the variability in weekly yield changes for bonds of maturities from six months to 18 years." Principal components analysis is used to "confirm" the results of Litterman and Scheinkman (1991) that there are three factors that determine the variation in yield changes. The first factor, which explains "88 percent of the variation," is the "level." The other two factors being "slope" and "curvature" explain a further 11 percent of the variation.

In addition to these "facts" Chapman and Pearson also present other "facts" such as: "the volatility of yields was decreasing for increasing terms to maturity." However, this empirical result does not hold over all sub-periods. There does appear to be high contemporaneous correlations between yield volatilities for all adjacent maturities "even for 1-month versus 10-year yields." In addition to these, "facts," Chapman and Pearson also present potentially illuminating results from working papers and forthcoming

studies that are "very preliminary and, consequently, must be interpreted with caution." It seems this preliminary evidence indicates "patterns in term premiums do not permit distinguishing among competing models. It may be necessary to examine data on the dynamics of interest options to narrow the field of acceptable dynamic models" (p. 91). Perhaps the inability of techniques such as "two-factor regime switching models" to capture the *ex ante* behavior of interest rates and the term structure is due more to the failure of the assumptions of logical positivism?

In the end, Chapman and Pearson (2001) stands as an example of the extent to which logical positivism dominates the methodology of modern Finance. The discussion is predicated on the assumption that by examining more data with "better" statistical techniques the "true" model generating interest rates and the term structure can be identified. Some indication of the validity of this approach can be gained by considering the various models of the interest rate process that have been suggested in numerous studies over the last three decades. Almost a decade prior to Chapman and Pearson, Chan et al. (1992) provide a description and empirical evaluation of eight different diffusion specifications. In contrast to one of the "facts" provided by Chapman and Pearson, Chan et al. (1992, pp. 1224-5) find "no evidence of a structural shift in the interest rate process in October 1979 for the models which capture the conditional volatility of the interest rate process". In addition, Chan et al. stress the failures of available models to explain "the dependence of volatility on the level of the interest rate".

Despite the "progress" identified by Chapman and Pearson, Chan et al. (1992) still provides a useful taxonomy of the SDE models that are used to model the short term (riskless) interest rate. Eight such models are identified. These first of these specifications is the absolute Brownian motion used in Merton (1973) to model discount bond prices: $dr = \mu dt + \sigma dW$. This SDE suffers from a number of defects. One obvious limitation is the possibility of paths with negative interest rates. Another limitation is empirical, Chan et al. (p. 1210) find that the "the most important feature differentiating interest models" is the ability to capture the volatility of interest rates which is sensitive to the level of r . By modelling volatility as a constant, absolute Brownian motion does not recognize this empirical result. As such, this SDE is of interest primarily for the simplicity of the process in solving pricing problems.

A more plausible SDE is geometric Brownian motion: $dr = \mu r dt + \sigma r dW$. This process has been used in a number of studies including Rendleman and Bartter (1980). The empirical validity of this process has been questioned for allowing paths where the interest rate increases to levels that are not practical. The process does have desirable features such as sample paths that are only positive and, for option pricing purposes, the ability to access pricing results that have been derived for other cases, e.g., the Black-Scholes European option on a non-dividend paying stock. To value an option written on an interest rate, the Black (1976) version of the Black-Scholes formula is used where the implied forward rate is substituted for the forward price, the strike rate is used for the exercise price and the volatility is for the implied forward rate, e.g., Chance (1998, pp. 590-1).

Though there is considerable debate in economics about whether interest rate processes are unit root or mean reverting, e.g., Wu and Zhang (1996), at least since

Vasicek (1977) modern Finance has favored modelling interest rates as mean reverting processes. The Vasicek mean-reverting model of the short term interest rate has the OU form:

$$dr = \kappa(\mu - r)dt + \sigma dW$$

In this model, μ is the long run mean, κ is the speed of adjustment to the long run mean, and σ is the volatility of the short rate. Vasicek used this model to solve for discount bond prices and for a one factor model of the term structure (see end of chapter questions). Included among a large number of studies that have used the "Vasicek model," Jamshidian (1989) used this process to solve for options on zero coupon and coupon bearing bonds, e.g., Hull (2003, pp. 540–1). Gibson and Schwartz (1990) used this process to value the convenience yield for oil futures and forwards contracts.

While the Vasicek model accounts for the possibility of mean reversion in interest rates, this process suffers from the possibility that some sample paths may contain negative interest rates, though the presence of a mean reverting drift does mitigate this possibility. Another limitation is the assumption of constancy in the volatility. Both of these problems are at least partly addressed in the square root diffusion model employed by Cox, Ingersoll and Ross (1985), often referred to as the CIR model:

$$dr = \kappa(\mu - r)dt + \sigma\sqrt{r} dW$$

CIR used this process to solve for an equilibrium model of the term structure (see end of chapter questions). The empirical properties of the model have been investigated in numerous studies including Gibbons and Ramaswamy (1993), Pearson and Sun (1994) and Lamoureux and Witte (2002). While Pearson and Sun find the CIR model fits poorly, Lamoureux and Witte are able to extend the model to allow for additional "factors" that permit the CIR model to do "a very good job of fitting the time-series properties of short rates."

UNIVARIATE ITO'S LEMMA

Calculus is an important mathematical technique with numerous useful applications. As conventionally presented, calculus is applied to functions that are deterministic. In effect, the familiar rules associated with dy/dx , such as $y = x^2 \rightarrow dy/dx = 2x$, only apply when x is known. When x is a random variable the usual rules of calculus no longer apply. The importance of Ito's lemma is that it specifies the procedures for applying calculus to functions which contain random variables. More precisely, Ito's lemma provides a method for evaluating the total derivative of a function of a stochastic variable which follows a Markov diffusion process. As discussed above, the diffusion class includes a wide range of stochastic processes. Given this, the *univariate* form of Ito's lemma can be stated:¹²

Ito's lemma

Let $u[X, t]$ be a continuous random function mapping from $\mathfrak{R}^1 \times [0, T] \rightarrow \mathfrak{R}^1$ with continuous partial derivatives:

$$u_t = \frac{\partial u}{\partial t} \quad u_x = \frac{\partial u}{\partial x} \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}$$

If $X(t)$ is a random process with a stochastic differential equation obeying a diffusion of the form:

$$dX(t) = a(t)dt + v(t)dW(t)$$

where $W(t)$ is a standard Wiener process and $a(t)$ and $v(t)$ are the drift and volatility of the diffusion, then the function $u(t) = u[X(t), t]$ also has a differential on $[0, T]$ given by:

$$du(t) = \{u_t + u_x a(t) + 1/2 u_{xx} v(t)^2\}dt + u_x v(t)dW(t)$$

This form of Ito's lemma generalizes in a natural fashion to the case where X is multidimensional, e.g., Malliaris and Brock (1982, pp. 85-6).¹³

While Ito's lemma has had numerous applications in financial economics, perhaps the most well known is the Black-Scholes application to call option valuation, where the non-dividend paying stock price $S(t)$ is assumed to follow a log-normal diffusion:

$$dS = \alpha S dt + \sigma S dW$$

In this case, $a(t) = \alpha S$ and $v(t) = \sigma S$ where α and σ are constants. The functional relationship between the call option price (C) and the stock takes the form: $C = C[S(t), t]$. Application of Ito's lemma gives:

$$dC = \{C_t + C_S \alpha S + 1/2 C_{SS} \sigma^2 S^2\}dt + \{C_S \sigma S\}dW$$

This solution plays a central role in the derivation of the Black-Scholes price of a European call. In particular, Black-Scholes are able to use the *riskless hedge portfolio* construction to provide an additional condition that permits elimination of the dW term. In this fashion, the call option pricing problem is transformed from an SDE problem which is not solvable in closed form with standard techniques, into a deterministic partial differential equation (PDE) problem that can be solved. However, complications associated with solving PDE's are such that closed form solutions are not always possible. This has at least two implications. First, there is an emphasis on problem specifications that can be solved in closed form, e.g., Black-Scholes, even though such solutions may not be fully realistic. And, secondly, there is the need to apply numerical methods to "solve" problems in which a precise specification of the problem is required.

Ito's lemma can also be used to transform one type of SDE into another. One useful example involves the function $G[S] = \ln[S]$ where the SDE for S is geometric Brownian motion. Application of Ito's lemma provides the SDE for $G[S]$, specified using the drift and volatility parameters for S . More precisely, $G_t = 0$, $G_S = 1/S$ and $G_{SS} = -(1/S)^2$. Using Ito's lemma it follows that:

$$dG = \{0 + \alpha - 1/2\sigma^2\}dt + \sigma dW = (\alpha - 1/2\sigma^2)dt + \sigma dW$$

The result that the log of a lognormally distributed random variable follows an arithmetic Brownian process is not surprising, but the specification of the drift is not obvious.

A straightforward extension of the univariate approach to the case of functions of two variables (and time) is provided by Fischer (1975) that examines the stochastic behavior of the real bond price $q = B/P$ where, using Fischer's notation, B is the nominal price of a riskless bond and P is the aggregate price level. In this case, the rate of change in the aggregate price level (dP/P) and the return on a zero-coupon continuously compounded nominal bond (dB/B) can be specified:

$$dP/P = \pi dt + \sigma dW \quad dB/B = R dt$$

Observing that when $q = u[B, P] = B/P$:

$$u_t = 0 \quad u_{BP} = -\{1/P\}^2 \quad u_{BB} = 0 \quad u_B = 1/P \quad u_P = -\{B/P^2\} \quad u_{PP} = 2\{B/P^3\}$$

It is now possible to apply the multivariate form of Ito's lemma to get:¹⁴

$$dq = \left\{ \frac{1}{P} RB + \frac{-B}{P^2} \pi P + \frac{B}{P^3} P^2 \sigma^2 \right\} dt + \frac{-B}{P^2} \sigma P dW$$

Factoring out B/P gives the desired result:

$$dq/q = \{R - \pi + \sigma^2\}dt - \sigma dW$$

More developed applications of the multivariate form of Ito's lemma can be found in various sources, e.g., Gibson and Schwartz (1990), Schwartz (1982).

MULTIVARIATE ITO'S LEMMA

It is useful to proceed by example to illustrate the multivariate form of Ito's lemma. Briys and Solnik (1992) provide an interesting application to the case where the random process is multidimensional, involving two random variables (and time). This requires a more involved form of Ito's lemma. The Briys and Solnik example involves the domestic currency value of a foreign security, $V^* = VS$, where V is the random foreign currency value of the foreign security and S is the random spot exchange rate, producing V^* which is the random domestic currency value of the foreign security. In this case, processes are given for dV and dS with dV^* to be calculated using Ito's lemma.

To evaluate Ito's lemma for this case requires application of the associated multivariate total derivative. The resulting solution for the two random variable case, $y(t) = u[t, \{x(t)\}] = u[t, \{x_1(t), x_2(t)\}]$, takes the form:

$$dy = u_t dt + u_x dx + \frac{1}{2} \text{tr}[\Sigma u_{xx}] dt$$

where Σ is the variance-covariance matrix of the state variables and $\text{tr}[\cdot]$ is the trace operator. Because there are n state variables in the general form, u_x is a $n \times 1$ column vector containing the first partial derivatives, dx is a row vector containing the diffusion processes and u_{xx} is a symmetric $n \times n$ matrix of the second partial derivatives.

