Investment Returns and Inflation: Some Australian Evidence

by

Michael Sherris, Leanna Tedesco and Ben Zehnwirth

School of Economic and Financial Studies
Macquarie University, SYDNEY, NSW, 2109
AUSTRALIA
Tel: 61 2 9850 8572 Fax: 61 2 9850 9481
email: msherris@efs.mq.edu.au
WWW: http://www.ocs.mq.edu.au/~msherris/

Abstract

The development of stochastic investment models for actuarial and investment applications has become an important area of interest to actuaries. This paper reports the application of some techniques of modern time series and econometric analysis to Australian investment return and inflation data. It considers unit roots, cointegration and state space models. Some results from this analysis are not reflected in published stochastic investment models.

Keywords

stochastic investment models, unit roots, state space models, investment returns, inflation

INTRODUCTION

The main aim of this paper is to apply some of the techniques of modern time series and econometric analysis to analyse investment data in order to better understand the nature of long run relationships in investment series typically used in stochastic investment models. The analysis is based on Australian data. Other stochastic investment modelling studies for actuarial applications have not formally investigated such relationships yet these are fundamental to the structure of any model to be used in actuarial applications.

The paper attempts to identify and address fundamental issues that need to be considered before developing a particular stochastic investment model. It has identified many structural features of investment models that should be included in such a model. Many of these features are not found in published stochastic investment models for actuarial applications. The paper does not present a stochastic investment model. The detail required for a stochastic investment model will depend on the application. Parameter estimates and initial values for a stochastic investment model will generally need to be based on the most recent data.

Transfer functions were used to fit models to Australian investment data. These were found not to be appropriate for investment modelling analysis for the Australian asset returns series data since there was evidence of feedback relationships between many of the series. State space models were then fitted to the asset returns data and the inflation series allowing for feedback between inflation and asset class returns. The fitted models are reported in this paper.

The analysis in this paper uses quarterly data. Stochastic investment models are used in practice to establish strategic asset allocations and to examine solvency and capital adequacy. Long run asset allocation strategies are often determined using an annual model on the assumption that cash flows occur at annual intervals. In practice this will be a crude approximation to the timing of cash flows and a higher frequency model will be preferred. Similarly the capital requirements for meeting a solvency test will be much less stringent when solvency is tested at annual time intervals than at quarterly intervals. In this context the difference between annual and quarterly models has yet to be investigated but in order to do this it will be necessary to develop an appropriate quarterly model.

There are many important issues in stochastic investment modelling that require further investigation. These include modelling structural changes that have occurred in the economy using regime-switching models (Garcia and Perron, 1996 and van Norden and Vigfusson, 1996) and allowing for other time varying components of the series such as heteroscedasticity. Parameter estimation and stability of parameters require further investigation. Parameter and model uncertainty also needs to be incorporated.

Structural time series models provide a framework for analysing and developing stochastic investment models. Such models can be developed using a state space formulation. Such an approach provides a number of significant advantages over more dated time series techniques. These are:

- Models reflect the salient characteristics of the data.
- Model parameters are interpretable.
- Feedback mechanisms (bi-directional causality) are included.
- Stationarity does not have to be assumed.
- On-line model maintenance and updating on receipt of additional information.

Further research is required in this area. It is hoped that this paper will provide a foundation for that research.

DATA

The structure of a model needs to be based on economic and financial theory as far as possible. Often theory will rely on empirical data for justification. The model structure and the parameters of a stochastic model will need to reflect historical data. Parameter estimation will usually be based on historical data. A statistical analysis of historical data will also provide useful insights into the features of past experience that the model will need to capture.

The data used for the empirical analysis in this research is taken from the Reserve Bank of Australia Bulletin database. The study uses quarterly data in contrast to most other studies in this area that use annual data. The reasons for using quarterly data, rather than monthly or some other higher frequency, is that this is the highest frequency for which many of the main economic and investment series are available. It is also a frequency that is suited to most practical applications and will be more realistic than annual data as currently used.

Different series are available over different time periods. The longest time period for which data was available on a quarterly basis for all of the financial and economic series was from September 1969. Individual series were available for differing time periods. The series considered were Consumer Price Index - All groups (CPI), the All Ordinaries Share Price Index (SPI), Average Weekly Earnings - adult males (AWE), share dividend yields, 90 day bank bill yields, 2 year Treasury bond yields, 5 year Treasury bond yields, and 10 year Treasury bond yields. An index of dividends was constructed from the dividend yield and the Share Price Index series. Logarithms and differences of the logarithms are used in the analysis of the CPI, SPI, AWE and dividends. The differences in the logarithms of the level of a series is the continuously compounded equivalent growth rate of the series.

Appendix A sets out summary statistics for the series used. It is important to note that none of the series is well represented by an assumption of independent and identically distributed normal variables. Mean-variance optimisation models that are often used in determining optimal asset allocation strategies by fund managers and asset consultants assume that the returns are correlated but serially independent and identically distributed normal variables. It is clear from the data that such assumptions are inappropriate and that mean-variance models should be used with caution.

UNIT ROOTS AND STATIONARY SERIES

Many of the series used in stochastic investment modelling are non-stationary. For example the level of the Consumer Price Index, the level of the Share Price Index and the level of a dividend index can be seen to be non-stationary by simple inspection of a time series plot. It is less clear whether or not interest rates have stationary distributions and this cannot easily be determined by inspection of a time series plot. Rates of changes in indices or rates of return have been used in stochastic investment models and this might be justified because they can be considered as "natural" variables to use.

Because the level of the Consumer Price Index, the level of the Share Price Index and the level of a dividend index are non-stationary this has lead researchers to difference the data, or some transformation of the data such as a logarithmic transform, in order to obtain a stationary series for modelling. Wilkie (1986) and Carter (1991) used differenced data, as does Harris (1994, 1995) for equity and inflation series. In Carter (1991) the order of differencing was decided using more traditional time series techniques based on the sample autocorrelations. FitzHerbert (1992) fits a deterministic trend to various index levels instead of taking differences. Neither FitzHerbert (1992) nor Harris (1994, 1995) conduct formal tests for stationarity of the series used in their models.

If the level of a series is non-stationary but the difference of the series is stationary then the series is said to contain a "unit root", be "integrated or order 1", or be "difference stationary". It is important to understand that the existence of unit roots determines the nature of the trends in the series. If a series contains a unit root then the trend in the series is stochastic and shocks to the series will be permanent. If the series does not contain a unit root then the series is "trend stationary". The trend in the series will be deterministic and shocks to the series will be transitory. This has major implications for investment models in actuarial applications.

The other aspect of unit roots is that if they exist in a series and differences are not used in model fitting and parameter estimation then the statistical properties of the parameter estimates for the model will not be standard and the use of standard results for model identification and parameter estimation can result in an incorrect model structure and unreliable parameter estimates. Insignificant parameters are more likely to be accepted as being significant.

These are all significant reasons that make testing for unit roots in a series for use in developing an investment model critical. Formal statistical tests for unit roots have been developed over the past decade in the econometric literature. Unit root tests are described in many articles and books including Dickey and Fuller (1979) and Mills (1993). These procedures are implemented in various statistical and econometric packages such as Shazam (1993).

In order to test for unit roots in the series x_t the following regressions are fitted:

$$\Delta x_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t \qquad (1)$$

$$\Delta x_t = \alpha_0 + \alpha_2 t + \alpha_1 x_t + \varepsilon_t \qquad (2)$$

where ε_t are assumed to be independent and identically distributed and $\Delta x_t = x_t - x_{t-1}$ is the first difference in the series. These are referred to as Dickey-Fuller regressions. If the value of α_1 is equal to zero then x_t is integrated or order 1. In this case (1) defines x_t as a random walk with drift and (2) defines x_t as a random walk around a non-linear time trend. This can be seen by substituting $\alpha_1 = 0$ and rearranging to get:

$$y_t^* = y_{t-1}^* + \varepsilon_t$$

with

$$y_t^* = x_t - \alpha_0 t \text{ for (1), and}$$

$$y_t^* = x_t - \left[\alpha_0 + \frac{\alpha_2}{2}\right]t - \frac{\alpha_2}{2}t^2 \text{ for (2).}$$

In the case that the ε_t are not i.i.d. then the following regressions are used:

$$\Delta x_t = \alpha_0 + \alpha_1 x_t + \sum_{j=1}^{p} \gamma_j \Delta x_{t-j} + \varepsilon_t \qquad (3)$$

$$\Delta x_t = \alpha_0 + \alpha_1 x_t + \alpha_2 t + \sum_{j=1}^{p} \gamma_j \Delta x_{t-j} + \varepsilon_t \qquad (4)$$

where p is selected to ensure the errors are uncorrelated. These are referred to as augmented Dickey-Fuller regressions.

