FORECASTING VALUE-WEIGHTED REAL RETURNS OF TSE PORTFOLIOS USING DIVIDEND YIELDS

by

JOSEPH WADE BLANCHARD

B.Sc., Dalhousie University, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Statistics

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of STATISTICS

The University of British Columbia
Vancouver, Canada

Date JULY 5 / 1993

DE-6 (2/88)

Abstract

We assess the ability of dividend yields denoted by DY_t, to forecast value-weighted real returns, denoted by $R_{t,T}$ of Toronto Stock Exchange (TSE) portfolios for following return horizons, T: monthly, quarterly, and one to four year. Fama and French [4] applied similar methods to the New York Stock Exchange and found the forecast power increases as the return horizon increases. We find that the Fama and French methods generalize to TSE portfolios, however, it does not apply to all portfolios. We also determine that the Fama and French approach may not lie on solid statistical ground, in that the residual variance is not time invariant.

With these drawbacks in mind we consider using the methods of Dynamic Linear Models as discussed in West and Harrison [13], which allow the model parameters to be time varying. We conclude that for the majority of the portfolios, the two methods agree, however, the regression DLM approach does slightly better in comparison with the methods of Fama and French in terms of standarized forecast errors.

Table of Contents

List of Tables

List of Figures

List of Financial Variables

CPI_t – A measure of inflation at time t based on the prices of produces the typical consumer purchases.

$I_{t,T}$ – The continuously compounded inflation rate at time t for return horizon T.

D_t – The dividends received in the time period $t-1$ to t.

P_t – The value of the portfolio at time t.

DY_t – The dividend yield at time t.

$r_{t,T}$ – The continuously compounded nominal return at time t for return horizon T.

$CD_{t,T}$ – The accumulated dividends in the time period t to $t+T$.

$R_{t,T}$ – The continuously compounded real return at time t for return horizon T

Acknowledgements

I would like to thank Dr. Jian Liu, my supervisor, for his patience and generousity in helping me complete the manuscript. I would also like to thank Dr. Harry Joe for his comments and suggestions on improving the manuscript. Finally, I would like to thank my employer, Dr. Chris Field of Dalhousie University, for his patience in my completing the thesis while working at Dalhousie.

Chapter 1

Introduction

Do stock market returns have predictable components? Several studies suggest that the answer is, in fact, affirmative. However, the predictable components typically account for approximately 3% of return variances.

Fama and Schwert [5] assess the predictability of one month U.S. treasury bill rate on the monthly return of the value–weighted portfolio of all New York Stock Exchange (NYSE) stocks for the period January 1953 to July 1971. Using regression techniques, Fama and Schwert [5] conclude that the one month treasury bill rate accounts for 3% of the stock market return variability.

Keim and Stambaugh [8] using data for the period January 1928 to November 1978 and the following variables:

(a) $(y_{UBBA} - y_{TB})$ = difference in yields on long–term under–BAA–rated (low-grade) corporate bonds and short–term (approximately one month) U.S. treasury bills;

(b) $-\log(SP_{t-1}/\overline{SP}_{t-1})$, where SP_{t-1} is the real Standard and Poor's (S & P) Composite Index and $\overline{SP}_{t-1}$ is the average of the year–end real index over the 45 years prior to the year containing month $t-1$;

(c) $-\overline{\log P_t}$ where P_t is the share price, averaged equally across the quintile of firms with the smallest market values on the NYSE.

to predict stock market returns on firms of various sizes. Specifically,

(1) Q5 – common stocks making up the fifth quintile of firms ranked by size on the NYSE, i.e. the quintile containing the largest firms trading on the NYSE;

(2) Q3 – common stocks making up the third quintile of size on the NYSE;

(3) Q1 – common stocks making up the first quintile of size on the NYSE.

Keim and Stambaugh [8] use weighted least squares, with weights corresponding to the variance of the daily returns of the S & P. In addition, they consider two sub-periods: January 1928 to December 1952 and January 1952 to November 1978.

Stock	$(y_{UBBA} - y_{TB})$ adj R^2	$-\log(SP_{t-1}/\overline{SP}_{t-1})$ adj R^2	$-\overline{\log P}_{Qt}$ adj R^2
	January 1928 – November 1978		
Q5	0.006	0.005	0.006
Q3	0.002	0.002	0.008
Q1	0.001	0.002	0.014
	January 1928 – December 1952		
Q5	-0.003	-0.001	-0.002
Q3	0.001	0.000	0.005
Q1	0.002	0.006	0.020
	January 1953 – November 1978		
Q5	0.000	-0.003	0.000
Q3	-0.003	-0.003	0.001
Q1	0.003	-0.003	-0.001

As is evident from the Table, Keim and Stambaugh [8] conclude that the predictable component of stock market returns accounts for less than 2% of the return variability.

French, Schwert and Stambaugh [6] use daily stock market returns for the period January 1928 to December 1984 S & P index, to arrive at an estimate of the monthly return volatility using autoregressive–integrated–moving average (ARIMA) models (Box and Jenkins [2]). The estimates of volatility are then used as a predictor for the monthly value–weighted returns of the NYSE. French, Schwert and Stambaugh conclude that the stock market volatility accounts for less than 2% of the return variance.

Fama and French [4] use dividend yields to predict returns on the value-weighted portfolios of NYSE stocks for return horizons of one month, one quarter and one year to four years. Fama and French [4] state the the amount of predictability in stock market returns increases as the return period increases.

The purpose of this thesis is two–fold: firstly, we apply the methods of Fama and French [4] to the Toronto Stock Exchange (TSE) value–weighted portfolios; secondly, we extend the methods of Fama and French [4] by using Dynamic Linear Regression Models described in West and Harrison [13].

Chapter 2 will describe in more detail the method of Fama and French [4] and its application to TSE portfolios. In chapter 3, we describe the basic theory of the Dynamic Linear Model (DLM) as given in West and Harrison [13]. In chapter 4, we apply the West and Harrison techniques to the TSE data. Finally in chapter 5, will give a summary of the major findings of the thesis.

1.1 Data

The data consist of monthly Index Values and dividend yields of the Toronto Stock Exchange (TSE) 300 Composite and its 14 industry sectors (see Table 1.1). Appendix 1 contains a description of the method of calculation of the Index and dividend yield for a given component of the TSE Composite. Appendix 2 contains the Relative weights that each of the 300 stocks have in determining the TSE Composite and its industry portfolios on the close of December 31, 1991.

Table 1.1: The TSE Composite Index and its fourteen industry sectors.

Series	Begin	End	Percentage in Dec 1991	Source
Composite	Jan 1956	Dec 1992		TSE Review [11]
(1) Metals and Minerals	Jan 1956	Dec 1992	7.70	TSE Review
(2) Gold and Silver	Jan 1956	Dec 1992	7.39	TSE Review
(3) Oil and Gas	Jan 1956	Dec 1992	6.71	TSE Review
(4) Paper and Forest Products	Jan 1956	Dec 1992	2.23	TSE Review
(5) Consumer Products	Jan 1956	Dec 1992	9.47	TSE Review
(6) Industrial Products	Jan 1956	Dec 1992	11.05	TSE Review
(7) Real Estate and Construction	Jan 1968	Dec 1992	0.91	TSE Review
(8) Transportation and Environmental Services	Jan 1956	Dec 1992	2.12	TSE Review
(9) Pipelines	Jan 1956	Dec 1992	2.06	TSE Review
(10) Utilities	Jan 1956	Dec 1992	13.72	TSE Review
(11) Communications and Media	Jan 1956	Dec 1992	4.74	TSE Review
(12) Merchandising	Jan 1956	Dec 1992	5.13	TSE Review
(13) Financial Services	Jan 1956	Dec 1992	20.98	TSE Review
(14) Conglomerates	Jan 1956	Dec 1992	5.76	TSE Review
Consumer Price Index	Jan 1956	Dec 1992		Bank of Canada Review

Chapter 2

Fama and French

2.1 Basic Definitions

This section contains the basic definitions of the return rate of a portfolio, the rate of inflation and dividend yields which will be used in the thesis.

Definition 2.1 *The continuously compounded inflation rate at time t, for return horizon T, is given by*

$$I_{t,T} = \log(CPI_{t+T}/CPI_t)$$

where CPI_t is the consumer price index at time t.

Definition 2.2 *The dividend yield of a stock portfolio at time t is given as follows:*

$$DY_t = D_t/P_t$$

where D_t is the dividend received in time period $t-1$ to t and P_t is the value of the portfolio at time t.

Definition 2.3 *The continuously compounded nominal return of a stock portfolio at time t, for return horizon T, is given by*

$$r_{t,T} = \log\left(\frac{P_{t+T} + CD_{t,T}}{P_t}\right)$$

where $CD_{t,T}$ are the accumulated dividends in the time period t to $t+T$.

Definition 2.4 *The continuously compounded real return of a stock portfolio at time t, for return horizon T, is given by*

$$R_{t,T} = r_{t,T} - I_{t,T}.$$

We can apply the above definitions to each of the portfolios given in Table 1.1 by identifying the index value at time t with the price of the portfolio at time t.

2.2 Motivation

This section will attempt to motivate the conjecture, proposed previously, that dividend yields predict returns. The following is taken from Fama and French [4].