In the Briys and Solnik example, $dy = dV^*$ and V and S are assumed to follow the log-normal diffusions:

$$\frac{dV}{V} = \mu_V dt + \sigma_V dW_V \quad \frac{dS}{S} = \mu_S dt + \sigma_S dW_S$$

The covariance between dW_S and dW_V per unit time is $\text{cov}(dW_S, dW_V) = \rho_{VS} dt$. It follows that $\sigma_{VS} = \rho_{VS} \sigma_V \sigma_S$. Recognizing that $y = u[x, t]$ in this case is $V^* = VS$, it follows that:

$$\begin{aligned} u_1 &= \frac{\partial V^*}{\partial V} = S & u_2 &= \frac{\partial V^*}{\partial S} = V & u_t &= \frac{\partial V^*}{\partial t} = 0 \\ u_{11} &= \frac{\partial^2 V^*}{\partial V^2} = 0 & u_{22} &= \frac{\partial^2 V^*}{\partial S^2} = 0 & u_{12} &= \frac{\partial^2 V^*}{\partial V \partial S} = 1 & u_{21} &= \frac{\partial^2 V^*}{\partial S \partial V} = 1 \end{aligned}$$

Because there are now two state variables, V and S , the variance-covariance matrix Σ is 2×2 with variances on the diagonal and covariances on the off-diagonal. Remembering from the univariate lognormal example that $v(t)^2 = \sigma^2 S^2$, evaluating $1/2 \text{tr}[\cdot]$ for this case gives:

$$\begin{aligned} \frac{1}{2} \text{tr} \left\{ \begin{bmatrix} \sigma_{11} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{1,2} \\ u_{2,1} & u_{22} \end{bmatrix} \right\} &= \frac{1}{2} \{ \sigma_{11} u_{11} + \sigma_{1,2} u_{2,1} + \sigma_{2,1} u_{1,2} + \sigma_{22} u_{22} \} \\ &= \sigma_{1,2} = \sigma_{VS} VS \end{aligned}$$

Using the result that $1/2 \text{tr}[\cdot] = \sigma_{VS} VS$. The solution to the total derivative is found to be:

$$dV^* = 0 + S(\mu_V V dt + \sigma_V V dW_V) + V(\mu_S S dt + \sigma_S S dW_S) + \sigma_{VS} VS dt$$

$$\frac{dV^*}{V^*} = (\mu_V + \mu_S + \sigma_{VS}) dt + \sigma_V dW_V + \sigma_S dW_S$$

Given dV and dS , this is dV^* , the SDE for the domestic currency value of the foreign security.

*Binomial models and other discrete processes**

Continuous time methods have a number of appealing properties in dealing with problems involving functions of random variables. For example, Ito's lemma provides a method for the intuition and analytical insight of deterministic calculus to be extended to functions of random variables. However, these appealing properties are obtained by making relatively strong restrictions on the structure of the valuation problem. In many cases, the complications associated with the continuous time approach are unnecessary to achieve a solution to the problem at hand. In other cases, the problem at hand is not fully solvable using continuous time methods. In these cases, solutions can be obtained using discrete time processes and the associated techniques of numerical methods. These techniques are particularly useful in valuing securities where the payout is *path dependent*, a property that is applicable to many of the important fixed income securities with embedded options, such as callable bonds and mortgage pass throughs. Path dependence occurs because the realized value of such securities depend on the specific stochastic path taken by, say, interest rates.

The simplest form of a discrete time process is the binomial process, sometimes called a two state process. As depicted in figure 5.3, given an initial starting point, the binomial process proceeds by specifying the probability of an up (down) move. Because there are only two possible states of the world, the probability of a down (up) move is determined by observing that the sum of the probabilities equals one. The binomial process is characterized by the number of periods or steps that are evaluated and by whether the process is *recombining* or *non-recombining*. A recombining process has the property that an up move followed by a down move will end up at the same outcome as a down move followed by an up move. For a two period recombining binomial process there will be three possible outcomes while for a two period non-recombining process there will be four possible outcomes. The computational advantages of a recombining process are considerable if the number of periods is large. In addition, there are absence of arbitrage restrictions that need to be considered in determining the probabilities of the binomial process.

In figure 5.3, the state variable used is the bond price. Specifying the appropriate state variable for modeling fixed income securities is not as obvious as for equity securities. Consider the case of bond valuation. It is possible to use the bond price as the state variable but, unlike equities, a bond has a fixed maturity date. As the maturity date is approached, the price will converge to the maturity value. This convergence property will affect the volatility along the binomial paths. This difficulty is avoided if the short term interest rate is used as the state variable. In a one factor model of the term structure, if enough periods are evaluated then the prices of zero coupon bonds can be determined from the short term interest rate paths. Following the discussion in section 4.3, this is typically done by assuming the unbiased expectations hypothesis and treating the future short term interest rates as implied forward rates. This procedure is complicated by the need to ensure that the interest rate paths do not admit arbitrage opportunities. This can happen, for example, if the future interest rates are not restricted from attaining negative values.

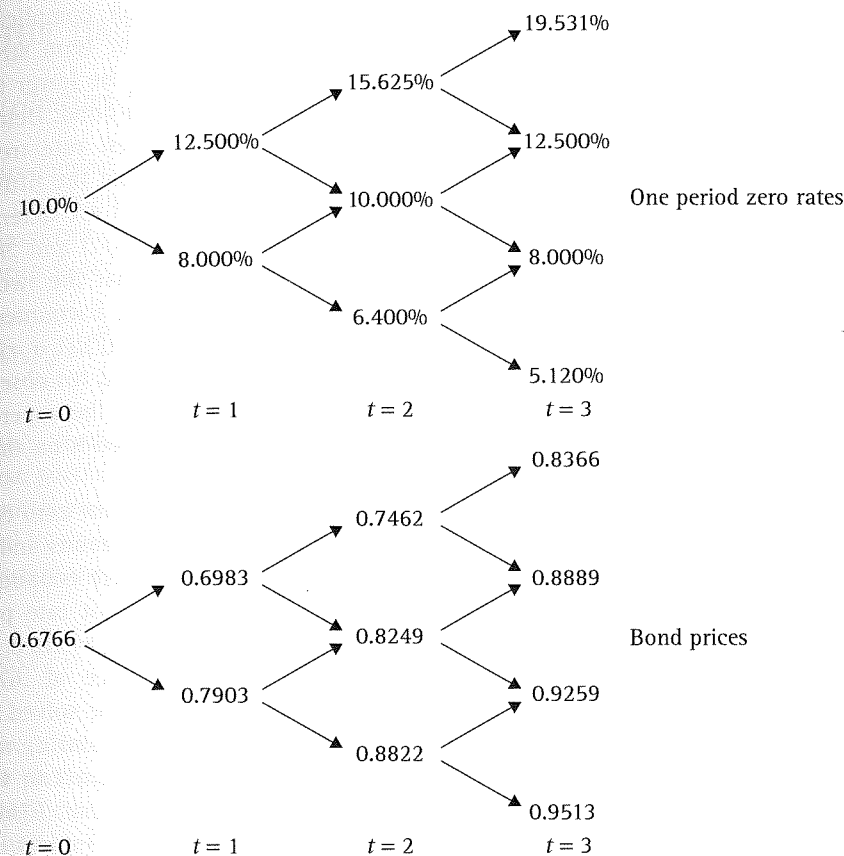


Figure 5.3 Binomial process of the one-period zero and associated bond prices

Following Sundaresan (2002, pp. 599–604), the upper tree in figure 5.3 illustrates a 3 year evolution of a recombining binomial process for one year zero coupon interest rates. At $t = 0$, the one period interest rate is 10% and it is assumed that the probability of rates increasing or decreasing is 50%. If rates increase then the new rate will be 1.25 times the initial rate and if rates go down the factor is 0.8. As indicated in figure 5.3, this implies that an up move will produce a rate of 12.5% in one year and a down move will produce a rate of 8%. The probabilities and change factors are selected to produce a recombining tree. More complicated discrete processes, such as the self-calibrating processes, use more complicated methods of determining the size of the changes and, possibly, the probabilities. For example, if the process is restricted to be mean reverting the probabilities of up (down) moves will be reduced as the process moves farther above (below) the long run mean, i.e., the probabilities will also evolve over time (e.g., Sundaresan 2002, p. 607). Another example occurs in the Black, Derman and Toy model where the size of the changes are set using observed interest rate volatility and further subjected to an absence of arbitrage condition.

Even though figure 5.3 does not incorporate all the various possible complications, it does capture essential elements of the procedure for modelling the term structure using discrete time models. Given the binomial interest rate process, the lower tree in figure 5.3 provides the evolution of the price of a 4 year zero coupon bond. Determining bond prices involves starting from the last nodes of the interest rate tree and working backward to determine the various bond prices. Observing that the fifth node (not given) is the $t = 4$ maturity value of \$1, the four nodes at $t = 3$ are determined by evaluating $P(3, 4) = 1/(1 + r(3, 4))$. For example, $.8366 = 1/(1.19531)$ and $.9513 = 1/(1.0512)$. The $t = 2$ values can now be calculated by observing that the two connected $t = 3$ nodes are equally likely and that the expected value can be discounted back at the rate applicable at that node, i.e., $P(2, 4) = (.5P(3, 4)^u + .5P(3, 4)^d)/(1 + r(2, 3))$, where u and d represent up and down moves. This condition is required for *absence of arbitrage*. For example, $.746162 = (.5(.8366) + .5(.8889))/(1.15625)$ gives the upper value on the bond price tree. This process is repeated to determine all the nodes on the zero coupon bond price tree.

Given the zero coupon bond price process, it is possible to calculate the zero coupon yield curve that, in turn, can be used to construct the price of coupon bonds and other types of fixed income securities. To see this let $P(0, T)$ represent the zero coupon bond price observed at $t = 0$ and maturing at $t = T$. For the binomial spot rate process it follows that $P(0, 4) = .6766 = 1/(1 + z_4)^4$ which can be solved to get $z_4 = .1026$. Using the same process that was used to construct the 4 year bond price of .6766, a three period tree can be constructed to get $P(0, 3) = .7475$ which can be solved to get $z_3 = .10187$ and again for a two period tree to get $P(0, 2) = .8249$ and $z_2 = .101$. Together with the initial starting value of $z_1 = .10$ the term structure from 1 to 4 years has been constructed. Following the discussion in section 4.3 observe that $.6766 = 1/\{(1.1)((1.101)^2/(1.1))((1.10187)^3/(1.101)^2)((1.1026)^4/(1.10187)^3)\} = 1/\{(1 + z_1)(1 + f_{1,2})(1 + f_{2,3})(1 + f_{3,4})\}$, i.e., the implied forward rates can be determined from the zero coupon interest rates. This absence of arbitrage condition is ensured by the method used to construct the binomial process. Sundaresan (2002, p. 604) demonstrates that the limit of the binomial process in figure 5.3 is lognormal, a process that does not have the most desirable properties for modelling the term structure.

The observation that the binomial process can be taken to the limit to derive the continuous time representation of the discrete process is significant. It recognizes the connection between the continuous and discrete approaches to modeling interest rate processes. The discrete and continuous models are complementary, not contradictory. Continuous time processes can lead to powerful analytical results but are difficult to apply in some situations. Binomial processes are useful in constructing solutions where continuous time methods are intractable. It is possible to extend the basic binomial model captured in figure 5.3 in a number of ways. In addition to the self-calibrating models, the binomial process can be used as the basis for Monte Carlo simulations (see section 6.3). In these models, a large number of binomial paths are generated using, say, a particular volatility assumption to generate the size of the steps. These various paths can be used to identify exercise decisions associated with embedded options. The specification of exercise decisions is explored further in chapter 6.