The procedure for testing for unit roots and determining the order of integration uses the t-statistic of the coefficient α_1 of x_{t-1} in the regressions given by (3) and (4). The null hypothesis is that the series is non-stationary with $\alpha_1 = 0$ and the alternative is that $\alpha_1 < 0$. The t-statistic under the null hypothesis has a non-standard distribution and is compared with the table of critical values found in Fuller (1976, p. 373). If the null is rejected then this is evidence that x_t is a stationary series. If the null is not rejected then differences of the series are taken and the differences tested for a unit root. When the null is eventually rejected the level of differencing required to reject the null determines the order of integration of the series. Usually this requires only one order of differencing for financial and economic series.

The critical values used for testing for a unit root depend on whether or not α_2 is zero in the regression given by (4). If α_0 and α_2 are zero then the t-statistic for α_1 has the non-standard limiting distribution as in Dickey and Fuller (1979). This distribution applies if α_0 is non-zero. If α_2 is non-zero then the limiting distribution is standard normal. It is therefore necessary to determine if α_2 is non-zero in order to determine critical values for testing for unit roots. Dickey and Fuller (1981) provide critical values for a range of tests as follows:

ϕ_1 using Equation (3) with Null $\alpha_0=0$, $\alpha_1=0$
ϕ_2 using Equation (4) with Null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$
ϕ_3 using Equation (4) with Null $\alpha_0\neq0$, $\alpha_2=0$, $\alpha_1=0$

The procedure for testing for unit roots can be set out as follows (Holden and Perman, 1994):

Step 1: Estimate the regression (4).
Step 2: Use ϕ_3 to test the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ against the alternative $\alpha_0 \neq 0$, $\alpha_2 \neq 0$, $\alpha_1 \neq 0$. If the null can not be rejected then go to Step 5.
Step 3: Need to determine if $\alpha_2 \neq 0$, $\alpha_1=0$, or $\alpha_2=0$, $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$, $\alpha_1 \neq 0$. Test for $\alpha_1=0$. If this null is not rejected then conclude that $\alpha_2 \neq 0$, $\alpha_1=0$ and the series has a unit root and linear trend.
Step 4: If the null is rejected in Step 3 then there is no unit root and the series is stationary. A conventional t-test for $\alpha_2=0$ is used to test for a trend. If this null is rejected then the series is stationary with a linear trend. A conventional t-test is used to test for a constant $\alpha_0 \neq 0$.
Step 5: In this case the series has a unit root with no trend and possibly with drift. This can be confirmed using the non-standard critical values for the null $\alpha_1=0$.
Step 6: To test for non-zero drift use ϕ_2 since this tests the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ and previous tests have not rejected $\alpha_2=0$, $\alpha_1=0$. If the null is not rejected then the evidence suggests that the series is a random walk without drift. If the null is rejected then the series is a random walk with drift.
Step 7: Regression (3) is used and ϕ_1 used to test the null $\alpha_0=0$, $\alpha_1=0$. This will confirm the results of earlier steps.

Table 1 sets out the unit root test statistics using the augmented Dickey-Fuller procedure for Australian quarterly data over the period September 1969 to December 1994. Table 2 gives the parameter estimates and t statistics for regressions (3) and (4) for this data.

Applying the procedure to the unit root test statistics in Tables 1 and 2 gives the following results:

logCPI - logarithm of the Consumer Price Index
ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so the series has a unit root. ϕ_1 rejects the null $\alpha_0=0$, $\alpha_1=0$ suggesting the drift is significant. This is confirmed by the regression (3) where the estimate of α_0 is 0.0244 with a t statistic of 2.94 which is significant at the 0.4% level.

ΔLogCPI - first difference of logCPI
ϕ_3 rejects the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ for the differences. The τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ so there is evidence that the differences are stationary.

Conclusion
On the assumption of i.i.d. errors, the logarithm of the CPI is integrated of order 1 and differences in the logarithm of the CPI is a stationary series with drift.

LogSPI - logarithm of the Share Price Index
ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so the series has a unit root with no trend. ϕ_2 does not reject the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ so that this is evidence of no drift. Using the more powerful test with ϕ_1 does not reject the null $\alpha_0=0$, $\alpha_1=0$. From the regression (3) the estimate of α_0 is 0.0547 with a t statistic of 0.52 which is not significant suggesting that the drift is zero.

ΔLogSPI

ϕ_3 rejects the null hypothesis that $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ for the differences. The τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ so there is evidence that the differences are stationary. From regression (4) the estimate of α_0 is 0.0128 and of α_2 is 0.0002 and neither of these is significant.

Conclusion

The logarithm of the SPI is integrated of order 1 and differences in the logarithm of the SPI is a stationary series. The results suggest that the drift in the difference of log(SPI) is not signicantly different from zero. This could be because the test used has little power against an alternative of a positive drift close to zero.

LogAWE - logarithm of Average Weekly Earnings

ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so the series has a unit root with no trend. ϕ_2 does not reject the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ so this does not reject the zero drift hypothesis. Using the more powerful test with ϕ_1 does reject the null $\alpha_0=0$, $\alpha_1=0$ suggesting that the drift is significant. From the regression (3) the estimate of α_0 is 0.1182 with a t statistic of 3.04 which confirms this.

ΔLogAWE

ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ for the differences which suggests that the differences have a unit root. However the τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ and this is evidence that the differences are stationary. Note that if the differences in the logarithm of AWE are not stationary then this means that a random shock to the continuously compounding growth rate of AWE would be permanent. A model with this feature would not be sensible since it would allow the continuously compounding growth rate to become arbitrarily large or small.

Conclusion

The logarithm of AWE is most likely integrated of order 1 although the statistical tests suggest it could be of higher order.

LogSDiv - logarithm of the Share Index Dividend Series

ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so the series has a unit root with no trend. ϕ_2 does not reject the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ so this does not reject the zero drift hypothesis. Using the more powerful test with ϕ_1 does not reject the null $\alpha_0=0$, $\alpha_1=0$ suggesting that the drift is not significant. From the regression (3) the estimate of α_0 is 0.0940 with a t statistic of 1.188 which confirms this.

Table 1 Test Statistics for Unit Roots - ADF Regressions

Variable	n	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value		*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
LogCPI	96	-2.5394	5.0773*	-0.14401	3.3658	3.2153
ΔLogCPI	95	-1.9654	1.9387	-3.2620*	3.8593	5.7811*
LogSPI	101	-0.36670	1.0593	-2.5374	2.9977	3.4468
ΔLogSPI	93	-4.0574*	8.2403*	-4.0353*	5.5106*	8.2573*
LogAWE	94	-2.9235*	5.0615*	-1.1134	3.3407	4.2320
ΔLogAWE	94	-1.6971	1.5320	-3.1763*	3.4963*	5.1460
LogSDiv	91	-0.94375	3.2947	-2.2644	3.8328	2.7697
ΔLogSDiv	97	-3.6085*	6.5145*	-3.5827*	4.3077*	6.4580*
SDyields	94	-2.6065*	3.4282	-2.4752	2.4863	3.6984
ΔSDyields	93	-4.2322*	8.9565*	-4.3210*	6.2671*	9.3998*
BB90	95	-2.0651	2.1326	-1.8133	1.4128	2.1190
ΔBB90	93	-4.3252*	9.3662*	-4.5146*	6.8251*	10.225*
TB2	98	-2.1987	2.4757	-2.2523	1.7769	2.6071
ΔTB2	97	-3.5883*	6.4738*	-3.5303*	4.3129*	6.4337*
TB5	98	-1.9812	2.0334	-1.8892	1.3792	1.9985
ΔTB5	96	-3.5858*	6.4526*	-3.6100*	4.4730*	6.6862*
TB10	101	-1.8629	1.9083	-1.3939	1.3380	1.8353
ΔTB10	98	-4.8847*	11.930*	-4.9430*	8.2041*	12.306*

* indicates significant at 10% level

Table 2 Tests for Unit Roots - Parameters of ADF Regressions
(t statistics in brackets beneath the estimate)