Consider a discrete-time deterministic model in which D_t, the dividend per share for the time period from t to $t+1$, grows at the constant rate g, and the market interest rate that relates the stream of future dividends to the stock price P_t at time t is the constant r. In this model the price, P_t is

$$\begin{aligned} P_t &= D_t \left[\frac{1+g}{1+r} + \frac{(1+g)^2}{(1+r)^2} + \dots \right] \\ &= D_t \frac{1+g}{1+r} \frac{1}{1 - \frac{1+g}{1+r}} \\ &= D_t \frac{1+g}{r-g} \end{aligned}$$

$$\frac{D_t}{P_t} = \frac{(r-g)}{1+g}$$

The interest rate r is the discount rate for dividends and the period by period return on the stock, this can be seen as follows:

$$\frac{D_t}{P_t} = \frac{r-g}{1+g} \Rightarrow \frac{D_{t+1}}{P_{t+1}} = \frac{r-g}{1+g}$$

however,

$$D_{t+1} = (1+g)D_t \Rightarrow P_{t+1} = \frac{(1+g)^2}{r-g}D_t$$

and the return r is

$$\begin{aligned}
r &= \frac{P_{t+1} + D_{t+1} - P_t}{P_t} \\
&= \frac{[(1+g)^2/(r-g) + (1+g)]D_t - P_t}{P_t} \\
&= \frac{[(1+g)^2/(r-g) + (1+g)](r-g)/(1+g)P_t - P_t}{P_t} \\
&= \frac{(1+r)P_t - P_t}{P_t}.
\end{aligned}$$

The transition form the deterministic model to a model that (a) allows uncertain future dividends and discount rates and (b) shows the relationship between discount rates and time–varying expected returns is difficult. The deterministic model given above, at the very least, lends plausibility to the conjecture that dividend yields predict expected returns.

2.3 Regressions

Following Fama and French [4], the following linear regression model is proposed to model expected returns:

$$R_{t,T} = \alpha_T + \beta_T DY_t + \varepsilon_{t,T} \quad t = 1, 2, \ldots, N_T \tag{2.1}$$

where

- $R_{t,T}$ is the continuously compounded real return for return horizon T;
- DY_t is the dividend yield at time t;

- α_T is the intercept for return horizon T;
- β_T is the slope for return horizon T;
- $\varepsilon_{t,T} \sim N[0, \sigma_T^2]$ is the error term;
- N_T is the number of observations in for return horizon T.

This model will be applied to the value-weighted real returns for the TSE portfolios for return horizons of one month, one quarter, and one to four years. The monthly, quarterly and annual returns are non-overlapping, the other return horizons are overlapping year–end values.

2.3.1 Parameter Estimation

The usual regression estimators for equation (2.1) are given as follows:

$$\begin{aligned}
\widehat{\beta_T} &= \frac{\sum_{t=1}^{N_T}(R_{t,T} - \overline{R_T})(DY_t - \overline{DY})}{\sum_{t=1}^{N_T}(DY_t - \overline{DY})^2}, \\
\widehat{\alpha_T} &= \overline{R_T} - \widehat{\beta_T}\overline{DY}, \\
\widehat{\sigma_T^2} &= \frac{1}{N_T - 2}\sum_{t=1}^{N_T}\widehat{\varepsilon}_{t,T}^2
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}(\widehat{\alpha_T}) &= \widehat{\sigma_T^2}\left(\frac{1}{N_T} + \frac{\overline{DY}^2}{\sum_{t=1}^{N_T}(DY_t - \overline{DY})^2}\right) \\
\mathrm{Var}(\widehat{\beta_T}) &= \widehat{\sigma_T^2}\frac{1}{\sum_{t=1}^{N_T}(DY_t - \overline{DY})^2}
\end{aligned}$$

or, in matrix notation:

$$\widehat{(\alpha_T, \beta_T)} = (\mathbf{X}_T'\mathbf{X}_T)^{-1}\mathbf{X}_T'\mathbf{R_T}$$

where

$$\mathbf{X}_T = \begin{bmatrix} 1 & DY_1 \\ 1 & DY_2 \\ \vdots & \vdots \\ 1 & DY_{N_T} \end{bmatrix} \quad and \quad \mathbf{R_T} = \begin{bmatrix} R_{1,T} \\ R_{2,T} \\ \vdots \\ R_{N_T,T} \end{bmatrix}$$

and

$$\mathrm{Var}(\widehat{(\alpha_T, \beta_T)}) = \widehat{\sigma_T^2}(\mathbf{X}_T'\mathbf{X}_T)^{-1}$$

for each return horizon T.

2.3.2 Assumptions

The estimates given above are valid provided the assumptions about the error sequence $\boldsymbol{\varepsilon}_T' = (\varepsilon_{1,T}, \varepsilon_{2,T}, \ldots, \varepsilon_{N_T,T})$ are not violated *i.e.*

$$\mathrm{Var}(\boldsymbol{\varepsilon}_T) = \sigma_T^2 I_{(N_T-T)\times(N_T-T)}$$

where $I_{s\times s}$ is the $s \times s$ identity matrix. That is, the errors are un–correlated and have constant variance.

In order to check the assumption of lack of serial dependence, we define the following:

Definition 2.5 *The auto–covariance function γ_τ of a second order stationary time series, $\{Z_t\}_1^N$, is given as follows:*

$$\gamma_\tau = Cov(Z_t, Z_{t+\tau})$$

and the auto–correlation function ρ_τ is given by

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}.$$

A second order stationary time series is one in which the mean is independent of time and the auto–covariance function depends only on the time difference τ between any two time points. The reader is referred to Box and Jenkins [2] for a more detailed discussion of stationarity.

Definition 2.6 *The sample analogue of γ_τ, denoted by $\widehat{\gamma}_\tau$, based on a sample of size N, is given by*

$$\widehat{\gamma}_\tau = \frac{1}{N}\sum_{i=1}^{N-\tau}(Z_t - \overline{Z})(Z_{t+\tau} - \overline{Z})$$

and hence the sample estimate of ρ_τ, denoted by $\widehat{\rho}_\tau$, is

$$\widehat{\rho}_\tau = \frac{\widehat{\gamma}_\tau}{\widehat{\gamma}_0}$$

Now, if the underlying process has $\rho_\tau = 0$ for $\tau > 0$, then

$$\mathrm{Var}(\widehat{\rho}_\tau) = \frac{1}{N}.$$

For a justification see Box and Jenkins [2].

Thus we have a method of determining whether there is serial dependence present in the residuals. Simply plot the sample auto–correlation function and determine if an inordinate number of sample auto-correlations fall outside the limits $\pm 2/\sqrt{N}$. We define E to be the number of times the sample auto–correlation fails outside the range $\pm 2/\sqrt{N}$.

What about the homogeneity of variance assumption will be discussed in a later section. The next section gives a correction applied to the variance of the estimators when the returns are over–lapping.

2.3.3 Corrections

Hansen and Hodrick [7] suggest a correction to the variances of the estimates given Section 2.3.1 when the return horizon is larger than the sampling period. Specifically, the correction should be applied to the 2 year, 3 year and 4 year return horizons. Hansen and Hodrick suggest the following modified covariance matrix for $(\widehat{(\alpha_T, \beta_T)})$

$$
\begin{aligned}
\Theta &= \frac{1}{N_T}\mathrm{Var}(\widehat{(\alpha_T, \beta_T)}) = \frac{1}{N_T} R_0^{\mathbf{x}\,-1}\,\Omega\, R_0^{\mathbf{x}\,-1},\\
\Omega &= \sum_{\tau=-T+1}^{T-1} R_\tau^{\varepsilon} R_\tau^{\mathbf{x}},\\
R_\tau^{\varepsilon} &= E(\varepsilon_{t,T}\varepsilon_{t+\tau,T}),\\
R_\tau^{\mathbf{x}} &= E(\mathbf{x}_t'\mathbf{x}_{t+\tau})
\end{aligned}
$$

where $\mathbf{x}_t' = (1, DY_t)$ and $\varepsilon_{t,T}$ are the regression errors for return horizon T.

Now we must have estimates of $R_\tau^{\mathbf{x}}$ and R_τ^{ε} for $\tau = -T+1, \ldots, T-1$. The estimators are given by:

$$
\begin{aligned}
\widehat{R}_\tau^{\mathbf{x}} &= \frac{1}{N_T}\sum_{t=1}^{N_T-\tau} \mathbf{x}_t'\mathbf{x}_{t+\tau}\\
\widehat{R}_\tau^{\varepsilon} &= \frac{1}{N_T}\sum_{t=1}^{N_T-\tau} \widehat{\varepsilon}_{t,T}\widehat{\varepsilon}_{t+\tau,T}
\end{aligned}
$$

where, $\widehat{\varepsilon}$ are the residuals, which estimate the errors ε.

Thus, an estimator of Θ, denoted by $\widehat{\Theta}$, is

$$
\widehat{\Theta} = \frac{1}{N_T}\widehat{R}_0^{\mathbf{x}\,-1}\,\widehat{\Omega}\,\widehat{R}_0^{\mathbf{x}\,-1}
$$

where

$$
\widehat{\Omega} = \sum_{\tau=-T+1}^{T-1} \widehat{R}_\tau^{\varepsilon}\widehat{R}_\tau^{\mathbf{x}}.
$$

2.3.4 Out of Sample Forecasts

It is well known in regression analysis that R^2 tends to be overly optimistic (see Draper and Smith [3] or Weisberg [12]). To illustrate what is meant by this statement consider the following:

- split the data randomly into two parts, a construction sample and a validation sample;
- compute the parameter estimates using only the construction sample and compute an $R^2_{\text{construction}}$ for the construction sample;
- using the estimates of the parameters from the construction sample, compute predicted values for the validation sample to form the following out of sample R^2

$$R^2_{\text{validation}} = 1 - \frac{\sum_{\text{validation}}(y_{\text{validation}} - y_{\text{pred}})^2}{\sum_{\text{validation}}(y_{\text{validation}} - \overline{y}_{\text{validation}})^2}.$$

Then $R^2_{\text{construction}}$ will be, in general, larger than $R^2_{\text{validation}}$ provided the model assumptions apply.