5.3 THE TIME VALUE-CONVEXITY TRADEOFF[#]

The contribution from convexity

The introduction of the convexity concept generated considerable initial enthusiasm among practitioners, e.g., Grantier (1988), Bierwag (1987, p. 306). The basis for this enthusiasm can be illustrated by examining the Taylor series expansion for the univariate bond price function $P_B[y]$:

$$\% \Delta P_B = \frac{P_B[y] - P_B[y_0]}{P_B[y_0]} \cong -DUR(y - y_0) + \frac{1}{2} CON(y - y_0)^2$$

Given this, consider a comparison of two bond portfolios (A and B), with values P_A and P_B . These portfolios are constructed to have equal duration ($D_A = DUR_A = DUR_B = D_B$), equal initial yield to maturity (y_0) and $CON_A > CON_B$. The initial durations and yields are set equal by appropriate selection of the bonds included in the two portfolios. The impact of an instantaneous interest rate change on these two portfolios would be approximately:

$$\% \Delta P_A - \% \Delta P_B \cong \frac{1}{2} (CON_A - CON_B)(y - y_0)^2$$

Observing that $CON > 0$ because the bond pricing function(s) are convex, it follows that whether yields go up or down, the portfolio with the higher convexity will have a better percentage change in price. This effect is further illustrated in figure 5.4.

To illustrate how the Taylor series expansion works to construct a function, consider table 5.4 which evaluates the impact of interest rate changes varying from 1 to 300 basis points on the price of a 20 year, 10% coupon par bond. Both increases and decreases in yield are considered. Because there is only one bond involved, it is straightforward to calculate the actual percentage price changes. Consistent with the convexity of the price function for a straight bond, the percentage change in price for a given change in yield is greater for a decrease in yield than for an increase in yield, e.g., -21.22% for a 300 bp increase versus +32.03 for a 300 bp decrease. Consistent with duration involving the derivative of the bond price function at a point, the price impact of small changes in yield are accurately estimated by duration. However, for changes as small as 10 bp there is some convexity impact. Consistent with convexity involving the second derivative, the impact of convexity on the (convex) default and option free bond price function is always positive.

Table 5.4 provides useful information about the properties of the Taylor series expansion of a convex function such as $P_B[y]$. In particular, for a given change in yield, the absolute value of the predicted change due to duration is the same for up moves as down moves. As a consequence, duration over-predicts the price impact of yield increases and under-predicts for yield decreases. Addition of the convexity term more than compensates for the yield increase case and does not sufficiently compensate for yield decrease. The result is that the first two terms of the Taylor series expansion do

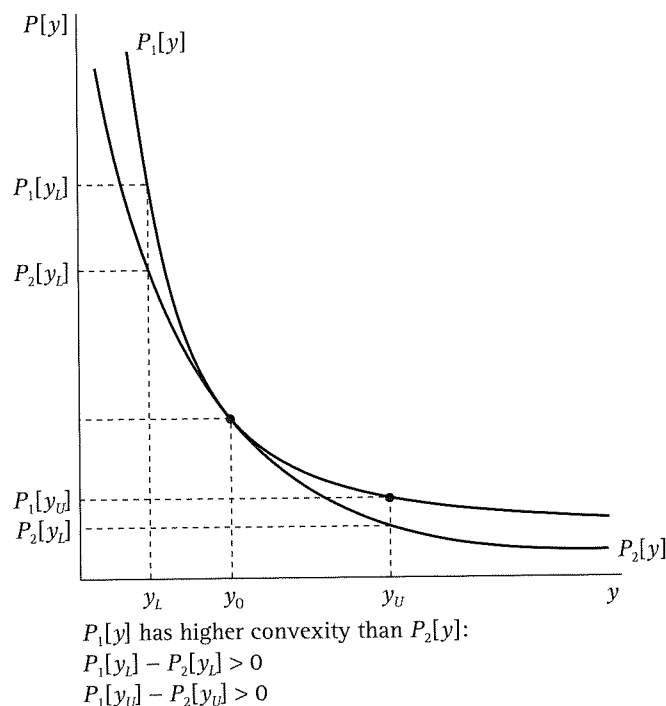


Figure 5.4 Higher convexity provides superior return

not completely explain all of the percentage change in price. Actual evaluation of the expansion follows from observing that $D = 8.5795$ and $CON = 117.478$. It follows that for, say, a 100 bp increase:

$$\begin{aligned}\% \Delta P_B &= -D(.01) + \frac{1}{2} CON(.01)^2 = -8.58(.01) + \frac{1}{2} 117.478(.01)^2 \\ &= -8.58\% + .5873\% = 7.99\%\end{aligned}$$

This value compares to the actual price change of 8.023%.

To provide an illustration of the impact of convexity, consider the gains associated with the following self-financing portfolio:

- 1 Purchase one pure discount (zero coupon) bond with a 20 year maturity, with par value of \$1 million.
- 2 Sell short 1.92 pure discount bonds with a 5 year maturity and par value of \$1 million per unit. The full value of the short sale proceeds is immediately available.
- 3 Invest \$1,127,490 in bonds with zero duration (cash has $D = 0$).

Without loss of generality, assume that the yield to maturity on all bonds is 5% initially. What is the initial investment in the portfolio? From an inspection of the balance sheet, it is apparent that no investment is required in the portfolio:

Table 5.4 Example of the % price change using modified duration and convexity for a 20-year 10% coupon bond selling at par to yield 10%

<i>Increase in required yield</i>						
Required yield	10.01%	10.10%	10.50%	11.00%	12.00%	13.00%
Change in basis points	1	10	50	100	200	300
Estimated % change						
Duration	-0.09	-0.86	-4.29	-8.58	-17.16	-25.74
Convexity	0.00	0.01	0.14	0.57	2.27	5.10
Total	-0.09	-0.85	-4.15	-8.01	-14.89	-20.64
Actual % change	-0.09	-0.85	-4.15	-8.02	-15.05	-21.22
<i>Decrease in required yield</i>						
Required yield	9.99%	9.90%	9.50%	9.00%	8.00%	7.00%
Change in basis points	-1	-10	-50	-100	-200	-300
Estimated % change						
Duration	0.09	0.86	4.29	8.58	17.16	25.74
Convexity	0.00	0.01	0.14	0.57	2.27	5.10
Total	0.09	0.87	4.43	9.15	19.43	30.84
Actual % change	0.09	0.86	4.44	9.20	19.79	32.03

Assets		Liabilities plus net worth
20 year bond	$\$376,890 = 1 \text{ million}/(1.05)^{20}$	Short sale $\$1,504,370 = 1.92 \text{ million}/(1.05)^5$
Cash	$\$1,127,480$	Net worth = \$0 (Self Financing)
Total:	$\$1,504,370$	

Computing the Macaulay duration of the assets and of the liability reveals that $D_A = (376,890/1,504,370)(20) + 0 = 5 = D_L$. Treating the liabilities as one portfolio and the assets as another, the position conforms to the requirement of equal duration and yield. Calculation of the convexity reveals that the convexity of the assets is $(21/6)$ times larger than the convexity of the liabilities. The portfolio is immunized according to the classical Redington conditions.

Consider what happens if the yield to maturity for all assets and liabilities immediately rises to 6%. In this case, the net value of the portfolio (value of assets less value of liabilities) changes to:

Assets		Liabilities plus net worth	
20 year bond	$\$311,800 = 1 \text{ million}/(1.06)^{20}$	Short sale	$\$1,434,740 = 1.92 \text{ million}/(1.06)^5$
Cash	<u>$\\$1,127,480$</u>	Net worth	<u>4,540</u>
Total:	$\$1,439,280$		$\$1,439,280$

Now consider what happens if the yield to maturity of all assets and liabilities immediately falls to 4%:

Assets		Liabilities plus net worth	
20 year bond	$\$456,390 = 1 \text{ million}/(1.04)^{20}$	Short sale	$\$1,573,100 = 1.92 \text{ million}/(1.04)^5$
Cash	<u>$\\$1,127,480$</u>	Net worth	<u>10,770</u>
Total:	$\$1,583,870$		$\$1,583,870$

This example illustrates that basic point: starting from a position of equal duration and yield for the assets and liabilities of the portfolio, whether yields increased or decreased the net worth of the portfolio increased.

Given this background, the initial enthusiasm for convexity is understandable. It appears as though there is a free lunch gain to be achieved from portfolios with higher convexity. In the context of portfolio immunization, this translates into a recommendation for using barbell asset positions to immunize single cash flow liabilities. The resolution of this apparent free lunch can be found in the time value-convexity tradeoff. This tradeoff can be illustrated by considering what happens to the portfolio in the example if yields remain unchanged. Consider the value of the portfolio in 1 year if yields are the same as on the initial portfolio creation date:

Assets		Liabilities plus net worth	
20 year bond	$\$395,734 = 1 \text{ million}/(1.05)^{19}$	Short sale	$\$1,579,590 = 1.92 \text{ million}/(1.05)^4$
Cash	<u>$\\$1,127,480$</u>	Net Worth	<u>$(\\$55,276)$</u>
Total:	$\$1,524,214$		$\$1,524,214$

This illustrates the role of convexity. Setting the durations of the portfolios equal, higher convexity portfolios outperform low convexity portfolios whether interest rates go up or down. However, if rates do not change then higher convexity portfolios will underperform lower convexity portfolios. The large time value impact arises because this example does not adequately impose the condition that the initial yields of the two portfolios be equal (see end of chapter questions).

The Taylor series representation of the time value-convexity tradeoff*

The multivariate Taylor series expansion of the function $P_B[y, t]$ can be manipulated to produce the result:

$$\begin{aligned} \frac{P_B[y, t] - P_B[y_0, t_0]}{P_B[y_0, t_0]} &= \frac{1}{P} \frac{\partial P}{\partial y} (y - y_0) + \frac{1}{P} \frac{\partial P}{\partial t} (t - t_0) \\ &+ \frac{1}{2} \left\{ \frac{1}{P} \frac{\partial^2 P}{\partial y^2} (y - y_0)^2 + \frac{1}{P} \frac{\partial^2 P}{\partial t^2} (t - t_0)^2 \right\} \\ &+ \frac{1}{P} \frac{\partial^2 P}{\partial y \partial t} (y - y_0)(t - t_0) + \dots H.O.T. \\ \% \Delta P &\equiv -DUR(y - y_0) + \theta(t - t_0) \\ &+ \frac{1}{2} [CON (y - y_0)^2 + \Theta_2(t - t_0)^2] + CROSS(y - y_0)(t - t_0) \end{aligned}$$

where $DUR = D$, θ is the theta of the bond, Θ_2 relates to the second derivative with respect to time and $CROSS$ relates to the cross product terms. In this presentation, the sign of θ requires interpretation. Much like an option, time counts backwards for a bond. Yet, as time increases, the time to maturity goes down and the value of the bond position increases, i.e., the time derivative is positive. Unlike options where the premium value falls as time increases, for θ to be taken as a positive value it is not necessary to define θ by multiplying by minus one.¹⁵

Evaluation of the relevant coefficients in the expansion is simplified considerably by using par bonds. This assumption permits the appropriate derivative to be used without having to deal with the valuation of the bond. For example, consider a semi-annual coupon par bond with $C = \$8$ ($C/2 = \$4$), $M = \$100$, $y = .08$ ($y/2 = .04$), $T = 5$ years:

$$P_B[y, t] = C \left[\frac{1}{\frac{y}{2}} - \frac{1}{\frac{y}{2} \left(1 + \frac{y}{2}\right)^{2T}} \right] + \frac{M}{\left(1 + \frac{y}{2}\right)^{2T}}$$

$$\frac{1}{P} \frac{\partial P}{\partial \left(1 + \frac{y}{2}\right)} = \frac{-811.09}{100} \text{ (in half years)}$$

$$\rightarrow D = \frac{8.1109}{2} = 4.055 \rightarrow D^* = D(1.04) = 4.218$$

Evaluation of convexity follows appropriately:

$$\frac{1}{P} \frac{\partial^2 P}{\partial \left(1 + \frac{y}{2}\right)^2} = \frac{8,075.43}{100} \text{ (in half years}^2\text{)} \rightarrow CON = \frac{8,075.43}{4(100)} = 20.1886 \text{ (in years)}$$

Division of the convexity value by four (instead of 2 as in the case of the calculation of duration for semi-annual coupon bonds) is because the derivative is expressed in units of half years squared.