	Regression 3			Regression 4			
Variable	**p**	α_0	α_1	**p**	α_0	α_1	α_2
LogCPI	5	0.02438	-0.00446	5	0.017387	-0.00175	-0.00006
		(2.94)	(-2.539)		(0.5417)	(-0.1440)	(-0.2255)
ΔLogCPI	5	0.00390	-0.20633	5	0.01406	-0.41302	-0.00011
		(1.744)	(-1.965)		(3.262)	(-3.262)	(-2.724)
LogSPI	0	0.05472	-0.00594	0	0.54693	-0.10166	0.00251
		(0.5161)	(-0.3667)		(2.536)	(-2.537)	(2.598)
ΔLogSPI	7	0.02421	-1.2959	7	0.01283	-1.3302	0.00021
		(1.803)	(-4.057)		(0.4519)	(-4.035)	(0.4551)
LogAWE	7	0.11815	-0.01577	7	0.12763	-0.01766	0.00005
		(3.041)	(-2.923)		(1.513)	(-1.113)	(0.1267)
ΔLogAWE	6	0.00520	-0.29492	6	0.03463	-0.80799	-0.00034
		(1.204)	(-1.697)		(2.952)	(-3.176)	(-2.684)
LogSDiv	10	0.09404	-0.00092	10	0.99212	-0.14792	0.00361
		(1.188)	(-0.9437)		(2.332)	(-2.264)	(2.147)
ΔLogSDiv	3	0.01354	-0.67085	3	0.01587	-0.67512	-0.00004
		(1.789)	(-3.609)		(1.043)	(-3.583)	(-0.1772)
SDyields	7	0.84903	-0.18277	7	0.91397	-0.17543	-0.00176
		(2.610)	(-2.607)		(2.719)	(-2.475)	(-0.7951)
ΔSDyields	7	0.01025	-1.3678	7	0.13634	-1.4646	-0.00224
		(0.1704)	(-4.232)		(0.9379)	(-4.321)	(-0.9529)
BB90	6	1.3958	-0.12360	6	1.4177	-0.11965	-0.00121
		(1.987)	(-2.065)		(1.963)	(-1.813)	(-0.1463)
ΔBB90	7	0.00596	-1.5594	7	0.58113	-1.7217	-0.01032
		(0.02885)	(-4.325)		(1.167)	(-4.515)	(-1.268)
TB2	3	0.74802	-0.06935	3	0.71168	-0.07914	0.00255
		(2.203)	(-2.199)		(2.061)	(-2.252)	(0.6407)
ΔTB2	3	0.03630	-0.71139	3	0.10645	-0.73037	-0.00130
		(0.3556)	(-3.588)		(0.4620)	(-3.530)	(-0.3400)
TB5	3	0.61876	-0.05561	3	0.61260	-0.06071	0.00111
		(2.011)	(-1.981)		(1.977)	(-1.889)	(0.3311)
ΔTB5	4	0.03165	-0.79979	4	0.16815	-0.86536	-0.00249
		(0.3752)	(-3.586)		(0.8454)	(-3.610)	(-0.7582)
TB10	0	0.54248	-0.04698	0	0.54105	-0.04031	-0.00134
		(1.951)	(-1.863)		(1.939)	(-1.394)	(-0.4767)
ΔTB10	2	0.02894	-0.84539	2	0.15870	-0.88785	-0.00240
		(0.3953)	(-4.885)		(0.9776)	(-4.943)	(0.8958)

ΔLogSDiv

ϕ_3 does reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ which is evidence that the differences in the series are stationary without trend. The τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ so this is further evidence that the differences are stationary.

Conclusion

The logarithm of the dividend series is integrated of order 1 and the differences in the series have no drift.

SDyields - Dividend yields on the Share Price Index

ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so the series has a unit root with no trend. ϕ_2 does not reject the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ so this does not reject the zero drift hypothesis. Using the more powerful test with ϕ_1 does not reject the null $\alpha_0=0$, $\alpha_1=0$ *suggesting that the drift is not significant.* However regression (3) shows an estimate for α_0 of 0.849 and this is significant.

ΔSDyields

ϕ_3 does reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ for the differences of the series which is evidence that the differences are stationary. The τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ so this is further evidence that the differences are stationary.

Conclusion

The dividend yield series is integrated of order 1 and the differences in the series are stationary with zero drift.

Interest rates - BB90 - 90 day bank bill yields, TB2 - 2 year Treasury bond yields, TB5 - 5 year Treasury bond yields, TB10 - 10 year Treasury bond yields

The same test statistics are significant for all of the interest rate series. ϕ_3 does not reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ so this is evidence that each of the series has a unit root with no trend. ϕ_2 does not reject the null $\alpha_0=0$, $\alpha_2=0$, $\alpha_1=0$ so this does not reject the zero drift hypothesis. Using the more powerful test with ϕ_1 does not reject the null $\alpha_0=0$, $\alpha_1=0$ suggesting that the drift is not significant.

ΔInterest rates - ΔBB90, ΔTB2, ΔTB5, ΔTB10

ϕ_3 does reject the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ for all interest rate series which is evidence that the differences of each of the series is stationary without trend. The τ_2 test statistic rejects the hypothesis that $\alpha_1=0$ so this is further evidence that the differences of the series are stationary.

Conclusion

Each of the interest rate series is integrated of order 1 so the changes in yields are stationary with no drift.

Table 3 Test Statistics for Unit Roots -Phillips-Perron Tests (n=101)

Variable	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value	*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
LogCPI	-3.8378*	119.83*	2.2660	98.752*	13.073*
ΔLogCPI	-5.2039*	13.482*	-6.3676*	13.538*	20.307*
LogSPI	-0.3098	0.9029	-2.8153	3.4695	4.3031
ΔLogSPI	-9.8017*	48.052*	-9.8160*	32.128*	48.178*
LogAWE	-4.1848*	48.665*	0.0434	33.830*	9.6581*
ΔLogAWE	-8.4424*	35.636*	-9.9542*	33.037*	49.544*
LogSDiv	-1.1221	6.5966*	-1.5004	4.8385*	1.4252
ΔLogSDiv	-10.607*	56.279*	-10.613*	37.587*	56.374*
SDyields	-3.1195*	4.9308*	-3.0255	3.3472	4.9644
ΔSDyields	-9.1122*	41.529*	-9.4328*	27.813*	41.712*
BB90	-2.7201	3.7059	-2.6124	2.4444	3.6589
ΔBB90	-10.773*	58.023*	-10.775*	38.702*	58.047*
TB2	-1.9289	1.9504	-1.7487	1.2882	1.8453
ΔTB2	-9.1718*	42.071*	-9.1772*	28.106*	42.152*
TB5	-1.8815	1.8962	-1.5613	1.2723	1.7848
ΔTB5	-9.2738*	43.005*	-9.3308*	29.027*	43.541*
TB10	-4.8920*	1.9395	-1.4685	1.3388	1.8608
ΔTB10	-9.0689*	41.130*	-9.1662*	28.013*	42.018*

* indicates significant at 10% level

Phillips and Perron (1988) have proposed non-parametric procedures for testing for unit roots with more general assumptions concerning ε_t than the i.i.d. assumptions for the Dickey-Fuller and augmented Dickey-Fuller tests. Table 3 sets out the equivalent test statistics to those in Table 1 using the Phillips-Perron test procedures. These were calculated using the procedures in Shazam (1993).

The conclusions already drawn for the SPI and for the interest rate series are supported by these test statistics. There are however some differences apparent for the other series. For the logCPI series the Phillips-Perron ϕ_3 rejects the null hypothesis $\alpha_0 \neq 0$, $\alpha_2=0$, $\alpha_1=0$ and the hypothesis that $\alpha_1=0$ is not rejected. The conclusion is that the series has a unit root with a trend. The differences in the series are stationary. Similar conclusions are reached for logAWE using the Phillips-Perron statistics. In the case of LogSDiv, the logarithm of the dividends series, the Phillips-Perron statistics suggest that this series is difference stationary with drift. Dividend yields are difference stationary without drift.

So far the data period used has been common to all the series covering the period September 1969 to December 1994. Some of the series are available for longer time periods. Tables 4 and 5 report unit root test statistics for these longer time periods for the relevant series.