It is the goal of this section to provide estimates of the forecast performance, *i.e.* R^2, that are not totally model based. That is, we wish to have estimates of the forecast performance of the model given in equation (2.1) when applied to future values of the real returns. Consider the following strategy:

(a) Choose a window length W which will provide the estimates of the parameters of the model (2.1).

(b) For the next time point, $W+1$ compute the forecast using the parameter estimates from (a) and the DY_{W+1} to forecast $R_{W+1,T}$. Which results in a forecast error $e_{W+1,T} = R_{W+1,T} - \widehat{R}_{W+1,T}$.

(c) Move the window forward one point in time. Now we will base our parameter estimates on the data points $2, \ldots, W+1$ and then we repeat (b), using the next time point and continue with steps (b) and (c) until the end of the series is reached.

Now we can compute the following measure of forecast performance

$$R^2_{\text{out}} = 1 - \frac{MSE_{\text{out}}}{s^2_{\text{out}}}$$

where

$$MSE = \textstyle\sum_{t=W+1}^{N_T} e^2_{t,T} \quad \text{and} \quad s^2_{out} = \sum_{t=W+1}^{N_T} (R_{t,T} - \overline{R}_{\text{out}})^2$$

with

$$\overline{R}_{\text{out}} = \frac{1}{N_T - W} \sum_{t=W+1}^{N_T} R_{t,T}.$$

2.3.5 Stationarity of Parameter Estimates

Is it reasonable to assume that the same model given in equation (2.1) should apply to the entire sampling period. That is, should the parameters of the model (α_T, β_T and σ^2_T) be time varying. This section will discuss an ad hoc approach to this problem. The next chapter on Dynamic Linear Models will deal with this problem in a more rigorous fashion.

We assume that the model (2.1) holds for some fixed time period, say L. We then use the following strategy:

(a) Using the first L data points,*i.e.* $t = 1, \ldots, L$, to provide estimates $\hat{\alpha}_T$, $\hat{\beta}_T$ and $\hat{\sigma}^2_T$ of the parameters.

(b) Move the data window forward, U time points.

(c) Use the data provided at time points $t = 1 + U, \ldots, L + U$ to get new estimates for the parameters.

(d) Repeat steps (b) and (c) until we reach the end of the time series.

One could then examine the evolution of the parameter values as we move through time.

2.4 Results

We now apply the techniques discussed in the previous sections to value–weighted real returns of TSE portfolios given in table 1.1. The real returns for return horizon T at time t are denoted by $R_{t,T}$, see definition 2.4. The availability of the returns are given in the following table:

Return Horizon	N_T	Start	End
M	443(299)	January 1956(1968)	November 1992
Q	147(99)	1st Quarter 1956(1968)	3rd Quarter 1992
1	36(24)	1956(1968)	1991
2	35(23)	1956(1968)	1990
3	34(22)	1556(1968)	1989
4	33(21)	1956(1968)	1988

The values in parentheses apply to the Real Estate and Construction portfolio, while the others apply to remaining TSE portfolios.

Figures 2.1, 2.2 and 2.3 display plots of the real returns $\{R_{t,T}\}$, dividend yields $\{DY_t\}$ and real returns $\{R_{t,T}\}$ vs dividend yields $\{DY_t\}$ for of each of the six return horizons T for the TSE Composite portfolio.

It is apparent by examining the three sets of plots that the dividends yields appear to track real returns more closely as the return horizon increases. This indicates, at least for the Composite portfolio, that the results of Fama and French [4] seem to apply.

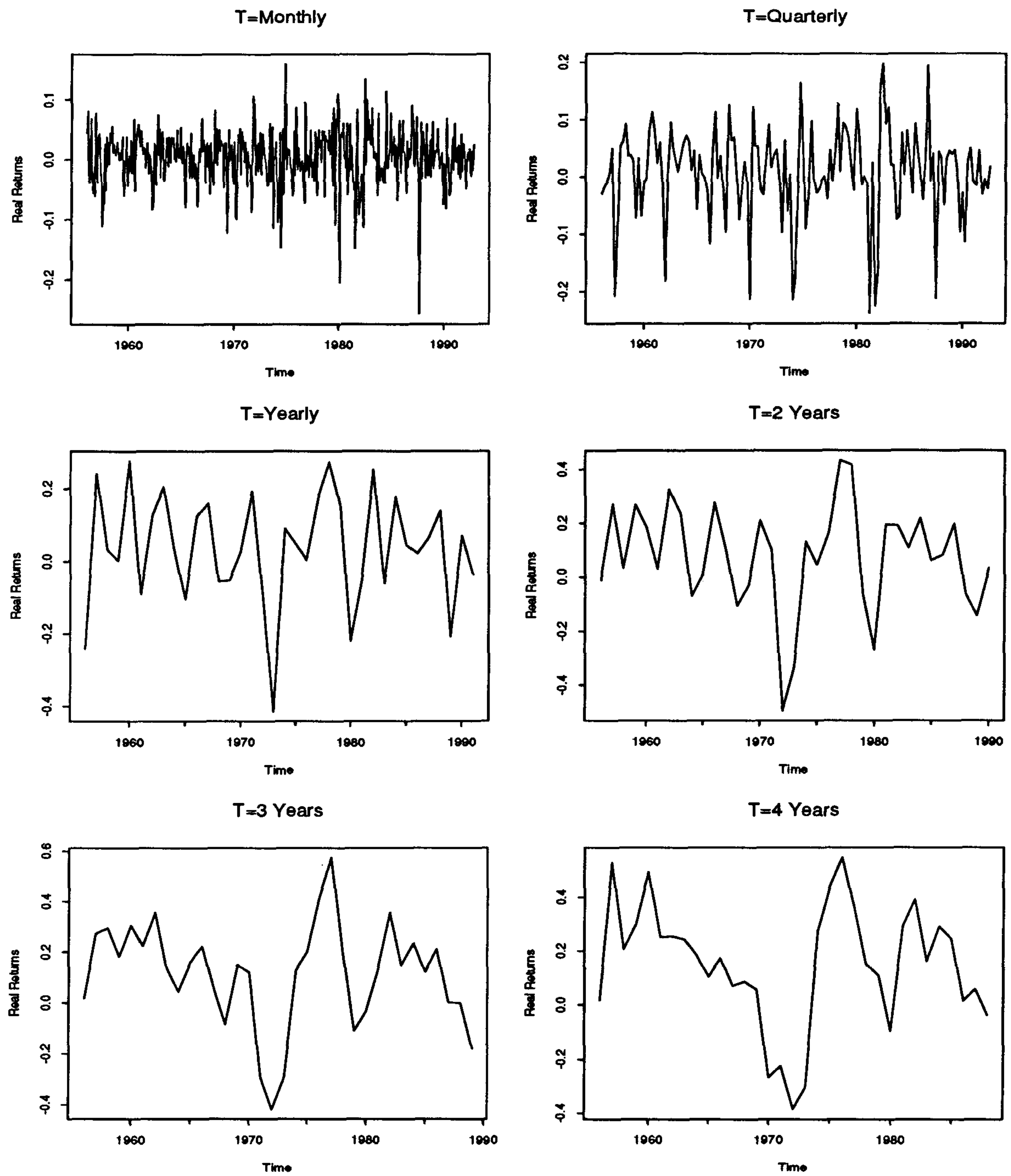

Figure 2.1: Plots of TSE Composite portfolio $\{R_{t,T}\}$ for return horizons, T.

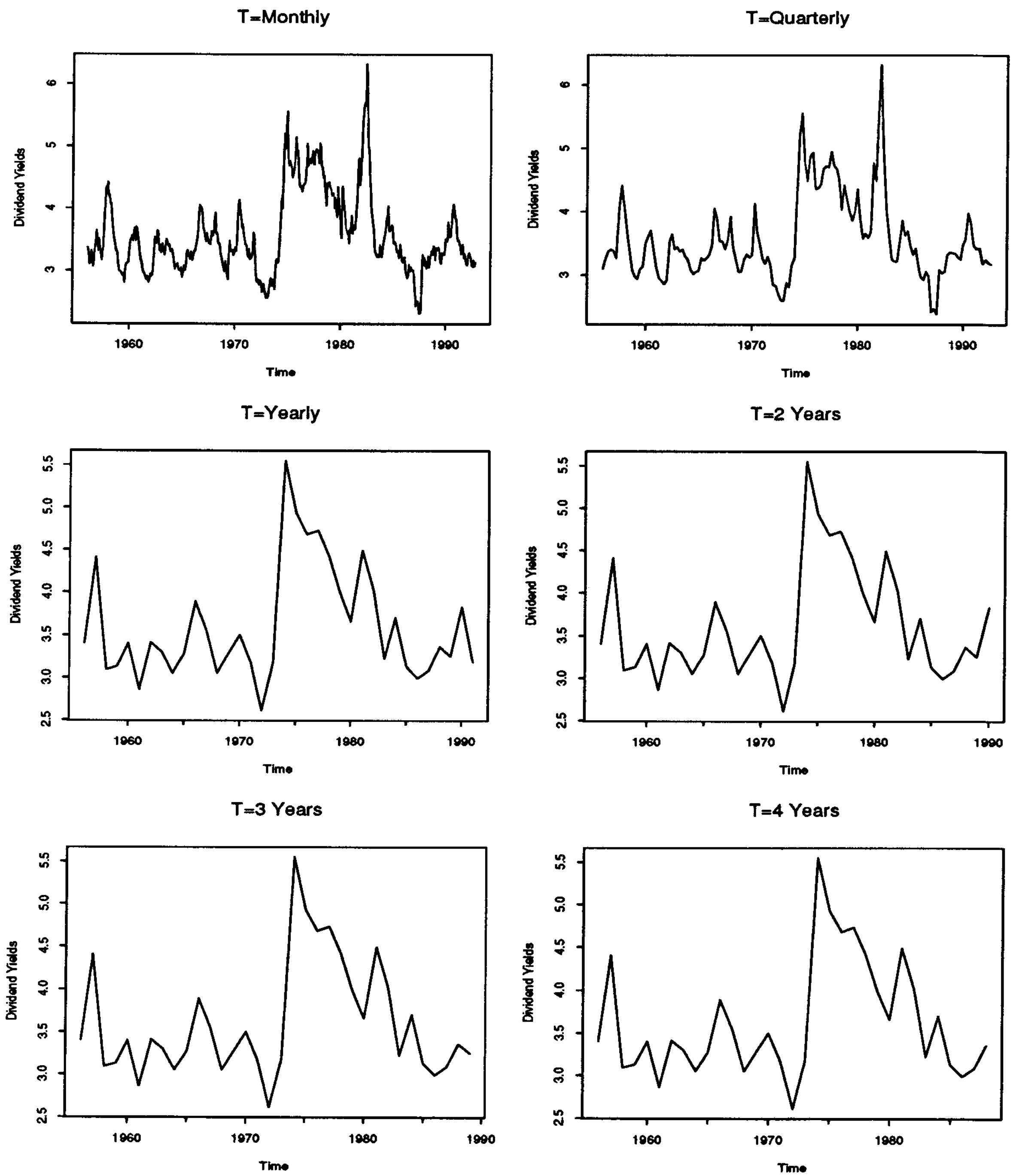

Figure 2.2: Plots of TSE Composite portfolio $\{DY_t\}$ for return horizons, T.