Evaluation of theta is somewhat different. Duration and convexity deal with the change in price for an instantaneous change in yield. Theta deals with the change in value as time passes. Instead of the price impact, theta is concerned with the return earned from holding the bond over a given time period:

$$V[y, t] = P_B = \frac{C/2}{1 + y/2} + \frac{C/2}{(1 + y/2)^2} + \dots + \frac{M + C/2}{(1 + y/2)^{2T}}$$

$$-\frac{\partial V}{\partial (T - t)} = \frac{\partial V}{\partial t} = \ln \left[1 + \frac{y}{2} \right] (M + C/2) \text{ (in half years)}$$

This result applies to all bonds, independent of the term to maturity. Evaluating θ for the par bond example given previously $y/2 = .04$, $C/2 = \$4$ gives a half year value of $\theta = 0.04079$. In effect, the time value for a par bond is approximately equal to the coupon payment that is earned from holding the bond. For a non-par bond there will be an additional term associated with the capital gain (loss) due to the price change associated with the impact of time change on a discount (premium) bond, holding interest rates constant.

Adapting an exercise conducted in Chance and Jordan (1995), it is possible to compare the relative contributions for each component of the Taylor series to changes in price for various assumptions about changes in yield and time. Table 5.5 provides a benchmark table that can be used to compare with the results of table 5.4. Table 5.5 provides the values calculated above for D , CON and θ for the four par bond maturities being considered. The initial yield to maturity is 8% for all bonds. The percentage changes in price result from an instantaneous change in yield at all maturities; terms involving changes in time have no impact. The yield change considered is a 10 basis point increase. In table 5.4 it was demonstrated that, for such a small yield change, the D and CON terms in the Taylor series explain virtually all of the price change. Table 5.4 also illustrates that duration will over predict the fall in price due to the increase in yields and the positive impact of CON will offset this over prediction.

These results from table 5.4 are also indicated in table 5.5 where the prediction from D is more than the actual price change, i.e., values of $(D\Delta y/\% \Delta P) > 100\%$. For example, consider the 5 year bond, $D = 4.2177/1.04 = 4.05548$. The predicted $\% \Delta P = 4.05548(.001) = .00405548$. The actual $\% \Delta P$ is $99.5955 - 100/100 = -.004045$. Evaluating the absolute value of the predicted change divided by the actual change gives: $.00405548/.004045 = 1.00259 = 100.259\%$. Similarly, for the convexity adjustment: $\frac{1}{2} CON(\Delta y^2) = \frac{1}{2}(20.819)(.00001)$. Evaluating the convexity component divided by the actual price change gives: $-0.0000104095/.00405548 = -.2573\%$. The sum of

Table 5.5 Decomposition of the Taylor series expansion: default free par bonds from 1 year to 30 years to maturity*

Bond maturity (years)	1	5	15	30
Duration (D^*)	0.9808	4.2177	8.9919	11.7642
Convexity	1.3512	20.819	104.978	214.24
Theta (θ)	0.04079	0.04079	0.04079	0.04079
New price	99.9058	99.5955	99.1406	98.8795
$\% \Delta P$	-0.0009	-0.0041	-0.0086	-0.0112
$-(D\Delta y)/\% \Delta P$	100.0717	100.259	100.608	100.9486
$(\frac{1}{2} CON \Delta y^2)/\% \Delta P$	-0.0717	-0.2573	-0.6108	-0.956

* $\Delta y = 10$ basis points, $\Delta t = 0$

Table 5.6 Decomposition of the Taylor series expansion: default free par bonds from 1 year to 30 years to maturity*

Bond maturity (years)	1	5	15	30
Duration (D^*)	0.9808	4.2177	8.9919	11.7642
Convexity	1.3512	20.819	104.978	214.24
Theta (θ)	0.04079	0.04079	0.04079	0.04079
New price	99.9275	99.6174	99.1622	98.901
$\% \Delta V$	-0.0007	-0.0038	-0.0084	-0.011
$-(D\Delta y)/\% \Delta V$	130.0734	106.006	103.1982	102.9245
$(\theta \Delta t)/\% \Delta V$	-29.642	-5.646	-2.5651	-1.9554
$-(Cross \Delta y \Delta t)/\% \Delta V$	-0.3354	-0.040	-0.0093	-0.0019
$(\frac{1}{2} CON \Delta y^2)/\% \Delta V$	-0.0932	-0.2637	-0.625	-0.9747
$(\frac{1}{2} \Theta_2 \Delta t^2)/\% \Delta V$	-0.0032	-0.0006	-0.0003	-0.0002

* $\Delta y = 10$ basis points, $\Delta t = 1$ day

the duration and the convexity terms in the Taylor series reproduces the actual price change almost exactly. Such calculations depend on the size of the assumed change in yield. Increasing the change to, say, 100 basis points will reduce the accuracy of the approximation.

Extending table 5.5 to incorporate a one-day change in time produces the results in table 5.6. Converting θ to an annual basis and evaluating the change in value for one day gives $.08/365 = .000219$. Observing the par value is \$100 gives the difference in the "new prices" between tables 5.5 and 5.6. For example, for the five year par bond the actual price change in table 5.5 was $-.4045$. Taking into account the time value in table 5.6 gives an actual price change of $-.4045 + (.000219)(100) = -.3826$. This gives a new price of $100 - .3826 = 99.6174$. The various contributions follow appropriately: $(.00405548)/(.003826) = 106.006\%$. Recognizing the convention that theta has been defined as a positive value, the time value component can be calculated as $.000216/.003826 =$

Table 5.7 Decomposition of the Taylor series expansion: default free par bonds from 1 year to 30 years to maturity*

Bond maturity (years)	1	5	15	30
Duration (D^*)	0.9809	4.2177	8.9919	11.7642
Convexity	1.3512	20.819	104.978	214.24
Theta (θ)	0.04079	0.04079	0.04079	0.04079
Δy (%)	0.6500	0.4700	0.3875	0.3500
New price	100.084	98.7873	97.3825	96.8174
$\% \Delta V$	0.0008	-0.0121	-0.0262	-0.0318
$-(D\Delta y)/\% \Delta V$	-729.3730	157.1763	127.9986	124.3981
$(\theta \Delta t)/\% \Delta V$	767.1437	-53.1648	-24.6315	-20.2577
$-(\text{Cross} \Delta y \Delta t)/\% \Delta V$	56.4215	-2.0496	-0.3448	-0.0671
$(\frac{1}{2} \text{CON} \Delta y^2)/\% \Delta V$	3.3966	-1.8388	-3.0111	-4.1231
$(\frac{1}{2} \Theta_2 \Delta t^2)/\% \Delta V$	2.473	-0.1714	-0.0794	-0.0653

* Δy = one standard deviation (historical), Δt = 30 days

5.646%. Table 5.6 illustrates the importance of the time value component relative to the convexity component. Even for a 30 year bond, the one day impact of interest income exceeds the contribution of convexity. For the type of scenario considered in table 5.6, the contributions from *CROSS* and Θ_2 are insignificant.

Though interesting, table 5.6 poses the difficulty that the values used for the time and yield change were arbitrary. For example, yield changes for short term bonds were assumed to be the same as for long term bonds when, in practice, short term yields are typically more variable than long term yields. In addition, bond trading horizons are often longer than one day. Table 5.7 extends the analysis to incorporate a longer period of time, 30 days, and yield changes that are one standard deviation in size for that time interval, i.e., are typical yield changes. One interesting feature of table 5.7 is the implication of the offsetting effects for time and yield on the results for the 1 year bond. The percentage change in the value of the position is so small, due to the offsetting impacts, that the individual duration and theta components are over seven times larger. Comparing the relative sizes of these two components reveals that for a 1 year bond the impact of a 30-day theta move exceeds the impact from a one-standard deviation duration move. In addition, the *CROSS* term is much larger than the convexity component, a result that extends to the 5 year bond. Even Θ_2 rivals convexity in size for the 1 year bond.

Another interesting result in table 5.7 is that, for a 30 year bond, the theta component of the change in value is about five times larger than the convexity value. For the 15 year bond theta is about eight times larger. Even when a two standard deviation change is considered (not reported), the theta term for the 30 year bond is still larger (9.55 to 7.75) than the convexity impact. This confirms results presented in Kang and Chen (2002) for Treasury bonds and other government bonds that demonstrated a greater role for theta than convexity in the changing value of bond positions. Table 5.7 also illustrates that

for the longer term bonds, Θ_2 and $CROSS$ are insignificant for both one or two standard deviation change scenarios. The upshot of the information contained in table 5.7 is that the shape of the yield curve imbeds a cost of convexity that appears as a time value. Setting duration and yield for two portfolios to be equal will have a time value cost that will be traded off against the convexity value. This point is discussed further below.

*A continuous time solution of the time value-convexity tradeoff**

As demonstrated by Ingersoll (1987, pp. 348-9) and others, Ito's lemma can be derived by applying a Taylor series expansion to an appropriately specified function of a random variable. Given this, it is not surprising that the continuous time solution has a close resemblance to the multivariate Taylor series result. In a one-factor model, this approach starts with an assumption about the SDE driving the spot rate. In general, the SDE takes the form:

$$dr(t) = \mu_r[r, t]dt + \sigma_r[r, t]dW(t)$$

Following the discussion in section 4.3, this one factor interest rate model can be used to generate a set of current spot interest rates. It follows that $P_B = P[r, t]$ and application of Ito's lemma leads to:

$$\frac{dP[r, t]}{P[r, t]} = \mu_p[r, t]dt + \sigma_p[r, t]dW(t)$$

where:

$$\mu_p[r, t] = \frac{1}{2}\sigma_r^2 CON - \mu_r[r, t]D + \theta \quad \sigma_p[r, t] = \sigma_r[r, t]D$$

In words, Ito's lemma has been used to derive the SDE for the bond price from the SDE for the interest rate process.