Table 4 Test Statistics for Unit Roots - ADF Regressions
Various Periods

Variable	n	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value		*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
Data from March 1939 to March 1995						
LogSPI	223	-0.0369	3.2802	-2.7539	4.9204*	3.9974
ΔLogSPI	210	-4.4210*	9.7739*	-4.4454*	6.5980*	9.8960*
LogAWE	210	-0.6899	2.1214	-2.4390	3.3084	3.0349
ΔLogAWE	210	-2.3994	2.9010	-2.3545	1.9627	2.9216
Data from March 1958 to December 1994						
LogCPI	179	0.18249	3.3767	-1.6925	3.3323	1.5961
ΔLogCPI	180	-3.2500*	5.2817*	-3.2444*	3.5113	5.2666
LogSPI	185	-0.29941	2.8000	-2.4794	3.9967	3.1611
ΔLogSPI	172	-4.0135*	8.0617*	-4.0168*	5.4166*	8.1171*
LogAWE	172	-0.07160	1.4983	-2.6340	3.4014	3.5482
ΔLogAWE	173	-2.6096*	3.7562	-2.6057	2.5520	3.4787
Data from March 1958 to December 1994						
LogCPI	147	-0.51786	1.3032	-2.6342	3.1287	3.4721
ΔLogCPI	141	-2.1290	2.2675	-1.9831	1.5080	2.2608
LogSPI	146	-0.60057	2.3429	-2.1793	3.0718	2.3939
ΔLogSPI	138	-4.3209*	9.3540*	-4.3673*	6.3799*	9.5508*
LogAWE	134	-1.0694	2.1296	-1.4421	1.9822	1.4069
ΔLogAWE	134	-1.6881	1.4316	-1.6724	1.2924	1.9319
TB10	143	-1.4655	1.2807	-1.6998	1.2059	1.6020
ΔTB10	142	-4.4264*	9.8343*	-4.4309*	6.6185*	9.8903*

* indicates significant at 10% level

For the period March 1939 to March 1995 Tables 4 and 5 provide support for the hypothesis that logSPI is difference stationary with drift. This longer period of data provides a more reliable estimate of the drift so the conclusion is that the logSPI is difference stationary with positive drift There is evidence in Table 4 that logAWE and logCPI are integrated of a higher order than 1 but the results in Table 5 suggest that they are integrated of order 1.

Because the data used is quarterly it is necessary to test for seasonal integration. In quarterly data there could be a bi-annual or annual frequency seasonal unit root as well as the quarterly unit root tested for already. Hylleberg et al (1990) develop tests and test statistics for seasonal unit roots. Shazam (1993) provides procedures for implementing these tests. These procedures were applied and bi-annual and annual unit roots are convincingly rejected for all of these series.

It is worth noting that structural breaks in any series can result in a stationary series appearing to have a unit root. This will lead to differencing the data when a model using the levels of the data and explicitly capturing the structural break would be more appropriate. Differencing series results in the loss of information about the long run level of the series so that care has to be taken to ensure that the series is not stationary.

It could be argued that deregulation of financial markets during the 1980's resulted in a structural break in many of the series. For instance the method used to sell government securities changed during this period and the bond market became more active. The requirements for life insurance companies, superannuation funds and banks to hold government securities were also relaxed. During this period an imputation tax system was introduced for share investments. All of these factors could well have resulted in structural changes in rates of return and the levels of the series used in this study.

Table 5 Test Statistics for Unit Roots - Phillips-Perron Tests Various Periods

Variable	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value	*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
Data from March 1939 to March 1995 (n=224)					
LogSPI	-0.066319	3.1281	-2.8455	4.9236*	4.2426
ΔLogSPI	-14.151*	100.13*	-14.143*	66.678*	100.02*
LogAWE	0.020279	57.654*	-1.0626	38.751*	0.58240
ΔLogAWE	-11.358*	64.474*	-11.332*	42.781*	64.171*
Data from March 1948 to December 1994 (n=186)					
LogCPI	0.23384	73.724*	-0.64481	49.144*	0.28228
ΔLogCPI	-5.5568*	15.382*	-5.5431*	10.185*	15.278*
LogSPI	-0.32714	2.6717	-2.5693	4.0070*	3.3852
ΔLogSPI	-12.851*	82.575*	-12.827*	54.846*	82.268*
LogAWE	-0.96316	48.872*	-0.52415	32.424*	0.52388
ΔLogAWE	-10.584*	55.994*	-10.604*	37.463*	56.192*
Data from March 1958 to December 1994 (n=147)					
LogCPI	1.5225	75.700*	-2.5431	57.601*	5.4467*
ΔLogCPI	-5.1516*	13.206*	-5.2291*	9.1058*	13.655*
LogSPI	-0.61547	2.2874	-2.2381	3.0804	2.5254
ΔLogSPI	-11.650*	67.861*	-11.611*	44.949*	67.421*
LogAWE	-0.42646	39.283*	-0.52395	26.044*	0.19468
ΔLogAWE	-10.069*	50.672*	-10.041*	33.597*	50.395*
TB10	-1.3495	1.1305	-1.3746	0.90826	1.1456
ΔTB10	-10.920*	-59.631	-10.913*	39.706*	59.559*

* indicates significant at 10% level

COINTEGRATION

The differencing operation used to achieve stationarity, often used in developing stochastic investment models for actuarial applications, involves a loss of information about long-run movements in the series. The theory of cointegration explains how to study the inter-relationships between the long-term trends in the series. These long-term trends are differenced away in the standard Box-Jenkins approach. The inter-relationships between the long-term trends in the series can be interpreted as equilibrium relationships between the series.

Empirical studies have demonstrated that financial markets generally move quickly to an equilibrium since informed investors act quickly on new information particularly when transaction costs are low and markets are liquid. Financial markets can be out of apparent equilibrium as evidenced by "speculative bubbles" that occur when the share market booms and subsequently "crashes" even though these events are consistent with rational expectations. Economic systems are less likely to be in equilibrium since friction and price stickiness in goods and labour markets can cause the adjustment process to equilibrium to occur over an extended time frame. This suggests that if equilibrium relationships exist between financial and economic variables then these will only be detected by examining data over long time periods.

Rates of return on different investments would be expected to have long run equilibrium relationships determining their relative values. For example the spread between the return on a short term investment and the return on a longer term investment should fluctuate around some long term relationship that reflects the risk premium investors require for the longer term investment over the shorter term investment. If a long term relationship does hold then the difference between the returns should have a stationary distribution. The rates of return themselves might not be stationary but a linear combination of them will be stationary if such a long run equilibrium holds. Rates of return adjusted for expected rates of inflation, referred to as "real" rates of return as compared with nominal rates of return, might also be expected to have a stationary distribution.

Similarly the level of the share price index (SPI) and the level of an inflation index (CPI) could have an equilibrium such that they do not "wander" too far away from each other even though each is non-stationary. Thus if there is excessive inflation then it is often argued for a variety of reasons that the level of the share market should eventually increase in line with inflation and vice versa. In some cases a well developed theory might not exist to specify the nature of the equilibrium relationships between different series or there might be conflicting theories. In this case it will be the empirical relationships in the data that will support one or the other theories.

Most actuaries assume that there is a relationship between equity returns and inflation. This assumption is usually implicit in the use of "real" rates of return for projecting asset values and for valuation purposes. If a constant "real" rate of return is used then this implicitly assumes that asset returns are perfectly correlated with inflation. The Wilkie model uses inflation as the main factor driving asset returns. Investment model studies by Carter (1991) and Harris (1995) include results derived from fitting Wilkie's model to Australian data and find no statistically significant empirical relationship between equity returns and rates of inflation. This conflict between often used actuarial assumptions and empirical results clearly requires investigation since it will be fundamental to investment modelling and modelling the interaction between liabilities and assets of insurance companies and pension (superannuation) funds.

It is important to recognise that equity values and inflation can have a long-run equilibrium relationship and for there to be no statistically significant relationship between equity returns and rates of inflation. This could be the case if the series are co-integrated. Since rates of inflation and equity (capital) returns are differences in the logarithm of the level of the inflation index and differences in the logarithm of the equity index respectively, these rates of change in the levels of the indexes might appear to have no statistical relationship even though the levels of the indexes might be co-integrated with a long run equilibrium relationship. Each of the index series would be difference stationary containing a unit root consistent with the notion of market efficiency and with studies of Australian data such as Carter (1991), Harris (1994) and the results in this paper.

If variables are non-stationary but an equilibrium relationship represented by a linear combination of the variables exists such that this linear combination is stationary then the variables are said to be co-integrated. Engle and Granger (1987) suggested the concept of cointegration and developed tests for cointegration. The concept of cointegration captures the notion that two or more series "move together" in some fashion. Each series, if looked at individually, need not have a long run equilibrium but their relative values might. The series have common stochastic trends.