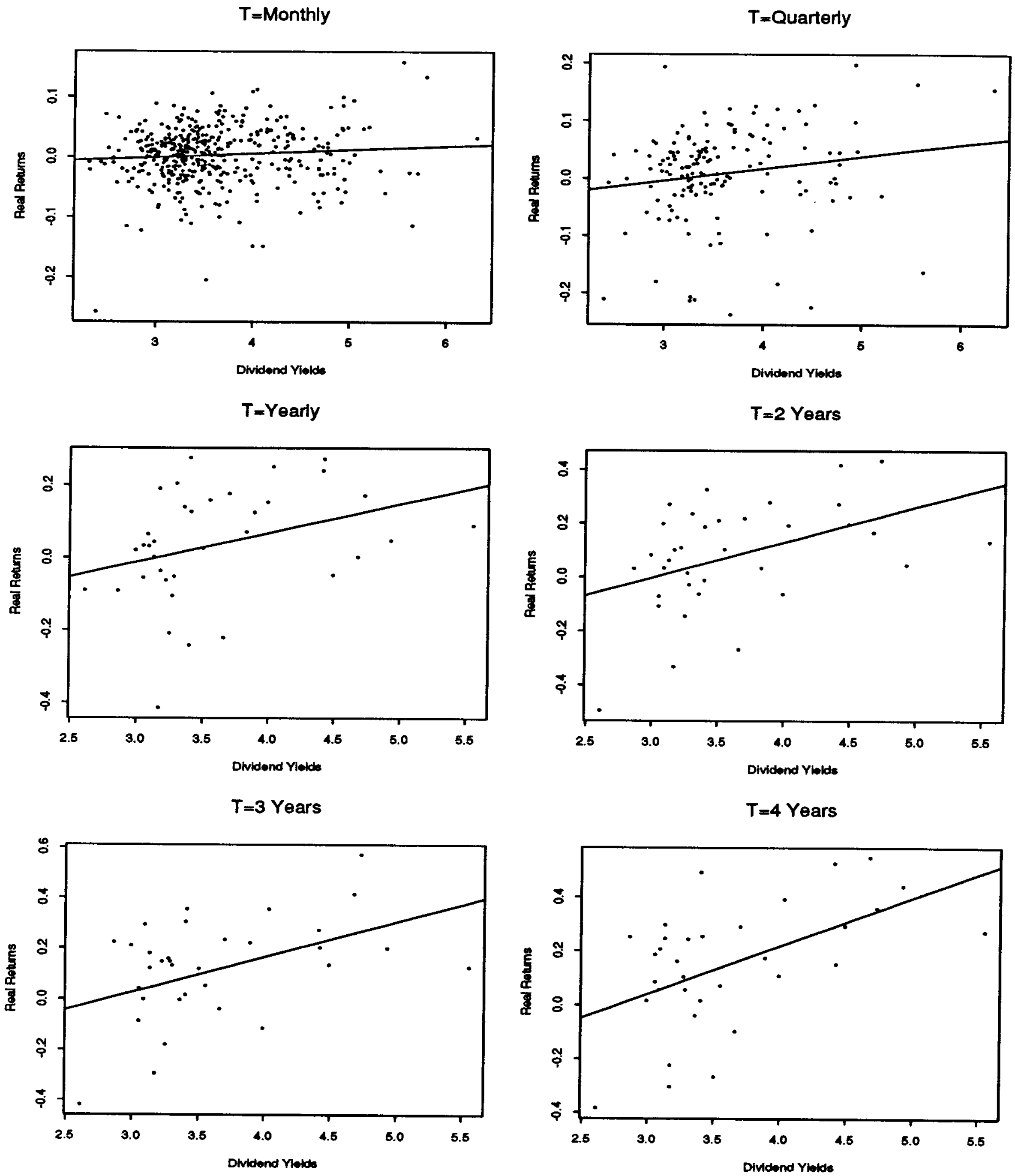

Figure 2.3: Plots of TSE Composite portfolio $\{R_{t,T}\}$ vs $\{DY_t\}$ for return horizons, T.

2.4.1 Regressions

Table 2.1 contains the results of applying the regression model (2.1) to the real returns $R_{t,T}$ for the TSE portfolios. The following statistics are given in the table:

- the slope estimate, $\widehat{\beta}_T$;
- the standard error of the slope estimate se($\widehat{\beta}_T$) as well as the Hansen and Hodrick [7] correction given in parenthesis for the overlapping returns;
- forecast performance, R^2;
- the residual variance $\widehat{\sigma_T^2}$;
- the first four auto–correlation estimates $\widehat{\rho}_1$, $\widehat{\rho}_2$, $\widehat{\rho}_3$, $\widehat{\rho}_4$;
- standard error of the auto–correlation estimate assuming lack of serial dependence, $se(\widehat{\rho}) = 1/\sqrt{N_T}$;
- the number of times the sample auto–correlation function exceeds $2 \times se(\widehat{\rho})$. The number of auto correlations assessed depends on the return horizon: monthly series – 40, quarterly – 20, annual – 10.

Table 2.1: Regression of real value–weighted TSE portfolio returns $(R_{t,T})$ on dividend yields (D_t/P_t), for differing return horizons T.

		Real Returns $R_{t,T}$				Auto–correlations					
T	N_T	$\widehat{\beta}_T$	se($\widehat{\beta}_T$)	R^2	$\widehat{\sigma_T^2}$	$se(\widehat{\rho}_\tau)$	$\widehat{\rho}_1$	$\widehat{\rho}_2$	$\widehat{\rho}_3$	$\widehat{\rho}_4$	E
Composite											
M	443	0.006	0.003	0.008	0.002	0.048	0.104	-0.053	0.071	0.028	4
Q	147	0.021	0.010	0.030	0.006	0.082	0.205	-0.014	-0.065	-0.077	1
1	36	0.081	0.038	0.118	0.022	0.167	0.020	-0.196	0.091	-0.013	0
2	35	0.132	0.046(0.053)	0.197	0.033	0.169	0.432	-0.108	-0.019	0.056	1
3	34	0.138	0.049(0.062)	0.200	0.036	0.171	0.616	0.199	-0.051	-0.047	1
4	33	0.178	0.051(0.071)	0.280	0.039	0.174	0.688	0.427	0.236	0.023	2
Metals and Minerals											
M	443	0.001	0.003	0.000	0.004	0.048	0.044	-0.057	0.010	-0.040	5
Q	147	0.004	0.009	0.001	0.011	0.082	0.077	-0.013	-0.105	-0.041	1
1	36	0.023	0.033	0.014	0.043	0.167	-0.111	-0.240	0.152	0.053	0
2	35	0.038	0.041(0.043)	0.025	0.066	0.169	0.310	-0.221	0.055	0.120	0
3	34	0.007	0.044(0.053)	0.001	0.074	0.171	0.550	0.179	-0.045	-0.026	1
4	33	0.034	0.050(0.066)	0.015	0.094	0.174	0.623	0.396	0.229	-0.137	2
Gold and Silver											
M	443	0.004	0.003	0.003	0.009	0.048	0.014	-0.069	0.000	0.081	1
Q	147	0.011	0.009	0.009	0.025	0.082	0.036	-0.070	-0.001	0.009	2
1	36	0.040	0.034	0.039	0.086	0.167	-0.186	-0.214	0.067	-0.190	0
2	35	0.092	0.042(0.043)	0.129	0.128	0.169	0.292	-0.284	-0.169	-0.237	1
3	34	0.102	0.044(0.047)	0.144	0.140	0.171	0.484	-0.020	-0.488	-0.350	4
4	33	0.108	0.048(0.048)	0.142	0.158	0.174	0.416	-0.004	-0.170	-0.383	3
Oil and Gas											
M	443	0.021	0.007	0.019	0.005	0.048	0.038	-0.009	0.096	-0.008	1
Q	147	0.068	0.021	0.064	0.015	0.082	0.091	0.073	0.028	0.014	0
1	36	0.251	0.079	0.228	0.051	0.167	0.190	-0.190	-0.168	-0.002	0
2	35	0.461	0.102(0.116)	0.384	0.084	0.169	0.406	-0.145	-0.202	-0.031	1
3	34	0.658	0.094(0.116)	0.603	0.072	0.171	0.542	0.029	-0.017	-0.028	1
4	33	0.730	0.088(0.101)	0.691	0.062	0.174	0.406	0.275	0.019	0.217	1
Paper and Forest Products											
M	443	0.000	0.002	0.000	0.004	0.048	0.141	-0.061	-0.036	0.032	6
Q	147	0.003	0.007	0.001	0.013	0.082	0.157	0.009	-0.030	-0.079	1
1	36	0.003	0.025	0.000	0.046	0.167	-0.120	-0.217	0.010	-0.138	0
2	35	0.010	0.033(0.036)	0.003	0.077	0.169	0.327	-0.285	-0.233	-0.139	0
3	34	-0.010	0.036(0.042)	0.003	0.092	0.171	0.523	-0.004	-0.319	-0.266	1
4	33	0.020	0.039(0.046)	0.009	0.106	0.174	0.538	0.134	-0.178	-0.391	2