The next step in the derivation of the continuous time solution to the tradeoff between time value and convexity involves a result that is often used in deriving option pricing models, e.g., Christensen and Sorensen (1994). For example, Garman and Kohlhagen (1983) use this result in deriving the pricing formula for a European currency option. The result can be stated:

$$\frac{\mu_p[r, t] - r(t)}{\sigma[r, t]} = \lambda[r, t]$$

where $\lambda[r, t]$ is the risk premium, measured per unit of standard deviation, that applies to all assets. Using absence of arbitrage arguments, it is possible to demonstrate that this λ is the same for all assets, e.g., Ingersoll (1987). Given the specification of λ , it is possible to substitute in the coefficients from the SDE for the bond price to produce a partial differential equation (PDE) that may be solvable.

Substituting the SDE coefficients into the definition for λ gives the PDE:

$$\frac{1}{2}\sigma_r^2 \frac{\partial^2 P}{\partial r^2} + [\mu_r - \lambda\sigma_r] \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0$$

$$\rightarrow \frac{\sigma_r^2}{2} CON - [\mu_r - \lambda\sigma_r]D + \theta = r$$

This PDE specification is sufficient to define a tradeoff between convexity and time value. To do this, recognize that P can be specified to refer not only to individual bonds. More generally, P can also refer to portfolios of fixed income securities. With appropriate adjustments, it is possible for these portfolios to contain bonds with embedded options or that are subject to default risk, but this extension will not be considered here. Rather, consider the exercise of comparing two portfolios that are specified to have the same duration. It is possible for each of these portfolios to contain a single bond.

Examining the PDE it is apparent that the coefficients involving r will be the same for different P because the stochastic process for interest rates is the same for fixed income securities. Consider the relationship between the PDE's for two distinct portfolios. Because the PDE's for both portfolios are equal to r , it follows:

$$\frac{\sigma_r^2}{2} CON_1 - [\mu_r - \lambda\sigma_r]D_1 + \theta_1 = r = \frac{\sigma_r^2}{2} CON_2 - [\mu_r - \lambda\sigma_r]D_2 + \theta_2$$

For two portfolios with equal durations, $D_1 = D_2$, it follows that:

$$\frac{\sigma_r^2}{2} [CON_1 - CON_2] = \theta_2 - \theta_1$$

This is the continuous time version of the tradeoff between convexity and time value.

A number of different interpretations can be made of the tradeoff. One interpretation is straight forward: given two portfolios with equal durations, it does not follow that the portfolio with the higher convexity will be necessarily outperform the lower convexity portfolio. Higher convexity comes at the cost of lower time value. The higher convexity portfolio will outperform if the interest rate changes are sufficiently large. If interest rates do not change sufficiently then the lower time value will cause the higher convexity portfolio to underperform. The level, slope and shape of the yield curve at a given point in time reflect the prevailing tradeoff between time value and convexity. Yet, the different possible scenarios for the future term structure make it difficult to infer from a given yield curve shape precise predictions about future yield curve shape and, as a consequence, the relative performance of high and low convexity portfolios.

To see this, compare a barbell portfolio that combines a 1 year and 15 year bond with a portfolio that holds only a 5 year bond. If the yield curve is upward sloping, equating the durations of these two portfolios will involve holding a larger fraction of the portfolio value in the 1 year bond than in the 15 year bond. It follows that the barbell portfolio, which has a higher convexity, will have a lower time value than the portfolio with the 5 year bond. Hence, the shape of the yield curve at a given point in time

represents the time value cost of convexity. By introducing fixed income derivative securities such as call options (for convexity adjustment) and futures (for duration adjustment), it may be possible to achieve both duration and convexity matching. Assuming that the one factor model of the term structure is valid, then such duration and convexity matching strategies will also be time value matching.

Another interpretation of the tradeoff can be made by observing that setting the durations of two portfolios equal will produce a given difference in convexity. Though it is possible to construct a large number of possible portfolios that will match the duration of a given portfolio, once a specific matching portfolio is selected the difference in convexities is given. As such, there will then be a tradeoff between the volatility of interest rates and the time value. In effect, the tradeoff involves not just convexity and time value but also interest rate volatility. At higher levels of interest rate volatility, the cost of convexity will be higher in terms of the time value that has to be paid. Following Phoa (1997), it is possible to invert this relationship to seek an explanation for the commonly observed hump-shaped yield curve where the yield on 30 year Treasuries is below the yield on, say, 20 year Treasuries. Because the 30 year Treasury has a higher convexity, the degree of hump in the yield curve could be due to the costs of convexity and used as a measure of the market's expectations of future interest rate volatility.

5.4 IMMUNIZATION WITH NON-PARALLEL YIELD CURVE SHIFTS*

*Partial durations and convexities**

Redington (1952) recognized that classical immunization theory may fail when shifts in the term structure are not parallel.¹⁶ At least since Bierwag (1977), this failure has led to the development of alternative concepts of fixed income portfolio management such as the maxmin portfolio, e.g., Bierwag and Khang (1979), Prisman (1986), and the recognition of a connection between immunization strategy specification and the type of assumed shocks, e.g., Fong and Vasicek (1984), Chambers et al. (1988).¹⁷ Risk measures associated with convexity, such as M^2 , were developed to select the best duration matching portfolio from the set of potential portfolios. Being derived using a specific assumption about the stochastic process generating the term structure, these theoretically attractive models encountered difficulties in practice. In particular, "minimum M^2 portfolios fail to hedge as effectively as portfolios including a bond maturing on the horizon date" (Bierwag et al. 1993, p. 1165).¹⁸ This empirical research led to the following result:

using different data bodies, different term structure estimation procedures, different duration measures, and different time periods – the duration-matching portfolio containing a bond having a maturity matching the planning period length has a smaller portfolio-promised return deviation than portfolios not containing a maturity-matching bond. This empirical regularity is so pervasive that we may view it as an appropriate risk-minimizing duration-matching portfolio strategy. (Bierwag et al. 1993, p. 1148)

This empirical result is referred to as the "duration puzzle." Is this puzzle due to failings of the stochastic process assumption underlying the theoretically derived immunization measures or is there some deeper properties of the immunization process which are not being accurately modelled?

Instead of assuming a specific stochastic process and deriving the optimal immunization conditions, it is possible to leave the process unspecified and work directly with the properties of the expansion of the spot rate pricing function, i.e., employ directional derivatives. The use of directional derivatives to derive worst possible shocks and the associated upper bound on the loss of an "immunized" portfolio is found in Reitano (1991b, 1992), Bowden (1997) and Balbas and Ibanez (1998). This line of development has the attractive feature of allowing the introduction of partial durations and convexities associated with the relevant spot interest rates. In Reitano (1991b, 1992), the bounds are derived from Cauchy-Schwarz and quadratic form restrictions inherited from the normed vector space. A closely related problem is adequately solved by Bowden (1997) through the use of the Gateaux differential.

Reitano (1991a,b, 1992, 1996) provides a seminal, if not widely recognized, analysis of duration and bond portfolio immunization. Reitano evaluates a multivariate Taylor series for the asset and liability price functions specified using spot interest rates. Defining a norm applicable to a parallel yield curve shift, Reitano is able to identify bounds on the possible deviations from classical immunization conditions. In other words, even though classical immunization conditions are violated for non-parallel yield curve shifts, it is still possible to put theoretical bounds on the deviations from classical conditions. In addition, it is also possible to identify the specific types of shifts that represent the greatest loss or gain. Developing the Gateaux differential approach introduced by Bowden (1997), Balbas and Ibanez (1998) rediscover the possibility of defining such bounds, albeit in a more sophisticated mathematical framework. Balbas and Ibanez also introduce a novel innovation: a linear dispersion measure that, when minimized, permits identification of the "best" portfolio within the class of immunizing portfolios. More precisely, a strategy of matching duration and minimizing the dispersion measure identifies the portfolio that will minimize immunization risk and, as a consequence, provides an optimal upper bound for possible loss on the portfolio.

The multivariate Taylor series expansion can be applied to motivate the directional derivative approach to fixed income portfolio immunization. Recognizing that P_B is a function of the T spot interest rates contained in $z = \{z_1, z_2 \dots z_T\}$, it is possible to apply a multivariate Taylor series expansion to this bond price formula, that leads to the concepts of partial duration and partial convexity:¹⁹

$$\begin{aligned}
 P(z) &= P(z_0) + \sum_{i=1}^T \frac{\partial P(z_{i,0})}{\partial z_i} (z_i - z_{i,0}) + \frac{1}{2!} \sum_{i=1}^T \sum_{j=1}^T \frac{\partial^2 P(\cdot)}{\partial z_i \partial z_j} (z_i - z_{i,0})(z_j - z_{j,0}) + H.O.T. \\
 \rightarrow \frac{P(z) - P(z_0)}{P(z_0)} &\cong - \sum_{i=1}^T D_i (z_i - z_{i,0}) + \frac{1}{2!} \sum_{i=1}^T \sum_{j=1}^T CON_{i,j} (z_i - z_{i,0})(z_j - z_{j,0}) \quad (1)
 \end{aligned}$$

where $z_0 = (z_{1,0}, z_{2,0}, \dots, z_{T,0})'$ is the $T \times 1$ vector of initial spot interest rates, D_i is partial duration associated with z_i , the spot interest rate for time t , and $CON_{i,j}$ is the partial

convexity associated with the spot interest rates z_i and z_j for i, j defined over $(1, 2, \dots, T)$.²⁰ This formulation is useful for examining the impact of non-parallel yield curve shifts on fixed income portfolio strategies.

Observing that the partial durations can be identified with a $T \times 1$ vector $D_T = (D_1, D_2, \dots, D_T)'$ and the partial convexities with a $T \times T$ matrix C_T with elements $C_{i,j}$, the model proceeds by applying results from the theory of normed linear vector spaces to identify theoretical bounds on D_T and C_T . In the case of D_T , the Cauchy-Schwarz inequality is used.²¹ For C_T the bounds are based on restrictions on the eigenvalues of C_T derived from the theory of quadratic forms. To access these results, it is necessary to specify a norm and the norm selected, often referred to as a *direction vector*, is intuitively appealing. More precisely, taken as a group, the $(z_t - z_{0,t})$ changes in the individual spot interest rates represent shifts in yield curve shape. These individual changes can be reexpressed as the product of a direction shift N and a magnitude Δi :

$$(z_t - z_{0,t}) = n_t \Delta i \quad \text{where: } N = (n_1, n_2, \dots, n_T)'$$

From this, the spot rate curve is shocked and the immunization bounds can be derived. This approach differs slightly from shocking the yield curve and deriving the associated change in the spot rate curve. Using the spot rate approach, it follows that $N_0 = (1, 1, \dots, 1)'$ represents a parallel shift in the spot rate curve, with the size of the shift determined by Δi . Using this substitution (1) becomes:

$$\frac{P(z) - P(z_0)}{P(z_0)} \cong -\Delta i [N' D_T] + \Delta i^2 [N' C_T N]$$

Because the unit shift has been applied to spot rate rates, it is only for a parallel shift of a flat term structure that the partial duration and convexity measures, $N_0' D_T$ and $N_0' C_T N_0$, reduce to the classical definitions of duration and convexity, e.g., Reitano (1992, p. 38).