Testing for cointegration between any two series, where there is only one co-integrating linear combination determining the equilibrium relationship between the series, requires only the unit root tests used earlier to determine the order of stationarity of the investment data. Consider two series x_t and y_t that are integrated of order 1 so that they are difference stationary. If a long term (linear) relationship exists between these then $x_t - \beta y_t$, for some constant β, will be stationary. If x_t is regressed on y_t and there is a long run equilibrium relationship between them, then the residuals from this regression will not have a unit root. Thus for these residuals the null hypothesis of a unit root should be rejected if the series are co-integrated. Otherwise there is no evidence of cointegration.

Table 6 reports the results of unit root cointegration tests for bi-variate series from Australian data using Augmented Dickey Fuller tests and Table 7 reports the results using Phillips-Perron tests. From the tests carried out earlier in this paper all of the series used were previously found to be integrated of order 1. The results in Tables 6 and 7 were calculated using procedures in Shazam (1993). They consider each of the bi-variate series over the longest time period available and also for shorter time periods.

There is no evidence that any of the bi-variate series considered, other than the 90 day bank bill yield and the 10 year Treasury bond yield, are co-integrated. In all cases other than for these two interest rates the test statistics for both ADF and Phillips-Perron tests given in Tables 6 and 7 do not reject the null hypothesis of a unit root. Thus there is no evidence that the SPI and the CPI "move together" nor that share index dividends and the CPI "move together". It is encouraging to find that the long and short interest rate are co-integrated since this is supported by the results of Ang and Moore (1994).

Table 6 Test Statistics for Co-integration - ADF Regression Tests
Various Periods

Variable	**n**	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value		*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
Data from September 1948 to March 1995						
RSC (SPI-CPI)	186	-2.2308	2.4892	-2.2755	1.7733	2.6588
ΔRSC	186	-3.9832*	7.9760*	-3.9460*	5.2886*	7.8902*
Data from March 1958 to December 1994						
RSC (SPI-CPI)	147	-2.2356	2.5917	-2.2261	1.7159	2.4817
ΔRSC	138	-4.1852*	8.7753*	-4.1839*	5.8690*	8.7864*
R10C(TB10-CPI)	144	-1.8067	1.6823	-1.7801	1.2556	1.8333
ΔR10C	147	-4.5473*	10.377*	-4.5761*	7.0730*	10.571*
Data from September 1967 to December 1994						
RSC (SPI-CPI)	109	-1.7755	1.5808	-1.8020	1.1877	1.7770
ΔRSC	100	-3.5151*	6.1879*	-3.7956*	4.8270*	7.2305*
RSD (SPI-DIVS)	102	-1.9911	1.9934	-2.0944	2.0728	3.0978
ΔRSD	100	-3.8000*	7.2352*	-4.1384*	5.7208*	8.5658*
RDC (DIVS-CPI)	99	-2.0153	2.0389	-2.0644	1.4597	2.1813
ΔRDC	99	-4.5093*	10.167*	-4.4580*	6.7409*	10.111*
Data from September 1969 to December 1994						
RSC (SPI-CPI)	101	-2.1646	2.3788	-2.2573	2.1308	3.1598
ΔRSC	93	-3.9151*	7.6758*	-4.0427*	5.4993*	8.2371*
RSD (SPI-DIVS)	94	-2.2297	2.4956	-2.4545	2.5298	3.7847
ΔRSD	93	-4.2881*	9.1979*	-4.3814*	6.4889*	9.7294*
RDC (DIVS-CPI)	91	-1.9483	1.9080	-2.0843	1.5118	2.2576
ΔRDC	97	-3.7729*	7.1208*	-3.7797*	4.8360*	7.2507*
RB90T10 (BB90-TB10)	95	-3.0086*	4.6227*	-3.3381*	3.7971	5.5976*
ΔRB90T10	93	-4.5139*	10.189*	-4.4986*	6.7474*	10.120*
RB90C (BB90-CPI)	95	-1.9732	1.9849	-1.9528	1.5621	2.3052
ΔRB90C	93	-4.3734*	9.5769*	-4.5307*	6.8770*	10.302*
RT10C (TB10-CPI)	101	-1.4796	1.0999	-1.6161	1.3038	1.9502
ΔRT10C	98	-4.9972*	12.486*	-5.0212*	8.4723*	12.708*

* indicates significant at 10% level

RSC are the residuals from the regression logSPI = $\alpha_0 + \alpha_1$ logCPI

RSD are the residuals from the regression logSPI = $\alpha_0 + \alpha_1$ logSDiv

RDC are the residuals from the regression logSDiv = $\alpha_0 + \alpha_1$ logCPI

RB90T10 are the residuals from the regression BB90 = $\alpha_0 + \alpha_1$ TB10

RB90C are the residuals from the regression BB90 = $\alpha_0 + \alpha_1$ logCPI

RT10C are the residuals from the regression TB10 = $\alpha_0 + \alpha_1$ logCPI

Table 7 Test Statistics for Co-integration - Phillips-Perron Tests Various Periods

Variable	n	τ_1	ϕ_1	τ_2	ϕ_2	ϕ_3
10% Critical Value		*(-2.57)*	*(3.78)*	*(-3.13)*	*(4.03)*	*(5.34)*
Data from September 1948 to March 1995						
RSC	187	-2.3507	2.7666	-2.3937	1.9552	2.9318
ΔRSC	186	-12.372*	76.536*	-12.345*	51.800*	76.200*
Data from March 1958 to December 1994						
RSC	148	-2.3100	2.7567	-2.3011	1.8269	2.6546
ΔRSC	147	-11.440*	65.442*	-11.400*	43.330*	64.991*
R10C	148	-1.5317	1.2507	-1.5376	1.0544	1.5053
ΔR10C	147	-11.194*	62.657*	-11.207*	41.872*	62.807*
Data from September 1967 to December 1994						
RSC	110	-1.8159	1.6544	-1.8381	1.2285	1.8384
ΔRSC	109	-10.128*	51.297*	-10.122*	34.158*	51.277*
RSD	110	-2.2646	2.5663	-2.2649	1.7906	2.6852
ΔRSD	109	-10.423*	54.332*	-10.413*	36.145*	54.207*
RDC	110	-1.9553	1.9130	-1.9500	1.2701	1.9015
ΔRDC	109	-11.337*	64.269*	-11.283*	42.449*	63.674*
Data from September 1969 to December 1994						
RSC	102	-2.1995	2.4551	-2.2783	2.1503	3.1902
ΔRSC	101	-9.6874*	46.938*	-9.7916*	31.968*	47.938*
RSD	102	-2.7287	3.7547	-2.7584	2.8158	4.1945
ΔRSD	101	-9.8873*	48.908*	-9.9629*	33.111*	49.644*
RDC	102	-1.8420	1.6966	-1.8487	1.1802	1.7688
ΔRDC	101	-10.755*	57.858*	-10.703*	38.215*	57.318*
R9010	102	-4.2355*	8.9906*	-4.5131*	6.8119*	10.201*
ΔR9010	101	-11.167*	62.349*	-11.111*	41.160*	61.726*
R90C	102	-2.6822	3.6009	-2.7375	2.6313	3.9422
ΔR90C	101	-10.789*	58.196*	-10.782*	38.750*	58.119*
R10C	102	-1.5535	1.2136	-1.6759	1.3390	2.0035
ΔR10C	101	-9.2265*	42.570*	-9.2868*	28.755*	43.130*

* indicates significant at 10% level

RSC are the residuals from the regression $\log SPI = \alpha_0 + \alpha_1 \log CPI$

RSD are the residuals from the regression $\log SPI = \alpha_0 + \alpha_1 \log SDiv$

RDC are the residuals from the regression $\log SDiv = \alpha_0 + \alpha_1 \log CPI$

R9010 are the residuals from the regression $BB90 = \alpha_0 + \alpha_1 TB10$

R90C are the residuals from the regression $BB90 = \alpha_0 + \alpha_1 \log CPI$

R10C are the residuals from the regression $TB10 = \alpha_0 + \alpha_1 \log CPI$

The conclusions that can be drawn from this analysis of co-integrating, or "long-run" equilibrium, relationships in the Australian returns data is that, with the exception of the interest rate series, the analysis finds no strong evidence that such equilibrium relationships exist between the series analysed. This has implications for the structure of stochastic investment models since it will be important to incorporate an equilibrium structure for interest rates in the model but differences in the logarithms of the SPI, CPI and dividends can be used in the model as stationary variables without the need to incorporate any specific equilibrium between these series. This also adds to the empirical evidence that there are no strong relationships between inflation and equity returns.