Table 2.1: Continued

		Real Returns $R_{t,T}$				Auto–correlations					
T	N_T	$\widehat{\beta}_T$	se($\widehat{\beta}_T$)	R^2	$\widehat{\sigma_T^2}$	$se(\widehat{\rho}_T)$	$\widehat{\rho}_1$	$\widehat{\rho}_1$	$\widehat{\rho}_3$	$\widehat{\rho}_4$	E
Consumer Products											
M	443	0.003	0.002	0.005	0.002	0.048	0.105	-0.053	0.042	0.080	3
Q	147	0.012	0.007	0.019	0.007	0.082	0.166	0.091	-0.022	-0.043	1
1	36	0.047	0.029	0.071	0.029	0.167	-0.014	-0.160	0.225	-0.124	0
2	35	0.118	0.038(0.047)	0.225	0.045	0.169	0.460	0.041	0.113	0.033	1
3	34	0.149	0.042(0.060)	0.284	0.053	0.171	0.715	0.357	0.134	0.030	2
4	33	0.200	0.050(0.071)	0.345	0.069	0.174	0.693	0.484	0.372	0.093	4
Industrial Products											
M	443	0.002	0.003	0.001	0.003	0.048	0.132	-0.010	0.004	-0.041	3
Q	147	0.009	0.009	0.006	0.010	0.082	0.084	-0.145	-0.075	-0.100	0
1	36	0.034	0.029	0.039	0.024	0.167	-0.126	-0.231	0.207	0.052	0
2	35	0.045	0.037(0.039)	0.043	0.036	0.169	0.292	-0.218	0.127	0.203	0
3	34	0.032	0.037(0.050)	0.023	0.037	0.171	0.568	0.203	0.033	0.132	1
4	33	0.059	0.041(0.063)	0.061	0.046	0.174	0.630	0.440	0.334	0.067	2
Real Estate and Construction											
M	299	-0.007	0.005	0.007	0.007	0.058	0.232	0.045	0.141	0.017	3
Q	99	-0.020	0.018	0.013	0.031	0.101	0.230	0.102	-0.041	0.028	1
1	24	-0.015	0.090	0.001	0.164	0.204	0.279	0.109	-0.109	-0.111	0
2	23	0.104	0.149(0.180)	0.023	0.362	0.209	0.514	0.005	-0.214	-0.112	1
3	22	0.384	0.189(0.231)	0.171	0.474	0.213	0.476	0.084	-0.269	-0.251	1
4	21	0.498	0.186(0.233)	0.275	0.455	0.218	0.497	0.120	-0.255	-0.356	1
Transportation and Environmental Services											
M	443	0.002	0.002	0.003	0.004	0.048	0.089	0.005	0.061	0.038	2
Q	147	0.007	0.006	0.010	0.014	0.082	0.186	0.014	-0.046	-0.034	3
1	36	0.029	0.022	0.051	0.055	0.167	0.174	-0.147	0.050	-0.154	0
2	35	0.066	0.032(0.042)	0.111	0.115	0.169	0.527	-0.034	-0.135	-0.265	2
3	34	0.085	0.038(0.052)	0.133	0.146	0.171	0.661	0.156	-0.242	-0.405	3
4	33	0.103	0.043(0.057)	0.156	0.161	0.174	0.676	0.191	-0.163	-0.389	3
Pipelines											
M	443	0.002	0.002	0.003	0.003	0.048	0.007	-0.015	0.095	0.050	1
Q	147	0.008	0.005	0.016	0.009	0.082	0.104	-0.079	-0.115	-0.047	0
1	36	0.036	0.018	0.105	0.028	0.167	0.051	-0.150	0.049	-0.092	0
2	35	0.056	0.024(0.028)	0.144	0.046	0.169	0.431	-0.109	-0.090	0.012	1
3	34	0.064	0.026(0.036)	0.162	0.055	0.171	0.623	0.133	-0.138	-0.133	1
4	33	0.077	0.028(0.042)	0.201	0.063	0.174	0.670	0.350	0.056	-0.205	2

Table 2.1: Continued

		Real Returns $R_{t,T}$				Auto–correlations					
T	N_T	$\widehat{\beta}_T$	se($\widehat{\beta}_T$)	R^2	$\widehat{\sigma_T^2}$	$se(\widehat{\rho}_\tau)$	$\widehat{\rho}_1$	$\widehat{\rho}_2$	$\widehat{\rho}_3$	$\widehat{\rho}_4$	E
Utilities											
M	443	0.002	0.001	0.010	0.001	0.048	0.060	-0.072	-0.020	0.036	7
Q	147	0.007	0.003	0.031	0.004	0.082	0.068	-0.081	-0.107	0.140	0
1	36	0.024	0.012	0.099	0.012	0.167	0.033	0.028	-0.077	-0.067	0
2	35	0.039	0.016(0.022)	0.148	0.022	0.169	0.529	0.003	-0.089	-0.123	2
3	34	0.054	0.019(0.030)	0.193	0.030	0.171	0.669	0.283	-0.102	-0.088	2
4	33	0.064	0.021(0.037)	0.226	0.036	0.174	0.750	0.413	0.134	-0.033	3
Communications and Media											
M	443	0.003	0.002	0.004	0.003	0.048	0.123	0.015	0.108	-0.013	2
Q	147	0.008	0.006	0.013	0.009	0.082	0.173	-0.017	-0.012	-0.061	1
1	36	0.046	0.027	0.079	0.041	0.167	-0.047	-0.283	0.152	-0.068	0
2	35	0.094	0.035(0.038)	0.185	0.068	0.169	0.359	-0.232	-0.083	0.012	1
3	34	0.107	0.037(0.048)	0.210	0.076	0.171	0.604	0.069	-0.170	-0.143	1
4	33	0.139	0.039(0.055)	0.286	0.086	0.174	0.594	0.298	0.063	-0.299	2
Merchandising											
M	443	0.005	0.003	0.004	0.002	0.048	0.183	-0.005	0.049	0.090	7
Q	147	0.017	0.011	0.014	0.009	0.082	0.191	0.017	0.029	-0.125	2
1	36	0.086	0.049	0.083	0.040	0.167	-0.036	-0.208	0.280	0.106	0
2	35	0.143	0.064(0.073)	0.132	0.065	0.169	0.358	-0.119	0.232	0.249	1
3	34	0.129	0.064(0.093)	0.113	0.065	0.171	0.621	0.261	0.157	0.205	2
4	33	0.185	0.072(0.119)	0.174	0.080	0.174	0.703	0.513	0.441	0.223	5
Financial Services											
M	443	0.003	0.002	0.007	0.002	0.048	0.116	-0.048	0.045	0.063	4
Q	147	0.012	0.006	0.028	0.007	0.082	0.209	-0.090	-0.156	-0.089	2
1	36	0.047	0.027	0.084	0.027	0.167	-0.120	-0.166	0.244	-0.018	0
2	35	0.058	0.032(0.033)	0.091	0.038	0.169	0.288	-0.212	0.109	0.104	0
3	34	0.051	0.033(0.044)	0.070	0.039	0.171	0.555	0.117	-0.124	-0.187	1
4	33	0.072	0.036(0.050)	0.114	0.047	0.174	0.561	0.306	0.058	-0.307	1
Conglomerates											
M	443	0.002	0.002	0.004	0.004	0.048	0.064	-0.031	0.074	0.000	1
Q	147	0.006	0.005	0.008	0.012	0.082	0.136	-0.023	-0.001	-0.064	1
1	36	0.014	0.018	0.018	0.041	0.167	0.177	-0.219	0.011	0.010	0
2	35	0.019	0.028(0.033)	0.014	0.091	0.169	0.469	-0.141	-0.109	-0.070	1
3	34	0.005	0.032(0.041)	0.001	0.122	0.171	0.638	0.124	-0.214	-0.278	1
4	33	0.001	0.034(0.045)	0.000	0.139	0.174	0.680	0.267	-0.120	-0.382	2

We make the following observations about Table 2.1:

Composite – forecast power for all return horizons, R^2 increases with return horizon;

(1) **Metals and Minerals** – no forecast power, R^2 does not increase with return horizon;

(2) **Gold and Silver** – forecast power for the 2 to 4 year returns, R^2 increases with return horizon;

(3) **Oil and Gas** – forecast power for all return horizons, R^2 increases with return horizon;

(4) **Paper and Forest Products** – no forecast power, R^2 does not increase with return horizon T;

(5) **Consumer Products** – forecast power for 2 to 4 year returns, R^2 increases with return horizon;

(6) **Industrial Products** – no forecast power, R^2 increases with return horizon T;

(7) **Real Estate and Construction** – forecast power for 4 year returns, R^2 increases with return horizon;

(8) **Transportation and Environmental Services** – no forecast power, R^2 increases with return horizon;

(9) **Pipelines** – forecast power for 1 and 2 year returns, R^2 increases with return horizon;

(10) **Utilities** – forecast power for monthly, quarterly and annual returns, R^2 increases with return horizon;

(11) **Communications and Media** – forecast power for 2 to 4 year returns, R^2 increases with return horizon;

(12) **Merchandising** – no forecast power, R^2 increases with return horizon;

(13) **Financial Services** – no forecast power, R^2 does not increase with return horizon;

(14) **Conglomerates** – no forecast power, R^2 does not increase with return horizon T;

where forecast power means the slope estimate, $\widehat{\beta}_T$, is more than two standard errors from zero. For the overlapping annual returns the Hansen and Hodrick [7] corrections to the standard errors are used.