Duration and convexity bounds

Starting with the bond price function is instructive because it establishes a direct connection to the basic building blocks of the portfolio, the cash flows associated with individual assets and liabilities. Using the objective of fund surplus immunization, e.g., Messmore (1990), Reitano (1992) produces a slightly more complicated problem than calculating the partial durations and convexities for individual assets and liabilities in isolation. Surplus immunization involves recognizing the balance sheet relationship: $A = L + S$, where A is the assets held by the fund, L is the fund liabilities and S is the accumulated surplus. The objective of immunizing the portfolio surplus, $S = A - L$, against future changes in interest rates requires:

$$S(z) = \sum_{t=1}^T \frac{CF_t}{(1+z_t)^t} \rightarrow \frac{S(z) - S(z_0)}{S(z_0)} = 0 \cong -\Delta i [N' D_T] + \Delta i^2 [N' C_T N] \quad (2)$$

where D_T and C_T are now defined for the fund surplus, $S(z)$, and the cash flows at time t , CF_t , are the sum of the asset (+) and liability (-) cash flows that occur at time t . Because CF_t will be negative when the liability cash outflow exceeds the asset cash inflow at time t , it is possible for either the duration of surplus or convexity of surplus to take negative values, depending on the selected portfolio composition and yield curve shape.

It is not difficult to identify the analytical importance of the directional derivative approach. For analytical reasons, it will be easier to work with unit shifts in the spot rate curves. This requires a connection between the key building blocks of the directional derivative approach, the partial durations and convexities, and the associated classical concepts, for both the unit yield curve and spot rate curve shifts. As noted previously, the concepts are only equivalent for parallel shifts of flat yield curves. If the yield curve is not flat to start, then a unit vector change in yields, $N_0 = (1, 1, \dots, 1)'$, is then spread differentially across the spot interest rates.²² Heuristically, this can be seen by evaluating the total derivative of $S(z)$:²³

$$dS(z) = \frac{\partial S}{\partial z_1} dz_1 + \frac{\partial S}{\partial z_2} dz_2 + \dots + \frac{\partial S}{\partial z_T} dz_T \rightarrow \frac{dS(z)}{S} = \sum_{t=1}^T \frac{1}{S} \frac{\partial S}{\partial z_t} dz_t$$

If it is possible to set $dz_1 = dz_2 = \dots = dz_T = dy$, then the unit yield curve shift result applies:

$$D^S = \frac{1}{S} \frac{dS}{dy} = \sum_{t=1}^T \frac{1}{S} \frac{\partial S}{\partial z_t} = \sum_{t=1}^T D_t = D_N^S$$

where D^S is the classical modified duration of surplus. Yet, the assumption of equal changes in spot rates and yields will only apply when starting from a flat yield curve. There is a dependence among the spot interest rates that cannot be arbitrarily ignored. Even when $N_0 = (1, 1, \dots, 1)'$ it is not generally true that the sum of the partial durations calculated with a norm using a parallel change in spot rates will equal the classical duration. For example, unless the yield curve is flat, the classical duration for a coupon bond will not equal the duration calculated using the sum of the partial durations.²⁴ A similar comment applies to the convexity measures.

Using a unit change in spot rates to specify the norm can also lead to other theoretical confusions that can only be resolved by making reference to market reality. For example, Reitano selects an example with three relevant maturities, one maturity applicable to a 5 year zero coupon liability and two maturities applicable to a 6 month zero coupon asset and a 10 year coupon bond. Even though there will be partial durations and convexities associated with the regular coupon payments, only three maturity dates are incorporated into the analysis. Assuming *semi-annual* coupon payments, incorporating all the dates with cash flows would increase the number of elements in N_0 to $2T = 20$. Doing this would increase the parallel shift norming vector from $N_0'N_0 = 3$ to $N_0'N_0 = 20$. This would immediately increase the Cauchy-Schwarz bounds on $N'D_T$, even where the bounds are calculated directly from the extreme shift vector.²⁵ Following

Poitras (2003), consider the form of the extreme value solutions for the elements of N , e.g., for n_1^* , where the norming vector only identifies three yield curve points, i.e., $N_0 = (1, 1, 1)'$:

$$n_1^* = \sqrt{\frac{3}{1 + \left(\frac{D_2}{D_1}\right)^2 + \left(\frac{D_3}{D_1}\right)^2}}$$

Now consider the solution where the norming vector admits all the 20 cash flow points:

$$n_1^* = \sqrt{\frac{20}{1 + \left(\frac{D_2}{D_1}\right)^2 + \left(\frac{D_3}{D_1}\right)^2 + \dots + \left(\frac{D_{20}}{D_1}\right)^2}}$$

Such N values would almost surely result in a substantial increase in the bounds with no substantive change in the underlying immunization situation.²⁶

A mathematical resolution to this norming problem is achieved by making reference to market reality. Similar to the selection of an investment horizon, the initial specification of the optimization problem fixes the dimension of the term structure and then conducts calculations relative to this initial term structure. For example, this is the motivation for the "key rate durations" in Phoa and Shearer (1997). Once the dimension of T is set, it follows that for the semiannual payment frequency:

$$N_1 = \left(\frac{1}{\sqrt{2T}}, \frac{1}{\sqrt{2T}}, \dots, \frac{1}{\sqrt{2T}} \right) \rightarrow N_1' N_1 = 1$$

where T is the longest maturity (in years) of the assets and liabilities contained in the portfolio. In this case, $N_1' N_1 = 1$ is independent of the number of cash flow payment dates.²⁷ This solution to the inherent problems with the use of $N_0 = (1, 1, \dots, 1)'$ to specify the norm now requires interpretation of $N_1' D_T$ to involve multiplication by $2T^{1/2}$ to be equal to $N_0' D_T$. Even though this adjustment is inconsequential when the initial objective is to set the duration of surplus equal to zero, where comparisons with classical duration results are involved the adjustment by $2T^{1/2}$ is required. This difficulty can be readily resolved by factoring out $2T^{1/2}$ and redefining the magnitude of the shift as $(\Delta i / (2T)^{1/2})$. Scaling the magnitude of the shift in this fashion, as in Bowden (1997) and Balbas and Ibanez (1998), facilitates comparisons to be made across portfolios with different cash flow patterns, where norming with $N_1' N_1 = 1$ is an essential step.

Another related issue concerns the use of a key rate element vector when there is actually a larger number of cash flow payment dates. The advantage of this seemingly inconsistent reduction in the number of maturity points in the yield curve vector is that the calculated extreme bound vectors have the appearance of actual yield curve shifts. Where the yield curve or spot rate vectors are specified with all the actual number of

cash flow payments, the optical appearance of the extreme shifts often seems unrealistic. More precisely, similar to the difficulties that led to the introduction of key rate durations, e.g., Ho (1992), Phoa and Shearer (1997), the yield curve shape associated with the extreme upper and lower bounds will have an unusual sawtooth pattern. The n_t will reflect the size of the cash flow in period t , small cash flow periods will have n_t close to zero and large cash flow periods will have relatively large n_t . Yet, as it turns out, exploring this apparent problem leads to insight into other quandaries in immunization analysis, such as the so-called "duration puzzle" (Bierwag et al. 1993). The essential step is to recognize that the duration matching (zero duration of surplus) portfolio is not unique. It is possible to pick the composition of the portfolio to reduce the sensitivity of the portfolio value against certain types of yield curve shifts. In terms of the directional derivative approach, it is possible to construct the portfolio to have the upper and lower extreme bounds that are "tighter" to the zero duration of surplus condition, e.g., Reitano (1996). However, achieving tightness does not necessarily undercut the possibility of extreme yield curves that possess the sawtooth pattern.

One final point of interest about the directional derivative concerns the associated convexity calculation. Consider the direct calculation of the partial convexity of surplus, $CON_{i,j}^S$, where $i \neq j$:

$$S(z) = \sum_{t=1}^T \frac{CF_t}{(1+z_t)^t} \rightarrow \frac{\partial S(z)}{\partial z_i} = \frac{iCF_i}{(1+z_i)^{i+1}} \rightarrow \frac{\partial^2 S(z)}{\partial z_i \partial z_j} = 0 \quad \text{for } i \neq j$$

where CF_t is the cash flow at time t for $t = (1, 2, \dots, T)$. From these, it follows that the quadratic form $N'C_T^S N$ reduces to:

$$N'C_T^S N = \sum_{t=1}^T n_t^2 CON_{t,t}^S = \sum_{t=1}^T n_t^2 \frac{1}{S(z_0)} \frac{\partial^2 S(z_0)}{\partial z_t^2}$$

In terms of the extreme bounds on convexity, this is a significant simplification. Because the TxT convexity matrix is diagonal, the extreme bounds are now given by the maximum and minimum diagonal ($CON_{i,i}$) elements. If the i th element is a maximal element, the associated N^{**} vector is a Tx1 with a one in the i th position and zeros elsewhere. Similar to the duration adjustment, there will be complications in directly comparing $N_i'C_T N_i$ with either the classical convexity or $N_0'C_T N_0$.

Some basic calculations and examples

To facilitate heuristic comparison with Reitano (1991a, 1991b, 1992, 1996), the fixed liability in the portfolio initially being examined is identical to that for the example of Reitano (1992): a \$100 million par value, zero coupon GIC maturing in year 5.²⁸ This liability is being hedged with the same two assets, 6 month commercial paper, with market value A_1 , and a 12% coupon (semi-annual) 10 year bond, with market value A_2 . Variation in the universe of assets and liabilities will eventually be introduced, though

asset selection is restricted to combinations of short term, zero coupon money market instruments and longer term coupon bearing bonds. This is consistent with the absence in fixed income markets of competitively priced, zero coupon long term securities, i.e., long term fixed income markets have a preference for coupon bonds. The portfolio management objective is to choose the asset composition of the balance sheet such that the duration of surplus equal to zero is retained. Because variation in the market/par values of the balance sheet items will be permitted, a methodology is needed to determine the relevant values. Initially, the yield/spot interest rate curve from Fabozzi (1993) is selected. This *par bond* curve has semi-annual yields from 6 months to 10 years, i.e., $y = (.08, .083, .089, .092, .094, .097, .10, .104, .106, .108, .109, .112, .114, .116, .118, .119, .12, .122, .124, .125)'$. The 6 month and 1 year yields are for zero coupon securities, with the remaining yields applying to par coupon bonds. This yield curve produces the associated spot rate curve, $z = (.08, .083, .0893, .0925, .0946, .0979, .1013, .106, .1083, .1107, .1118, .1159, .1186, .1214, .1243, .1256, .1271, .1305, .1341, .1358)'$.²⁹

The first point to illustrate involves the result that the sum of the partial durations will not typically equal the classical duration, which is the case here because the yield curve slopes up. Observe that the partial duration and classical duration formulas can be specified, using semi-annual cash flows, as:³⁰

$$N_0'D_T = D_N^S = \frac{1}{S} \sum_{t=1}^{2T} \frac{\{t/2\} CF_t}{(1 + \{z_t/2\})^{t+1}} \quad D^S = \frac{1}{S} \sum_{t=1}^{2T} \frac{\{t/2\} CF_t}{(1 + \{y/2\})^{t+1}}$$

where $N_0'D_T$ is the duration of portfolio surplus calculated using the partial durations and D^S is the classical duration of surplus measure. To use these formulas to calculate the durations of a coupon bond, with regular semi-annual coupon payments, set $C/2 = CF_t$ for $t = 1$ to $(2T - 1)$, and the coupon plus return of principal, $C/2 + M = CF_{2T}$. Observing that surplus is defined by the balance sheet identity $L + S = A$, set $L = 0$ and admit only one balance sheet asset, the coupon bond. This transforms the formulas into the appropriate form. Term by term comparison for the upward sloping yield curve case reveals that the sum of the partial durations for the coupon bond must be less than the classical duration, requiring that $N_0'D_T < D^S$. For example, for a $T = 10$ semi-annual coupon par bond with $C = 12.5\%$, for the Fabozzi yield curve the (modified) classical duration is 5.62 and $N_0'D_T = 5.42$. This type of discrepancy can produce significant differences in the composition of immunizing portfolios, both for the initial portfolio as well as when yield curves change location and shape.³¹