STATE SPACE MODELS

Wilkie (1986, 1995) and Carter (1991) use transfer functions to develop their models. This approach allows the estimation of a cascade structure for a stochastic investment model where causality in one direction is assumed. The main driving variable in these models is the rate of inflation.

Transfer functions were examined in this research. The results are not reported in any detail here since it was found that after fitting these models there was evidence of feedback between the different variables. This means that transfer functions will not adequately capture the relationship between the different series since they impose a uni-directional causality that is not supported by the empirical data.

An alternative model is the Vector Autoregressive or VAR model. These models are used in practice for asset models and have the advantage that they allow for feedback. VAR models were fitted and it was found that too many lags were required and the models were difficult to interpret. Introducing a moving average term into these models is equivalent to an infinite number of auto-regressive terms so that a Vector Autoregressive Moving Average (VARMA) model should provide a more parsimonious model than a VAR model.

State space models provide a more succinct method of representing a stochastic investment model. They have an equivalent (VARMA) representation which has fewer lagged variables than the VAR models. Transfer function models are nested in the VARMA models. A state space model can be written as a state equation:

$$\mathbf{z}_{t+1} = \mathbf{F}\,\mathbf{z}_t + \mathbf{G}\,\mathbf{e}_{t+1}$$

and an observation equation:

$$\mathbf{y}_t = \mathbf{H}\,\mathbf{z}_t$$

where $\mathbf{y}_t$ are actual observations at time t, $\mathbf{z}_t$ is the state of the model at time t, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are matrices of parameters and $\mathbf{e}_t$ is a vector of mean zero, serially uncorrelated disturbances with covariance matrix Σ. The statistical package SAS was used to fit state space models using its state space procedure that selects the best model using the AIC model selection criteria. We assume that the state of the system is observed without error and that the series used in the model are the relevant state variables.

Returns and inflation

To examine the relationships between asset returns and inflation, state space models were fitted using each of the individual asset returns series and inflation. Models were fitted to the return on the equity index, the growth rate of dividends, the 10 year bond rate and the rate of inflation since actuaries often focus on rates of return and inflation when assessing premiums and liabilities. These models also allow a comparison with the transfer function models fitted by others.

EQUITY INDEX (SPI) AND INFLATION (CPI)

The following state space model for equity index and inflation rates was fitted as the "best" model using the quarterly data from September 1948 to March 1995:

$$\mathbf{z}_t = \begin{bmatrix} y_t \\ x_t \\ y_{t+1|t} \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0.905 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.432 & 0 \end{bmatrix}, \mathbf{e}_{t+1} = \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix},$$

$$\text{with } \Sigma = \text{var}\begin{bmatrix} n_{t+1} \\ e_{t+1} \end{bmatrix} = \begin{bmatrix} 8.39\text{x}10^{-5} & -7.18\text{x}10^{-5} \\ -7.18\text{x}10^{-5} & 8.84\text{x}10^{-3} \end{bmatrix}.$$

where the variables are the differences in the logarithms of the series, or the continuously compounding returns, adjusted for the mean of the series as follows:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} (1-B)\mathbf{Y}_t - 0.0153 \\ (1-B)\mathbf{X}_t - 0.0161 \end{bmatrix}$$

with $\mathbf{Y}_t = \log(\text{CPI})_t$, $\mathbf{X}_t = \log(\text{SPI})_t$, and $y_{t+1|t} = y_{t+1} - n_{t+1}$ is the "predicted" value for time t+1 conditional on information at time t. Note that the vectors $\mathbf{e}_{t+1}$ are assumed to be a sequence of independent normally distributed random vectors with mean **0** and covariance matrix Σ.

From the covariance matrix the standard deviation of the residuals after fitting the model are 0.0092 for the quarterly continuously compounding rate of inflation and 0.094 for the quarterly continuously compounding rate of growth of the SPI with a correlation between the residuals of -0.0834.

The parameter estimates were:

Parameter	Estimate	Std. Error	t value
F(3,3)	0.905	0.041	22.326
G(3,1)	0.432	0.066	6.563

For simulation studies such as in asset-liability modelling it is important to recognise that this model does not capture parameter or model uncertainty and has been calibrated to historical data over the time period September 1948 to March 1995. The variances in asset returns and rates of inflation are assumed to be homoscedastic in this model whereas the analysis carried out later in this paper provides evidence of heteroscedasticity.

The model can be written as an equivalent VARMA model as follows:

$$\begin{bmatrix} y_{t+1} \\ x_{t+1} \end{bmatrix} = \begin{bmatrix} 0.905 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} + \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} + \begin{bmatrix} -0.473 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e_t \\ n_t \end{bmatrix}$$

Note that using this modelling approach the best model for $\log(\text{SPI})_t$ is a random walk with drift and an ARMA model is required for $\log(\text{CPI})_t$. This model is very different to that suggested by Wilkie.

EQUITY DIVIDENDS AND INFLATION (CPI)

The following state space model for equity dividends and inflation was fitted as the "best" model using the quarterly data from September 1967 to December 1994:

$$\mathbf{z}_t = \begin{bmatrix} x_t \\ y_t \\ y_{t+1|t} \\ y_{t+2|t} \\ y_{t+3|t} \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0.030 & 0.340 & -0.236 & 0.481 & 0 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0.258 \\ -0.021 & 0.382 \\ 0 & 0.440 \end{bmatrix},$$

$$\mathbf{e}_{t+1} = \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix}, \Sigma = \text{var}\begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} 4.26\text{x}10^{-3} & 4.93\text{x}10^{-5} \\ 4.93\text{x}10^{-5} & 6.22\text{x}10^{-5} \end{bmatrix}.$$

with

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} (1-B)\mathbf{X}_t - 0.0208 \\ (1-B)\mathbf{Y}_t - 0.0178 \end{bmatrix}$$

and $\mathbf{X}_t = \log(\text{DIVS})_t$, $\mathbf{Y}_t = \log(\text{CPI})_t$.

The parameter estimates were:

Parameter	Estimate	Std. Error	t-value
F(5,1)	0.030	0.010	3.038
F(5,2)	0.340	0.121	2.800
F(5,3)	-0.236	0.082	-2.864
F(5,4)	0.481	0.150	3.208
G(3,2)	0.258	0.091	2.837
G(4,1)	-0.021	0.010	-2.057
G(4,2)	0.382	0.089	4.290
G(5,2)	0.440	0.089	4.971

An equivalent VARMA model can be readily developed from the above state space model.

10-YEAR TREASURY BOND RATES (TB10) AND CPI

Using the quarterly series from March 1958 to December 1994 the state space model was:

$$\mathbf{z}_t = \begin{bmatrix} x_t \\ y_t \\ y_{t+1|t} \\ y_{t+2|t} \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.675 & -0.308 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0.126 \\ 0.382 & 0.182 \end{bmatrix},$$

$$\mathbf{e}_{t+1} = \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix}, \Sigma = \text{var}\begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} 7.15\text{x}10^{-5} & 5.09\text{x}10^{-6} \\ 5.09\text{x}10^{-6} & 7.79\text{x}10^{-5} \end{bmatrix}.$$

with

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} (1-B)\mathbf{X}_t - 0.0217 \\ (1-B)\mathbf{Y}_t - 0.0147 \end{bmatrix}$$

and $\mathbf{X}_t = (\text{TB10})_t$, $\mathbf{Y}_t = \log(\text{CPI})_t$.

The parameter estimates were:

Parameter	Estimate	Std. Error	t-value
F(4,3)	0.675	0.042	16.146
F(4,4)	-0.308	0.057	-5.448
G(3,2)	0.126	0.056	2.245
G(4,1)	0.382	0.073	5.232
G(4,2)	0.182	0.052	3.516

The equivalent VARMA model fit is given by the following equation.

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -0.308 \end{bmatrix}\begin{bmatrix} x_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0.675 \end{bmatrix}\begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0.434 \end{bmatrix}\begin{bmatrix} e_t \\ n_t \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 0 \\ 0.382 & -0.454 \end{bmatrix}\begin{bmatrix} e_{t-1} \\ n_{t-1} \end{bmatrix}$$

These bi-variate models capture the relationship between these variables and assist in understanding the nature of these series and their interrelationships. However it can be seen that the models for the inflation series differ in each of the above models. This suggests that a model incorporating all of the series could provide more information about the best model for inflation since it will incorporate the interrelationships between the series. Such a model was fitted and the resulting model was complex and difficult to interpret so it has not been set out in this paper.