It is interesting to note that the results of Fama and French [4] do not apply to all TSE portfolios we saw above. That is, the forecast power, as measured by R^2, increases with return horizon. Fama and French [4] give a two–part explanation as to why this happens

(a) If expected returns have strong positive auto-correlation, rational forecasts of one–year returns one to four years ahead are highly correlated. As a consequence, the variance of the expected returns grows faster with the return horizon than the variance of unexpected returns – the variation of expected returns becomes a larger fraction of the variation of returns.

(b) Fama and French [4] claim that residual variances for regressions of returns on yields increase less than in proportion to the return horizon. They base their explanation on the so called discount–rate effect, which simply put states that the offsetting adjustment of current prices triggered by shocks to discount rates and expected returns. They find that estimated shocks to expected returns are indeed associated with opposite shocks to prices. The cumulative price effect of these shocks is roughly zero; on average, the expected future price increases implied by higher expected returns are offset by the immediate decline in the current price.

The corrections, given in Hansen and Hodrick [7] seem to do the correct thing, *i.e.* increases the estimate of the standard error whenever the residual auto-correlations are large. We would expect E, the number of exceedances of the auto–correlation function $\hat{\rho}_T$ above twice its standard error, to be about two for the monthly series, one for the quarterly series, and about one half for the annual series. Assuming a null hypothesis of no auto–correlation then we would expect 5% of the values to fall outside the $\pm 2 \times 1/\sqrt{N_T}$ limits. In many cases, there seem to be an inordinate number of auto-correlations exceeding these limits. This indicates the model given in 2.1 may be inadequate for these data. The next chapter discusses an alternative approach.

2.4.2 Out of Sample Forecasts

As was mentioned in Section 2.4.2 the R^2 values can be over stated when based on a single sample. This section, applies the method of Section 2.4.2 to the value-weighted real returns of TSE portfolios except for the Real Estate and Construction portfolio, due to its late inclusion in the TSE composite (see Table 1.1).

It was decided to choose a sliding window of 20 years to base the estimates on, which would lead to out of sample forecasts using the next available data point. The 20 year period was chosen to closely resemble the 30 year period chosen by Fama and French [4]. Specifically the period 20 year period January 1956 to December 1975. Thus we are providing out of sample forecasts for the period January 1976 to the end of the series. The following table shows the number of in sample and out of sample observations available for each return horizon T.

T	$N_{\text{in sample}}$	$N_{\text{out sample}}$
M	240	203
Q	80	67
1	20	16
2	20	15
3	20	14
4	20	13

Table 2.2 contains the out of sample forecast performance measured by R^2. It is quite evident that we don't do very well when forecasting out of the range of the data for the majority of the portfolios. For example, consider the composite portfolio, the in–sample R^2 are as follows: 0.008, 0.030, 0.118, 0.197, 0.200, 0.280 for the monthly, quarterly, and one to four year return horizons respectively. These values are quite inflated compared to the ones given in Table 2.2. As a further example, consider the consumer products portfolio, the in–sample R^2 are 0.005, 0.019, 0.071, 0.225, 0.284 and 0.345 for the six return horizons while in Table 2.2 the R^2 are all negative. This due to the fact the the dividends are changing are low over the latter part of the out-of-sample period compared to the remaining period. Thus, negative R^2 values could alert us to changing model characteristics.

2.4.3 Stationarity of Parameter Estimates

In this section we determine whether the regression model (2.1) applies for entire sample period. We will attempt to answer this question by examining the two largest samples, namely the monthly returns and the quarterly returns. It was decided quite arbitrarily to choose a window length of 5 years. That is, we are assuming that the parameters remain approximately constant over a five year period. Some sort of validation of this

Table 2.2: Out of Sample forecast power as measured by R^2

Portfolio	Monthly	Quarterly	Yearly	2 Year	3 Year	4 Year
Composite	-0.006	-0.028	0.022	0.084	0.081	0.060
Metals	-0.016	-0.053	-0.186	-0.248	-0.326	-0.673
Gold	-0.015	-0.044	-0.177	-0.232	-0.223	-0.091
Oil	0.007	0.027	0.168	0.297	0.588	0.585
Paper	-0.010	-0.028	-0.111	-0.226	-0.218	-0.341
Consumer	-0.011	-0.031	-0.117	-0.342	-0.971	-1.061
Industrial	-0.016	-0.059	-0.195	0.046	0.268	0.387
Transportation	-0.019	-0.071	-0.203	-0.198	-0.208	-0.334
Pipelines	0.008	0.025	0.098	0.213	0.275	0.249
Utilities	-0.004	-0.014	0.088	0.061	0.264	0.296
Communications	-0.003	0.000	-0.023	0.022	0.085	0.194
Merchandising	-0.035	-0.098	-0.209	-0.475	-0.439	-0.372
Financial	-0.021	-0.073	-0.001	-0.299	-0.286	-0.030
Conglomerates	-0.010	-0.036	-0.156	-0.255	-0.226	-0.197

value should be done, however, we did not pursue this avenue any further, since the topic of stationarity of parameter estimates is discussed in more generality in the next chapter. We also, decided that we would slide the window along by one time period.

Note that these assumptions provide a sliding window of 60 data points for the monthly series and 20 data points for the quarterly series. Figure 2.4 contains plots (solid lines) of the estimated intercept, estimated slope and estimated variance at time t and return horizon T denoted by $\widehat{\alpha}_{tT}$, $\widehat{\beta}_{tT}$ and $\widehat{\sigma}^2_{tT}$ respectively for the TSE Composite portfolio. Also, plotted are $\widehat{\alpha}_{tT} \pm 2 \times se(\widehat{\alpha}_{tT})$ and $\widehat{\beta}_{tT} \pm 2 \times se(\widehat{\beta}_{tT})$ (dotted lines). Note that when examining Figure 2.4 that successive estimates of α_T, β_T and σ^2_T are not distinct, in that each successive estimate contains data used to form the previous estimate. With this in mind, it is evident that the parameter values do not simply fluctuate about some mean level which would support the hypothesis of stationary or constant parameter values. The behavior of the parameter estimates seem to exhibit some of the properties of a random walk.

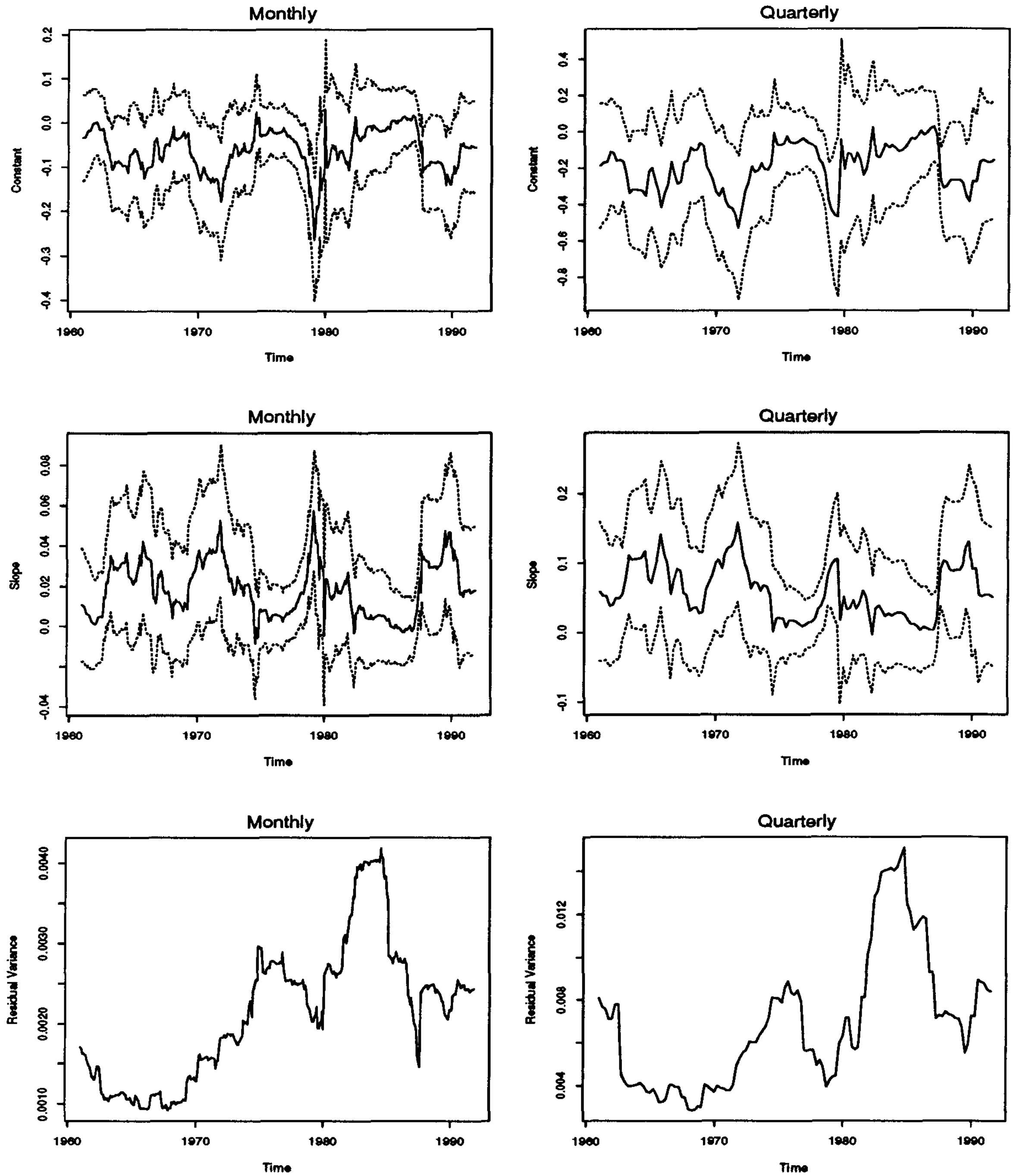

Figure 2.4: Plots of the time varying estimates of α_T, β_T and σ_T^2 for the TSE Composite portfolio for monthly and quarterly return horizons.