Recognizing that the liability is a zero coupon security, the 5 year spot rate is used to determine the price, $P_L = \$58.3427$ million, and the partial duration, $D_L = 4.738$. The example immunization problem is to pick A_1 and A_2 such that the duration of surplus, D_N^S , equals zero. The balance sheet identity $S + L = A_1 + A_2$ provides an additional restriction. Allowing for the value of the surplus to be positive, the zero duration of surplus condition requires:

$$w_1 D_{A_1} + w_2 D_{A_2} - w_3 D_L = 0 \rightarrow D_L = x_1 D_{A_1} + x_2 D_{A_2}$$

where $w_1 = A_1/S$, $w_2 = A_2/S$, $w_3 = P_L/S$, $S = A_1 + A_2 - P_L$, $x_1 = A_1/P_L$ and $x_2 = A_2/P_L$. The relevant asset durations are $D_{A1} = 0.481$ and $D_{A2} = 5.462$. For the problem at hand, if S is undetermined, the system is over-identified with 2 equations and 3 unknowns. This provides for a theoretically infinite number of possible combinations of the two assets that would achieve the immunization condition. In practice, the surplus would also be given, inherited from the previous history of the fund's performance. With both the surplus and liability given, there are now two equations to satisfy with only two variables to choose. Unless more assets are available, which would often be the case in practice, it is not possible to arbitrarily fix one of the asset values, permitting the surplus immunization condition to determine the (positive) value of the other asset. Both asset values are determined once the market values of the surplus and liability are given. In addition, for purposes of calculating this example, because the long term asset is a discount bond it is necessary to have a methodology for determining the par value associated with the given market price.

More precisely, an additional complication arises when the coupon bond is not a par bond. This happens in the current example where the bond sells at a discount. Because the relevant yield curve is for *par bonds*, it is not appropriate to apply the 10 year yield to the price to calculate the coupon cash flows. Even though it is possible to calculate the price using the spot interest rates, this creates calculation difficulties when the value of the asset is being changed to, say, achieve an immunization objective. To get around this difficulty, it is expedient to unbundle the cash flows from the discount bond into a combination of a par bond and a zero coupon bond. This recognizes that, relative to a par bond, a discount bond has a larger portion of the cash flow occurring at maturity. For example, assume that the par value, M , of the 10 year discount bond is \$48. The annual cash flow from the coupon will be $(.12)(48) = 5.76$ which is the same as the cash flow from a par bond with $M = 46.08$.³² The difference between $48 - 46.08 = 1.92$ is the M for the relevant 10 year zero coupon bond. The price of the discount bond follows appropriately:

$$P_{A2} = \$46.08 + \frac{\$1.92}{(1 + \{.1358/2\})^{20}} = \$46.5963$$

This price can now be used to calculate the duration of the asset as the value weighted sum of the par bond duration and the zero coupon duration.

The initial balance sheet being investigated is "high surplus": $S (= 33.8387) + L (= 58.3427) = A_1 (= 45.5851) + A_2 (= 46.5963)$. It is readily verified that these values satisfy the immunization of surplus condition:

$$D_N^S = \frac{45.5851}{33.8387}(.4808) + \frac{46.5963}{33.8387}(5.4617) - \frac{58.3427}{33.8387}(4.7377) = 0$$

Now consider what happens when the $L + S$ side of the balance sheet is reconstructed to be "low surplus" by increasing the par value of the liability to \$150 million, such that $L (= 87.5141) + S (= 4.6674)$. Even though the total of the $S + L$ side of the balance sheet is unchanged, the previous asset composition no longer immunizes the surplus:

$$D_N^s = \frac{45.5851}{4.6674}(.4808) + \frac{46.5963}{4.6674}(5.4617) - \frac{87.5141}{4.6674}(4.7377) = -29.61$$

The immunizing portfolio is now $A_1 = 17.8382$ and $A_2 = 74.3432$. The calculated par value for this bond price is 76.4928. This par value is needed for direct calculation of the coupon and principal cash flows in the partial durations.

Both the high and low surplus cases can be examined using the Fabozzi spots as the baseline spot rate curve. Table 5.8 reports the partial durations, the n_i and extreme duration bounds calculated from the Cauchy-Schwarz inequality. Comparison of the bounds between the low and high surplus cases depends crucially on the observation that the bounds relate to the percentage change in the surplus. Due to the smaller position in the 6 month asset, the larger bounds for the low surplus case still translate to a larger relative change when compared to the high surplus case. This result is calculated by multiplying the reported bound by the size of the surplus. As expected, the extreme shift vector, N^* , exhibits a sawtooth change, with about 80% of the worst shift concentrated on a fall in the 5 year yield and 17–20% on an increase in the 10 year rate.³³ This is an immediate implication of the limited exposure to cash flows in other time periods. However, even in this relatively simple portfolio management problem, the n_i provide useful information about the worst case shift. There is not much loss of content in heuristically “filling in” the sawtooth pattern, adjusting upward the small partial durations in the intervening periods. Consistent with basic intuition, the implication is that the worst type of shift has a sizeable fall in mid-term rates combined with smaller, but still significant rise in long term rates. As it turns out, filling in the sawtooth patterns in rates cannot be done arbitrarily but must satisfy basic absence of arbitrage conditions. When this is done, the extreme duration bounds are impacted.

Table 5.8 is constructed to be roughly comparable to the example in Reitano (1991, 1992, 1996). Of the range of possible extensions, increasing the number of assets to include the addition of a maturity matching bond is of topical interest. The impact of including such bonds is the source of the “duration puzzle” in Bierwag et al. (1993). Table 5.9 provides results for two cases with similar surplus levels but with somewhat different asset compositions. One case involves, a par bond with a maturity that matches that of the zero coupon liability ($T = 5$). This is referred to as the *maturity bond* portfolio. The other case does not include the maturity matching bond but, instead, uses 3 and 7 year par bonds. This is referred to as the *split maturity* portfolio. For both asset portfolios the 1/2 year and 10 year bonds of table 5.1 are included, with the position in the 10 year bond being the same in both asset portfolios. The 1/2 year bond position is permitted to vary, with the maturity matching portfolio holding a slightly higher market value of the 1/2 year asset. *A priori*, the split maturity portfolio would seem to have an advantage as four assets are being used to immunize instead of the three bonds in the maturity matching portfolio.

Given this, the results in table 5.9 are revealing. In particular, the asset portfolio with the maturity matching bond has much smaller extreme bounds even though more bonds are being selected in the split maturity portfolio. The partial durations reveal that, as expected, the presence of a maturity matching bond reduces the partial duration at $T = 5$ compared to the split maturity case. The partial durations at $T = 3$ and $T = 7$ are

Table 5.8 Partial durations, $\{n_t\}$ and extreme bounds for the high and low surplus examples^a

Date	High surplus		Low surplus	
	D	n^a	D	n^a
0.5	0.687	0.0775	2.292	0.0236
1.0	0.075	0.0085	0.870	0.0089
1.5	0.107	0.0121	1.238	0.0128
2.0	0.136	0.0153	1.569	0.0162
2.5	0.161	0.0182	1.863	0.0192
3.0	0.183	0.0206	2.111	0.0217
3.5	0.201	0.0226	2.318	0.0239
4.0	0.214	0.0241	2.472	0.0254
4.5	0.226	0.0255	2.611	0.0269
5.0	-7.933	-0.89545	-86.115	-0.88658
5.5	0.244	0.0275	2.815	0.0290
6.0	0.246	0.0277	2.837	0.0292
6.5	0.247	0.0279	2.854	0.0294
7.0	0.246	0.0278	2.845	0.0293
7.5	0.243	0.0275	2.810	0.0289
8.0	0.242	0.0273	2.792	0.0287
8.5	0.239	0.0269	2.758	0.0284
9.0	0.231	0.0260	2.664	0.0274
9.5	0.221	0.0249	2.551	0.0263
10.0	3.786	0.42733	43.742	0.45033
Extreme duration bounds:	Cauchy = $ D $ 8.860		Cauchy = $ D $ 97.131	
Surplus:	33.8387		4.66735	

^a The high surplus portfolio is composed of (\$45.581) 1/2 year and (\$46.5963) 10 year bonds. The low surplus portfolio is composed of (\$17.8382) 1/2 year and (\$74.343) 10 year bonds. The liability for the high surplus portfolio is a 5 year zero coupon bond with \$100 par value and market value of \$58.3427. The liability for the low surplus portfolio is a 5 year zero coupon bond with \$150 par value and market value of \$87.514. The extreme Cauchy bounds are derived using $|N| = 1$.

proportionately higher in the split maturity case to account for the difference at $T = 5$. The small difference in the partial duration at $T = 10$ is due solely to the small difference in the size of the surplus. Examining the n_t^* reveals that there is not a substantial difference in the sensitivity to changes in five year rates, as might be expected. Rather, the split maturity portfolio redistributes the interest rate sensitivity along the yield curve. In contrast, the maturity bond portfolio is more heavily exposed

Table 5.9 Partial durations, $\{n_i\}$ and extreme bounds for the maturity bond and split maturity examples^a

Date	Maturity bond		Split maturity	
	D	n^a	D	n^a
0.5	0.476	0.0183	0.251	0.0063
1.0	0.454	0.0174	0.446	0.0112
1.5	0.646	0.0248	0.635	0.0159
2.0	0.818	0.0314	0.804	0.0201
2.5	0.971	0.0373	0.955	0.0239
3.0	1.101	0.0422	7.737	0.1939
3.5	1.208	0.0464	0.834	0.0209
4.0	1.288	0.0494	0.889	0.0223
4.5	1.361	0.0522	0.939	0.0235
5.0	-23.866	-0.91558	-36.998	-0.92716
5.5	0.636	0.0244	1.012	0.0254
6.0	0.641	0.0246	1.020	0.0256
6.5	0.645	0.0247	1.026	0.0257
7.0	0.643	0.0247	8.177	0.2049
7.5	0.635	0.0244	0.601	0.0151
8.0	0.631	0.0242	0.597	0.0150
8.5	0.623	0.0239	0.590	0.0148
9.0	0.602	0.0231	0.569	0.0143
9.5	0.577	0.0221	0.545	0.0137
10.0	9.884	0.37921	9.350	0.23430
Extreme duration bounds:	Cauchy = $ D $ 26.07		Cauchy = $ D $ 39.905	
Surplus:	10.32685		10.91804	

^a The maturity bond portfolio is composed of (\$5.13) 1/2 year, (\$55.54) 5 year and (\$37.172) 10 year bonds. The split maturity portfolio is composed of (\$0.4105) 1/2 year, (\$33.8435) 3 year, (27.0) 7 year and (\$37.172) 10 year bonds. The liability is a 5 year zero coupon bond with \$150 par value and market value of \$87.5121. The extreme Cauchy bounds are derived using $|N| = 1$.

to changes in 10 year rates. This greater exposure along the yield curve by the split maturity portfolio results in wider extreme duration bounds because the norming restriction dampens the allowable movement in any individual interest rate. In other words, spreading interest rate exposure along the yield curve by picking assets across a greater number of maturities acts to increase the exposure to yield curve shifts of unit length.