Other models

Models were also fitted to the SPI and the dividend series as well as the SPI and the 10 year bond yield to examine the relationships between these series.

EQUITY INDEX (SPI) AND EQUITY DIVIDENDS

The best state-space model using the quarterly series from September 1967 to December 1994 was found to be:

$$\mathbf{z}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{e}_{t+1} = \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix},$$

$$\Sigma = \text{var}\begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} 4.26\text{x}10^{-3} & 5.67\text{x}10^{-4} \\ 5.67\text{x}10^{-4} & 1.28\text{x}10^{-2} \end{bmatrix}.$$

where

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} (1-B)\mathbf{X}_t - 0.0208 \\ (1-B)\mathbf{Y}_t - 0.0177 \end{bmatrix}$$

and $\mathbf{X}_t = \log(\text{DIVS})_t$, $\mathbf{Y}_t = \log(\text{SPI})_t$.

The model indicates that these series are random walks with drifts and correlated errors. In this case parameter estimates will be more efficient in a model that includes both series. Even though the statistical evidence supports random walk models for both series the correlation of the errors means that information is pooled by considering the two series simultaneously. Note that the model is for an equity dividend index and not for a dividend yield. The dividend yield is given by the difference in the dividend index divided by the value of the share price index. Models that assume that the dividend yield is stationary and mean-reverting will not necessarily be consistent with this fitted model.

EQUITY INDEX (SPI) AND 10-YEAR TREASURY BOND RATES (TB10)

The best state-space model for the quarterly series from March 1958 to December 1994 was:

$$\mathbf{z}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0.982 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{e}_{t+1} = \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix},$$

$$\Sigma = \text{var}\begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} 2.71\text{x}10^{-6} & -3.67\text{x}10^{-5} \\ -3.67\text{x}10^{-5} & 1.05\text{x}10^{-2} \end{bmatrix}.$$

where

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} (1-B)\mathbf{X}_t - 0.0217 \\ (1-B)\mathbf{Y}_t - 0.0177 \end{bmatrix}$$

and $\mathbf{X}_t = (\text{TB10})_t$, $\mathbf{Y}_t = \log(\text{SPI})_t$.

The parameters were:

Parameter	Estimate	Std. Error	T-Value
F(1,1)	0.982	0.016	63.029

Note that the fitted 10 year bond yield model is close to a random walk with drift and the equity index is a random walk with drift.

Summary

These models have all been bi-variate models. Ignoring heteroscedasticity, they provide support for the random walk model for the equity index. They also provide support for modelling the difference in interest rates as a stationary series and not the level of the series. These features are not found in many of the stochastic models that have details available in the public domain. For asset liability studies it will be important to have a model to project the equity returns, dividends, inflation and interest rates as a multi-variate system. As noted earlier such a model appears to be rather complex and difficult to interpret. This is an important issue that requires further investigation. The analysis of the variances of the series indicates the need for models that incorporate heteroscedasticity. In the state space approach the parameters in F and G can be allowed to be time varying. The models can incorporate parameter uncertainty. The Kalman filter maximum likelihood approach can be used with state space models to estimate these model parameters.

It should also be noted that tests of the assumptions of the model concerning the residuals have not been performed for these models. The results are based on the assumptions of i.i.d. and normally distributed errors.

CONCLUSIONS

This paper has set out the results of research into the structural features of a stochastic investment model for actuarial applications using Australian data. This analysis is fundamental to the construction of a soundly based model. It has analysed Australian investment data using a quarterly time period. It has formally tested for stationarity of all the series and tested to see which series are co-integrated and therefore maintain a long run equilibrium relationship. It has also examined the appropriateness of transfer function models that assume one way causality between series using Australian investment data.

The results of the research suggest that the stationary variables in the Australian investment data are the rate of (continuously compounding) growth in the Share Price Index (SPI), the rate of (continuously compounding) growth in the Consumer Price Index (SPI), the rate of (continuously compounding) growth in a Dividend index representing the dividends on the SPI, and differences in the interest rate series. The statistical analysis did not provide evidence that the interest rate levels were stationary.

The cointegration tests indicated a long-run equilibrium relationship exists between the interest rates whereas there was no evidence to support such a relationship between equity values, as measured by the SPI and a Dividend index, and the level of the inflation index (CPI).

Transfer functions models were fitted to the various series and inflation but they were not found to capture the relationships between these series. Thus it was necessary to use state space models to allow for feedback between the different series. State space models for the different series and inflation were fitted to provide a comparison with transfer function models fitted by other researchers.

This research demonstrates that further investigation is required in order to understand the appropriate structure for a stochastic investment model. It does highlight some important lessons for those wishing to construct and use stochastic investment models. It indicates the type of variables that should be used in these models and the nature of the relationships that should be built into them. These matters are fundamental to the construction of stochastic investment models.

APPENDIX A

AUSTRALIAN INVESTMENT DATA - SUMMARY STATISTICS

Different series were available for different time periods. In the analysis in the paper the longest time period available for the series have been used where possible. Statistics for these series over different time periods are summarised in this Appendix.

(1) Consumer Price Index - all groups (CPI) and All Ordinaries Share Price Index (SPI)

Quarterly data was available for these series for the period September 1948 to March 1995. Table A1 sets out summary statistics for these indices, the logarithm of the index and the change in the logarithm.

Table A1
Summary statistics of CPI and SPI, logarithm of CPI and SPI and first differences of logarithms of CPI and SPI.

Statistics	N	Min	Max	Mean	St Dev	Skewness	Kurtosis
CPI	187	6.70	114.70	39.1064	34.0734	1.0008	-0.4688
log(CPI)	187	1.9021	4.7423	3.2964	0.8573	0.3615	-1.2894
Δlog(CPI)	186	-0.0087	0.0704	0.0153	0.0132	1.1935	2.0125
SPI	187	84.60	2238.70	561.5283	567.3915	1.3879	0.6055
log(SPI)	187	4.4379	7.7137	5.8806	0.9354	0.3983	-0.9342
Δlog(SPI)	186	-0.5728	0.2613	0.0161	0.0940	-1.6993	8.7947

Note the negative skewness and high kurtosis for the continuously compounding return on the SPI - given by the variable Δlog(SPI).

(2) Consumer Price Index - all groups (CPI) and Share Dividends (DIVS)

Quarterly data for the period September 1967 to December 1994 was available for the CPI and dividend yields. The Dividend yield series is the Melbourne weighted (M.W.) series from September 1967 to December 1982. This was merged with the Australian dividend yield (A.Y.) series that is available from September 1983 to March 1995 by taking 2/3M.W.+1/3A.Y. for March 1983 and 1/3M.W.+2/3A.Y. for June 1983. The share dividend series (DIVS) is derived as the product of the SPI and the dividend yield for each quarter. It represents an annualised amount of dividends paid over the prior 12 months. Table A2 sets out summary statistics for these series, the logarithms of the series and the differences in the logarithms of the series.

Table A2
Summary statistics of CPI and DIVS, logarithm of CPI and DIVS and first differences of logarithms of CPI and DIVS.

Statistics	N	Min	Max	Mean	St Dev	Skewness	Kurtosis
CPI	110	16.20	112.80	56.9036	33.2526	0.3283	-1.3353
log(CPI)	110	2.7850	4.7256	3.8406	0.6678	-0.2351	-1.3759
Δlog(CPI)	109	-0.0046	0.0566	0.0178	0.0116	0.7168	1.0666
DIVS	110	747.40	9398.25	3526.76	2603.95	0.8094	-0.6369
log(DIVS)	110	6.6166	9.1483	7.8808	0.7802	0.0526	-1.3005
Δlog(DIVS)	109	-0.1987	0.2132	0.0208	0.0653	-0.3005	1.2734

(3) Interest Rates

Quarterly interest rate data is available for 90-day Bank Bills (BB90) from September 1969 to December 1994. The summary statistics of BB90, log(BB90) and $\Delta\log(BB90)_t = \log(BB90)_t - \log(BB90)_{t-1}$ are given in Table A3. Data for 5-year Treasury Bonds (TB5) is available for the period June 1969 to December 1994. The summary statistics of TB5 and $\Delta(TB5)_t = TB5_t - TB5_{t-1}$ are also given in Table A3. Data for 10-year Treasury Bonds (TB10) is available for the period March 1958 to December 1994. The summary statistics of TB10 and $\Delta(TB10)_t = TB10_t - TB10_{t-1}$ are also given in Table A3.