2.5 Conclusions

The results of Fama and French [4] which state that dividend yields show increased forecast power to predict real returns for increasing return horizons do not extend to all portfolios of the Toronto Stock Exchange. The Fama and French results only apply to the Composite portfolio and the Oil and Gas portfolio where there is forecast power for all return horizons as shown in Fama and French. However, the result of increasing forecast performance also apply to the following portfolios: Gold and Silver, Consumer Products, Industrial Products, Real Estate and Construction, Transportation, Pipelines, Utilities, Communications, and Merchandising.

However, the model suffers from some problems, most notably:

- residual auto–correlation
- dramatic decreases in out–of–sample R^2 indicating lack of stationarity
- Model 2.1 may not apply for all time periods as suggested in Fama and French [4] as evidenced by the changing residual variance results given in Section 2.4.3. It should be noted however, that Fama and French [4] mention the fact that the return variances are not constant throughout their sampling period. They present results for various sub periods of interest. We did not pursue this option since we had only a limited amount of data available.

For the reasons mentioned above, we should look for a more acceptable method of analysis which incorporates both of the deficiencies of the Fama and French [4] approach. This is the subject of the next chapter, namely Dynamic Linear Models.

Chapter 3

Theory of Dynamic Linear Models

3.1 Notation & Preliminaries

This section will introduce the notations, basic definitions and theorems used in the sequel.

Vectors and matrices will be denoted in bold face. For example if $\mathbf{X}$ is a p–dimensional vector and $\boldsymbol{\Sigma}$ is an $n \times p$ dimensional matrix, we denote them as follows

$$\mathbf{X} = (X_1, X_2, \ldots, _p)'$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{np} \end{bmatrix}$$

The density of a random vector will be denoted by $f(\mathbf{X})$. The joint density of two random vectors $\mathbf{X}$ and $\mathbf{Y}$ is $f(\mathbf{X}, \mathbf{Y})$ and $f(\mathbf{X}|\mathbf{Y})$ denotes the conditional density of $\mathbf{X}$ given $\mathbf{Y}$. Let E denote the expectation operator and let Var denote the variance operator.

Theorem 3.1 (Bayes) *Let $\mathbf{X}_1$ be a p_1–dimensional random vector and let $\mathbf{X}_2$ be a p_2–dimensional random vector and suppose we are given the conditional distribution $f(\mathbf{X}_2|\mathbf{X}_1)$. Then the conditional distribution $f(\mathbf{X}_1|\mathbf{X}_2)$ is given as follows:*

$$f(\mathbf{X}_1|\mathbf{X}_2) = \frac{f(\mathbf{X}_2|\mathbf{X}_1)f(\mathbf{X}_1)}{\int_{-\infty}^{\infty} f(\mathbf{X}_2|\mathbf{Z}_1)f(\mathbf{Z}_1)d\mathbf{Z}_1}$$

$$= \frac{f(\mathbf{X}_2|\mathbf{X}_1)f(\mathbf{X}_1)}{f(\mathbf{X}_2)}$$
$$\propto f(\mathbf{X}_1, \mathbf{X}_2)$$

Definition 3.1 *A p–dimensional random vector* $\mathbf{X}$ *is said to have a Multivariate Normal Distribution, with mean* $\boldsymbol{\mu}$ *and covariance matrix* $\boldsymbol{\Sigma}$*, denoted by* $\mathbf{X} \sim N[\boldsymbol{\mu}, \boldsymbol{\Sigma}]$ *iff*

$$f(\mathbf{X}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp(-(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})/2)$$

provide $\boldsymbol{\Sigma}$ *is positive definite.*

Definition 3.2 *A random variable* $\phi > 0$ *is said to have a Gamma distribution with parameters,* $n > 0$ *and* $d > 0$*, denoted by* $\phi \sim G[n, d]$ *iff*

$$f(\phi) = \frac{d^n}{\Gamma(n)} \phi^{n-1} \exp(-\phi d).$$

where $\Gamma(n)$ *is the gamma function*

$$\Gamma(n) = \int_0^{\infty} x^{n-1} \exp(-x) dx.$$

Note that $E[\phi] = n/d$ *and* $Var[\phi] = E[\phi]^2/n = n/d^2$*.*

Definition 3.3 *A p–dimensional random vector* $\mathbf{X}$ *is said to have a Multivariate Student's T Distribution with n degrees of freedom, mode* $\boldsymbol{\mu}$ *and scale matrix* $\boldsymbol{\Sigma}$*, denoted by* $\mathbf{X} \sim T_n[\boldsymbol{\mu}, \boldsymbol{\Sigma}]$ *iff*

$$f(\mathbf{X}) = \frac{n^{n/2}\Gamma((n+p)/2)}{\pi^{p/2}\Gamma(n/2)|\boldsymbol{\Sigma}|^{1/2}}(n+(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))^{-(n+p)/2} \; -\infty < X_j < \infty, i = 1, 2, \ldots, p \; .$$

Theorem 3.2 *Consider the following Multivariate Normal Distribution:*

$$\mathbf{X} \sim N[\boldsymbol{\mu}, \boldsymbol{\Sigma}]$$

which we can partition as follows:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \textit{and,} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Then we have the following conditional distribution:

$$(\mathbf{X}_1|\mathbf{X}_2) \sim N[\boldsymbol{\mu}_1(\mathbf{X}_2), \boldsymbol{\Sigma}_1(\mathbf{X}_2)]$$

where

$$\boldsymbol{\mu}_1(\mathbf{X}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$$

and

$$\boldsymbol{\Sigma}_1(\mathbf{X}_2) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Theorem 3.3 *Suppose that*

$$\phi \sim G[n/2, d/2],$$

and that the p–dimensional random vector $\mathbf{X}$ *is normally distributed conditional on* ϕ i.e.

$$(\mathbf{X}|\phi) \sim N[\boldsymbol{\mu}, \boldsymbol{\Sigma}\phi^{-1}].$$

Here the p–vector $\boldsymbol{\mu}$ *and the* $(p \times p)$ *matrix* $\boldsymbol{\Sigma}$ *are known. Then*

(a) $(\phi|\mathbf{X}) \sim G[n^*/2, d^*/2]$

where

$$n^* = n + p \quad \textit{and} \quad d^* = d + (\mathbf{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

(b) $\mathbf{X}$ *has a (marginal) multivariate Student's T distribution in p dimensions with n degrees of freedom, mode* $\boldsymbol{\mu}$ *and scale matrix* $\mathbf{R} = \boldsymbol{\Sigma}(d/n) = \boldsymbol{\Sigma}/E(\phi)$ *denoted by* $\mathbf{X} \sim T_n[\boldsymbol{\mu}, \mathbf{R}]$.

Proof: By Bayes Theorem (Theorem 3.1) we have:

$$\begin{aligned}
f(\phi|\mathbf{X}) &\propto f(\mathbf{X},\phi) \\
&\propto f(\mathbf{X}|\phi)f(\phi) \\
&\propto \frac{\phi^{p/2}}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp(-\phi(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})/2)\frac{d^{n/2}}{2^{n/2}\Gamma(n/2)}\phi^{n/2-1}\exp(-\phi d/2) \\
&\propto \phi^{(n+p)/2-1}\exp(-\phi\{(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))+d\}/2)
\end{aligned}$$

thus,

$$(\phi|\mathbf{X}) \sim G[n^*/2, d^*/2]$$

where $n^* = n+p$ and $d^* = d + (\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})$. This proves assertion (a).

Now to prove the assertion (b) consider the following:

$$\begin{aligned}
f(\mathbf{X}) &= f(\mathbf{X},\phi)/f(\phi|\mathbf{X}) \\
&= \frac{\frac{\phi^{p/2}}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp(-\phi/2(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))\frac{d^{n/2}}{2^{n/2}\Gamma(n/2)}\phi^{n/2-1}\exp(-\phi d/2)}{\frac{(d+(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))^{(n+p)/2}}{2^{(n+p)/2}\Gamma((n+p)/2)}\phi^{(n+p)/2-1}\exp(-\phi(d+(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))/2)} \\
&= \frac{\frac{\phi^{p/2}}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\frac{d^{n/2}}{2^{n/2}\Gamma(n/2)}\phi^{n/2-1}}{\frac{(d+(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))^{(n+p)/2}}{2^{(n+p)/2}\Gamma((n+p)/2)}\phi^{(n+p)/2-1}} \\
&\propto (d+(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}))^{-(n+p)/2} \\
&\propto (d+(\mathbf{X}-\boldsymbol{\mu})'\mathbf{R}^{-1}(d/n)(\mathbf{X}-\boldsymbol{\mu}))^{-(n+p)/2} \quad \text{letting } \mathbf{R} = d/n\boldsymbol{\Sigma} \\
&\propto (n+(\mathbf{X}-\boldsymbol{\mu})'\mathbf{R}^{-1}(\mathbf{X}-\boldsymbol{\mu}))^{-(n+p)/2} \quad \text{factoring out } d/n
\end{aligned}$$

Finally we have

$$f(\mathbf{X}) \sim T_n[\boldsymbol{\mu}, \mathbf{R}]$$

where $\mathbf{R} = \boldsymbol{\Sigma}(d/n) = \boldsymbol{\Sigma}/E[\phi]$. That is, $\mathbf{X}$ has a student multivariate student T distribution with n degrees of freedom, mode $\boldsymbol{\mu}$ and scale matrix $\mathbf{R}$. This proves assertion (b).