Questions

1*(a) The Vasicek (1977) mean-reverting SDE model of the short term rate can be used to solve for the price of a zero coupon bond that pays \$1 at time T as:

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}$$

where:

$$B(t, T) = \frac{1 - e^{\kappa(T-t)}}{\kappa}$$

$$A(t, T) = \exp \left[\frac{\{B(t, T) - (T - t)\} \{\kappa^2 \mu - (\sigma^2/2)\}}{\kappa^2} - \frac{\sigma^2 B(t, T)^2}{4\kappa} \right]$$

Using these results, solve for an equation specifying the term structure of interest rates. (Hint: see Hull 2003, pp. 538–40.)

(b) The CIR (1985) mean-reverting square root model of the short term rate can be used to solve for the price of a zero coupon bond that pays \$1 at time T as:

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}$$

where:

$$B(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{2\gamma + (\kappa + \gamma + \lambda)(e^{\gamma(T-t)} - 1)}$$

$$A(t, T) = \left[\frac{2\gamma(e^{\gamma(T-t)} - 1)}{2\gamma + (\kappa + \gamma + \lambda)(e^{\gamma(T-t)} - 1)} \right]^{\frac{2\kappa\mu}{\sigma^2}}$$

$$\gamma = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}$$

and λ is a parameter that is related to the risk aversion properties of the representative investor. Following the approach used in (a), use this result to solve for the term structure.

2 What is the relationship between Macaulay duration and the term to maturity for: (a) a zero coupon bond; (b) a perpetuity; (c) discount bonds;

(d) par and premium bonds. Plot the relationship between these four cases in a single diagram. (Hint: see Bierwag et al. 1983, p. 17.)

3 In the example of section 5.3 it was assumed that the asset portfolio contained a combination of a twenty year zero and cash. Rework this example using a one year zero coupon bond and a twenty year zero. If a flat yield curve is assumed, what is your estimate of the change in the value of the liability and the asset if yields are unchanged and one year is allowed to pass. Rework your answer assuming the yield curve slopes up with the one year yield = 4%, the five year = 5% and the 20 year equal to 6%.

4* "Whether the bond market moves up or down, high-convexity portfolios will always outperform low-convexity portfolios of equal duration and yield." Explain the argument supporting this statement. What factors would tend to undermine this position?

5* Explain this statement: "the larger the convexity on a portfolio, the less the value of the portfolio rises over time if the interest rate remains unchanged." What are the implications of this result for the asset/liability managers seeking to control interest rate risk? Is it true that "the cost of a higher convexity is a lower yield"?

6* Tables 5.6 and 5.7 were calculated using an increase in yield and a reduction in time. Because value increases with time changes, due to the payment of interest income, and falls in value when yields increase, the contributions from yield and time changes act in opposite directions. Recalculate the relevant values in these tables for a decrease in yield and a reduction in time.

7* "Classical immunization strategies, which explicitly assume parallel yield curve shifts, cannot in theory be expected to provide immunization when the yield curve shifts (are nonparallel) . . . However, these conditions readily generalize to conditions that insure immunization against any given yield curve shift assumption. Unfortunately, these conditions are not compatible in general. That is, immunization against a given type of shift will often create exposure to other types of shifts, causing immunization to fail as other shifts are realized."

Comment on the implications of this statement for asset and liability management. In your answer, be sure to identify what are "classical immunization strategies" and to explain how the generalized immunization conditions can be used to provide estimates of the degree of immunization risk.

NOTES

- 1 In Canada, a GIC (guaranteed investment certificate) can be issued by a financial institution such as a chartered bank. In the US such liabilities are referred to as certificates of deposit.
- 2 "Dispersion" is a description of the time profile of the cash flows. For example, a zero coupon bond has a cash flow without dispersion. The single cash flow is concentrated at a single maturity date. A coupon bond with the same maturity date has considerably more dispersion, as the coupon cash flows occur at dates prior to the maturity date. A fixed coupon annuity with the same maturity date has even more dispersion as the cash flows are concentrated evenly across all the payment dates. In theoretical examples, it is conventional to achieve greater dispersion by using a barbell portfolio of assets, where assets with only short and long maturities are used to immunize a liability with an intermediate term maturity.
- 3 Due to the absence of a 10 year instrument in Italy, Singapore and New Zealand, Dym (1992) uses the longest maturity active bond, without actually saying what was involved. As these three countries are involved in some of the atypical results, some caution must be taken because of possible yield curve distortions.
- 4 This expansion is sometimes presented by setting $a = x_0$ and considering a small region around x_0 by setting $x = x_0 + \varepsilon$. Using these substitutions, $x - a = \varepsilon$. Redington (1952) uses this form of the expansion.
- 5 Malliaris and Brock (1982) survey the range of applications in economics and finance.
- 6 To see this is not difficult. By construction, if y is lognormally distributed then for $\ln[y] = x$, x is normally distributed. A normally distributed variable is a real variable which can take values ranging over the real line from positive to negative infinity. Observing $\exp\{\ln[y]\} = y = \exp\{x\}$, if x takes the value of minus infinity, the lowest possible value for x on the real line, then y will take a value of zero. Hence, a lognormal variable is defined on the positive half-line, ranging from 0 to positive infinity.
- 7 The empirical literature on the distribution of security and derivative prices is voluminous. A useful introduction is included in Duffie (1989).
- 8 In the form given, the OU process is a form of arithmetic Gaussian process. Unlike the geometric case, this type of process admits the possibility of negative values for X . In certain cases, this difficulty can be rationalized away by arguing that only short term options are of interest. In other words, the probability of observing negative values increases with the length of the permissible time paths. If these paths are constrained to be short, then there will only be a negligible probability of observing negative values.
- 9 Analytically, the admissibility of a given diffusion for arbitrage free financial pricing depends on satisfaction of conditions required for Girsanov's theorem to hold. A discussion and application of this point to using Brownian bridge process to model bond prices is provided in Cheng (1991).
- 10 Extending analysis to this type of process can be motivated by observing that a diffusion process can be created from the countable combination of Poisson processes.
- 11 As discussed previously and elaborated in section 5.2, the interest rate is the most appropriate variable to model directly as a stochastic process. The bond price process can then be derived by applying Ito's lemma.
- 12 Various technical conditions associated with the lemma are suppressed, e.g., restrictions on $f[\cdot]$ and $\sigma[\cdot]$. These details can be found in various sources, e.g., Arnold (1974). It should also be recognized that other approaches to differentiation and integration of stochastic functions is possible, e.g., Meyer (1976).
- 13 Ingersoll (1987, pp. 348-9) motivates Ito's lemma using a Taylor series expansion and ignoring terms which are of order Δt .
- 14 Because the nominal bond price does not have a Brownian component, no cross product terms of the form $\{dB dP\}$ appear. This simplification is what makes Fischer's derivation uncomplicated.

- 15 A similar convention is used to define the theta for an option, e.g., Poitras (2002, ch. 9).
- 16 Classical immunization theory originates with Reddington (1952) where it was proposed that duration matching of assets and liabilities combined with a higher convexity of assets will minimize a life insurance company's possibility of loss from unexpected changes in market rates of interest.
- 17 Wu (2000) and Barber (1999) provide recent treatments of these alternative approaches to duration measures and the implications for immunization. Wu (2000) extends the use of CIR and Vasicek term structure models to derive duration measures and, as such, is a useful source for the relevant formulas.
- 18 Crack and Nawalkha (2000) and others continue to improve on the M^2 model by developing the properties of series solutions to capture changes in the level, slope and curvature of the yield curve.
- 19 The partial duration concept is closely related to the "key rate duration" of Ho (1992) and Phoa and Shearer (1997).
- 20 HOT refers to higher order terms which will be ignored. This assumption, which could be problematic for bonds with special features, such as call provisions, implies that the discussion centres on default free, straight bonds.
- 21 Reitano does not actually use the Cauchy-Schwarz inequality. Rather, the extreme yield curve shift is calculated and these values are used to calculate the bounds using N^*D_T . The result is tighter bounds than those provided by direct application of the Cauchy-Schwarz values. In Reitano's case, the upper and lower Cauchy-Schwarz bounds, calculated from the product of the inner products of N and D , are 303.6 and -303.6, respectively.
- 22 Conversely, unless the yield curve is flat, $N = (1, 1, \dots, 1)$ for the spot rates will not translate into a parallel shift in the yield curve. In addition, even though the unit shift vector does retain the classical property that the classical duration of surplus, D^S , continues to equal zero after the shift, a property also exhibited by $D_N^S = 0$, the composition of the initial immunizing portfolios will differ when the initial yield curve is not flat.
- 23 The result is heuristic because the changes in z_t are infinitesimal and not discrete. Extending the result to discrete changes is straight forward.
- 24 For surplus immunization this implies that the composition of the classical and Reitano immunizing portfolios will be different, at least when there are coupon bearing securities and yield curves are not flat.
- 25 The case of the actual Cauchy-Schwarz bounds:

$$-||N|| ||D|| \leq N^T D \leq ||N|| ||D||$$

is obvious, as the length of N will grow with T . As T goes to infinity, as in the case of a perpetuity, the bounds will also go to plus/minus infinity and be, effectively, non-binding.
- 26 Heuristically, the dates where cash flows are small would have small partial durations which would explode the denominator value, resulting in near zero n_t for those time periods. This leaves the key cash flow points to take on similar values to the $N = (1, 1, 1)^T$ case, albeit scaled up due to the larger value in the numerator.
- 27 This norm specification is used in Bowden (1997). Luenberger (1969) is an accessible treatment of optimization theory in vector spaces.
- 28 It is conventional in immunization analysis to work with simplified balance sheets even though such cases are remote from important practical applications such as life companies and pension plans. Practical situations will impose additional limitations on portfolio composition, e.g., due to the need to hold near cash reserves sufficient to meet unanticipated payout requirements.
- 29 In tables 5.8 and 5.9, indexing of these vectors is assumed to be done in half-units, i.e., $t = (.5, 1, 1.5, \dots, 9.5, 10)$. At other times, as in summation expressions, the indexing is done in integer units, i.e., $t = (1, 2, \dots, 19, 20)$.
- 30 The problem of resolving the relevant yield to use for discounting cash flows in the classical case is suppressed. In practice, the durations of assets and liabilities would be calculated using the yields associated with the maturity of the security, which would result in cash flows occurring at the same time being discounted at

different rates. The stated classical formula assumes that there is only one yield.

- 31 To see this for the initial portfolio composition, observe that the duration of surplus is calculated as: $D_S = w_1 D_{A1} + w_2 D_{A2} - w_L D_L$. The stated result follows by observing that the difference between classical durations and durations calculated as the sum of partial durations will be largest for the long maturity securities.
- 32 This par value can be determined directly by taking the ratio of the coupon percentage and the yield and multiplying by original par value, e.g., $(.12/.125)(48) = 46.08$.
- 33 To see this, recall that the length of the shift vector is one. As a consequence, the sum of the squares will equal one and the square of each n^* is the percentage contribution of that particular rate to the extreme directional shift vector.