Table A3
Summary statistics of Interest rates and first differences of Interest rates.

Statistics	N	Min	Max	Mean	St Dev	Skewness	Kurtosis
BB90	102	4.45	19.95	10.9093	4.1029	0.3310	-0.8313
log(BB90)	102	1.4929	2.9932	2.3148	0.3981	-0.2784	-0.8812
Δlog(BB90)	101	-0.4002	0.6213	0.0034	0.1712	0.6058	1.4652
TB5	103	0.0128	0.0394	0.0253	0.0072	-0.0909	-1.0977
Δ(TB5)	102	-0.0060	0.0050	0.0001	0.0019	-0.1966	1.2405
TB10	148	0.0106	0.0394	0.0216	0.0085	0.2654	-1.3368
Δ(TB10)	147	-0.0056	0.0048	0.00008	0.0014	-0.1179	3.7768

(4) All series

Quarterly data for all series was available from September 1969 to December 1994. Table A4 provides summary statistics for this time period. The data are CPI - Consumer Price Index, LogCPI - logarithm of (CPI), AWE - Average Weekly Earnings, LogAWE - logarithm of (AWE), SPI - Share Price Index, LogSPI - logarithm of (SPI), SDyields - Share dividend yields, SDiv - Share dividends series, BB90 - 90-day bank bills yields, TB2 - 2-year treasury bond yields, TB5 - 5-year treasury bond yields, TB10 - 10-year treasury bond yields.

Table A4
Summary statistics of all series
Quarterly Data from September 1969 to December 1994

Variable	Mean	St.Dev.	Max	Min	Median	Mode	Skewness	Kurtosis
CPI	60.074	32.462	112.80	17.000	55.300	107.60	0.2375	-1.3631
LogCPI	3.9220	0.62386	4.7256	2.8332	4.0128	4.6784	-0.3408	-1.2288
AWE	776.57	398.30	1364.3	176.90	796.33	1000.8	-0.0145	-1.3769
LogAWE	6.4806	0.64460	7.2184	5.1756	6.6800	6.9086	-0.6544	-0.8507
SPI	865.01	595.05	2238.7	194.30	603.40	2238.7	0.6797	-1.0008
LogSPI	6.5177	0.71137	7.7137	5.2694	6.4026	7.7137	0.1667	-1.4523
SD yields	4.4506	1.1496	7.7300	2.0700	4.5000	5.8500	0.2237	-0.1128
SDiv	3741.5	2584.0	9398.3	861.74	2877.4	9398.3	0.7365	-0.7603
BB90	10.909	4.1029	19.950	4.4500	10.350	15.450	0.3310	-0.8313
TB2	10.185	3.2623	16.400	4.6000	9.9400	15.150	0.0137	-1.1443
TB5	10.465	2.9845	16.400	5.2000	10.030	13.850	-0.0775	-1.0843
TB10	10.648	2.8299	16.400	5.7500	10.180	9.5000	-0.0997	-1.0091

Table A5
Jarque-Bera Asymptotic LM Normality Test
September 1969 - December 1994
Chi-squared 2DF 5% Critical Value 5.99

Variable	Chi-Square Statistic
CPI	8.74*
LogCPI	8.32*
AWE	7.96*
LogAWE	10.27*
SPI	11.97*
LogSPI	9.28*
SD yields	9.47*
SDiv	11.55*
BB90	4.87
TB2	5.60
TB5	5.15
TB10	4.57

*significant at 5% level

References

Ang, A. and D. Moore. (1994). *The Australian Yield Curve: A Cointegration Analysis*, Unpublished working paper, Macquarie University.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, 307-327.

Box, G.E.P. & Jenkins, G.M.. (1976). *Time series analysis, forecasting and control.* Holden-Day, San Francisco

Campbell, J.Y., and P. Perron. (1991), *Pitfalls and Opportunities: What Macroeconomists should know about Unit Roots.* NBER Macroeconomics Annual, 141-200.

Carter, Jon (1991). The derivation and application of an Australian stochastic Investment Model. *Transactions of The Institute of Actuaries of Australia*, 315-428.

Chow, G. (1960), Tests for Equality Between Sets of Coefficients in Two Linear Regressions, *Econometrica*, 28, 591-605.

Daykin, C.D. & G.B Hey. (1989). Modelling the Operations of a General Insurance Company by Simulation. *Journal of the Institute of Actuaries.* 116, 639-662.

Daykin, C.D. & G.B Hey. (1990). Managing uncertainty in a general insurance company. *Journal of the Institute of Actuaries.* 117, 173-259.

Dickey, D.A. and W.A. Fuller. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D.A. and W.A. Fuller. (1981). Likelihood Ratio Statistics for autoregressive Time Series with a Unit Root, *Econometrica*, 50, 1057-1072..

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987-1007.

Engle, R. F., and C. W. Granger. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, Vol 55, 251-276.

FitzHerbert, R. (1992). Stochastic Investment Models. *Transactions of The Institute of Actuaries of Australia.* 197-255.

Fuller, W.A., (1976). *Introduction to statistical time series.* Wiley, New York, U.S.A..

Garcia, R., and P. Perron (1996), An Analysis of the Real Interest Rate Under Regime Shifts, *Review of Economics and Statistics*, 111-126.

Geoghegan et al (1992). Report on the Wilkie Stochastic Investment Model. *Journal of the Institute of Actuaries*. Vol.119. Part II, 173-228.

Harris, G., (1994). On Australian Stochastic Share Return Models for Actuarial Use. *The Institute of Actuaries of Australia Quarterly Journal*, September 1994, 34-54.

Harris, G., (1995). A Comparison of Stochastic Asset Models for Long Term Studies. *The Institute of Actuaries of Australia Quarterly Journal*, September 1995, 43-75.

Harvey, A.C. (1989). *Forecasting structural time-series models, and the Kalman filter*. Cambridge University Press.

Holden, D. and R. Perman, (1994). Unit Roots and Cointegration for the Economist, Chapter 3 in B. B. Rao (ed.), *Cointegration for the Applied Economist*, Macmillan, London.

Huber, P. P., (1996), A Review of Wilkie's Stochastic Asset Model, to appear in British Actuarial Journal.

Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo. (1990). Seasonal Integration and Cointegration, *Journal of Econometrics*, 44, 215-238.

Johansen, S. and K Juselius. (1990), Maximum Likelihood estimation and Inference on Cointegration - with Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, 52, 2, 169-210.

Johansen, S. and K Juselius. (1994), Identification of the long-run and the short-run structure: An application to the ISLM model, *Journal of Econometrics*, 63, 7-36.

Mills, T.C. (1993). *The econometric modelling of financial time series*. Cambridge University Press, Great Britain.

Muscatelli, V. A. and S. Hurn. (1992). Cointegration and Dynamic Time Series Models, *Journal of Economic Surveys*, Vol. 6, No. 1, 1-43.

Perman, R. (1991), Cointegration: An Introduction to the Literature. *Journal of Economic Studies*. Vol 18, No 3, 3-30.

Phillips, P.C.B. and S. Ouliaris, (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58, 165-193.

Phillips, P.C.B. and P. Perron. (1988). Testing for Units Roots in Time Series Regressions, *Biometrika*, 75, 335-346.

Said, S. E., and D. A. Dickey. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika*, Vol. 71, 599-607.

SAS/ETS manual. (1991). *SAS/ETS Software*: Applications Guide 1. Version 6, First Edition, SAS Institute Inc.: Cary, NC, U.S.A..

SHAZAM User's Reference manual. (1993), *Shazam Econometrics Computer Program*, Version 7.0, Shazam, Vancouver, Canada.

Sherris, M. (1995), Interest Rate Risk Factors in the Australian Bond Market, *Proceedings of the 5th AFIR Colloquium.*

van Norden, S., and R. Vigfusson (1996), Regime-Switching Models: A Guide to the Bank of Canada Gauss Procedures, *Bank of Canada Working Paper 96-3.*

Wilkie, A. D. (1986). A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries.* 39, 341.

Wilkie, A. D. (1995). More on a Stochastic Asset Model for Actuarial Use. *British Actuarial Journal*, Vol 1, Part V, 777-946.