□

3.2 Dynamic Linear Models

Bayesian methods, not surprisingly, use Bayes theorem to update our knowledge about the current state of nature $\boldsymbol{\theta}$ once we have observed some data $\mathbf{Y}$. Let us assume that our initial state of knowledge can be expressed in terms of a prior distribution, $f(\boldsymbol{\theta})$. Let us also assume that we have a model which we assume generates our data $\mathbf{Y}$, given by $f(\mathbf{Y}|\boldsymbol{\theta})$. Note that once we observe $\mathbf{Y}$, our model can be interpreted as a likelihood for the various values of $\boldsymbol{\theta}$ and Bayes theorem tells us how to combine the prior and the likelihood to form the posterior:

$$
\begin{aligned}
f(\boldsymbol{\theta}|\mathbf{Y}) &\propto f(\mathbf{Y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) \\
\text{posterior} &\propto \text{likelihood} \times \text{prior}
\end{aligned}
$$

Note that the posterior updates our knowledge of the current state of nature $\boldsymbol{\theta}$, in light of the new data $\mathbf{Y}$ we have observed. That is, Bayes theorem gives us a sequential way of incorporating new information in our beliefs about the state of nature $\boldsymbol{\theta}$.

Dynamic models, as the name implies, are ones in which change is the driving force. Now applied to a time series, $\{Y_t\}_{t=1}^{N}$, dynamic models are models which change or adapt as time progresses. Now combining dynamic models with Bayesian methods gives one a unified approach to the problem of forecasting a time series $\{Y_t\}$, since the dynamic model tells us how the time series changes and the Bayesian approach tells us how the information we received about the time series can be incorporated into our state of knowledge.

Specifically, the approach given in West and Harrison [13] for Bayesian forecasting and dynamic modeling is,

(i) a sequential model definition;

(ii) structuring using parametric models with meaningful parameterization;

(iii) probabilistic representation of information about parameters;

(iv) forecasts derived as probability distributions.

Now, let D_t represent our state of knowledge at time t. The sequential approach bases our statements about Y_t, made at time $t-1$, on the information set available D_{t-1}. The parametric model at time t is $f(Y_t|\boldsymbol{\theta}_t, D_{t-1})$, which represents our belief as to how the data at time t is derived. As mentioned above we need a prior distribution about our parameters, *i.e.* $f(\boldsymbol{\theta}_t|D_{t-1})$. Now after we observe the data at time t we have a posterior, which in effect becomes the prior for the next time period. Thus, the prior–posterior pair effectively store the information about the defining parameter $\boldsymbol{\theta}_t$ as we move through time.

The following sections introduce various aspects of the Dynamic Linear Model and their Bayesian analysis.

3.3 Known Observational Variance V_t

This section will describe the most basic dynamic linear model, in which one assumes, that all defining parameters are known. The basic definition is as follows:

Definition 3.4 *The Univariate Normal Dynamic Linear Model is defined by:*
observation equation:

$$Y_t = \mathbf{F}_t\boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V_t],$$

system equation:

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N[0, \mathbf{W}_t],$$

initial prior:

$$(\boldsymbol{\theta}_0|D_0) \sim N[\mathbf{m}_0, \mathbf{C}_0],$$

where:

(a) $\mathbf{F}_t$ *is a known* $(p \times 1)$ *matrix;*

(b) $\mathbf{G}_t$ *is a known* $(p \times p)$ *matrix;*

(c) V_t *is a known constant;*

(d) $\mathbf{W}_t$ *is a known* $(p \times p)$ *matrix;*

for some prior moments $\mathbf{m}_0$ *and* $\mathbf{C}_0$*. The observational and evolution error sequences are assumed to be independent and mutually independent, and are independent of* $(\boldsymbol{\theta}_0|D_0)$*.*

Now we will describe the basic elements of the dynamic linear model:

- $\mathbf{F}_t$ plays the role of the regression matrix,
- $\boldsymbol{\theta}_t$ the vector of regression parameters also called the state vector,
- ν_t is the observational error with variance V_t,
- $\mathbf{G}_t$ is the evolution transfer matrix,
- $\boldsymbol{\omega}_t$ is the evolution error with known covariance matrix $\mathbf{W}_t$.

We assume that the information set at time t is updated as follows:

$$D_t = \{Y_t, D_{t-1}\}$$

The key result is given in the following theorem:

Theorem 3.4 *For the DLM of Definition 3.4, one–step forecast and posterior distributions are given, for each t, as follows:*

(a) Posterior at $t-1$*:*

$$(\boldsymbol{\theta}_{t-1}|D_{t-1}) \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}].$$

For some mean $\mathbf{m}_{t-1}$ *and variance matrix* $\mathbf{C}_{t-1}$*,*

(b) *Prior at t:*

$$(\boldsymbol{\theta}_t|D_{t-1}) \sim N[\mathbf{a}_t, \mathbf{R}_t],$$

where $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ *and* $\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t' + \mathbf{W}_t$

(c) *One-step forecast:*

$$(Y_t|D_{t-1}) \sim N[f_t, Q_t]$$

where $f_t = \mathbf{F}_t'\mathbf{a}_t$ *and* $Q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + V_t$

(d) *Posterior at t:*

$$(\boldsymbol{\theta}_t|D_t) \sim N[\mathbf{m}_t, \mathbf{C}_t],$$

with $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t$ *and* $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t'Q_t$*,*

where $\mathbf{A}_t = \mathbf{R}_t\mathbf{R}_tQ_t^{-1}$ *and* $e_t = Y_t - f_t$*.*

Proof: By induction on t. Assume a) holds, *i.e.*:

$$(\boldsymbol{\theta}_{t-1}|D_{t-1}) \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}].$$

And from the system equation we have:

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

where: $\boldsymbol{\theta}_{t-1} \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}]$ and $\boldsymbol{\omega}_t \sim N[0, \mathbf{W}_t]$ and $\boldsymbol{\theta}_{t-1}$ and $\boldsymbol{\omega}_t$ are independent of each other.

This implies that

$$\boldsymbol{\theta}_t \sim N[\mathbf{a}_t, \mathbf{R}_t]$$

where $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ and $\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t' + \mathbf{W}_t$.

Since the sum of two independent normal random variables is again a normal random variable with mean equaling the sum of the means and the variance equaling the sum of the variances. Thus (b) is established, *i.e.*

$$f(\boldsymbol{\theta}_t|D_{t-1}) = N[\mathbf{a}_t, \mathbf{R}_t]$$

Now using the observation equation $Y_t = \mathbf{F}_t'\boldsymbol{\theta}_t + \nu_t$ we have the following conditional density:

$$f(Y_t|\boldsymbol{\theta}_t, D_{t-1}) = N[f_t, Q_t]$$

where $f_t = \mathbf{F}_t'\mathbf{a}_t$ and $Q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + V_t$ since, Y_t is the sum of two independent normals. Thus (c) is established.

Now combining (b) $f(\boldsymbol{\theta}_t|D_{t-1})$ and (c) $f(Y_t|\boldsymbol{\theta}_t, D_{t-1})$ we have that $f(Y_t, \boldsymbol{\theta}_t|D_{t-1})$ is bivariate normal with covariance given as follows:

$$\begin{aligned}
\mathrm{Cov}(Y_t, \boldsymbol{\theta}_t|D_{t-1}) &= \mathrm{Cov}(\mathbf{F}_t'\boldsymbol{\theta}_t + \nu_t, \boldsymbol{\theta}_t|D_{t-1}) \\
&= \mathbf{F}_t'\mathrm{Cov}(\boldsymbol{\theta}_t, \boldsymbol{\theta}_t|D_{t-1}) + \mathrm{Cov}(\nu_t, \boldsymbol{\theta}_t|D_{t-1}) \\
&= \mathbf{F}_t'\mathrm{Var}(\boldsymbol{\theta}_t, \boldsymbol{\theta}_t|D_{t-1}) + 0 \\
&= \mathbf{F}_t'\mathbf{R}_t
\end{aligned}$$

Thus,

$$\left(\begin{matrix} Y_t \\ \boldsymbol{\theta}_t \end{matrix} \middle| D_{t-1} \right) \sim N\left[\begin{pmatrix} f_t \\ \mathbf{a}_t \end{pmatrix}, \begin{pmatrix} Q_t & \mathbf{F}_t'\mathbf{R}_t \\ \mathbf{F}_t\mathbf{R}_t & \mathbf{R}_t \end{pmatrix} \right]$$

Now we want to find the conditional density of $(\boldsymbol{\theta}_t|Y_t, D_{t-1})$ which is what we want in part (d) since $D_t = \{Y_t, D_{t-1}\}$. Note that since $(\boldsymbol{\theta}_t, Y_t|D_{t-1})$ is multivariate normal Theorem 3.2 applies. Thus if we label $\mathbf{X}_1 = \boldsymbol{\theta}_t$ and $\mathbf{X}_2 = Y_t$ we have

$$(\boldsymbol{\theta}_t|Y_t, D_{t-1}) \sim N[\mathbf{m}_t, \mathbf{C}_t]$$

where

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{F}_t\mathbf{R}_tQ_t^{-1}[Y_t - f_t]$$

and

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{F}_t\mathbf{R}_tQ_t^{-1}\mathbf{F}_t'\mathbf{R}_t$$

Now if we define $e_t = Y_t - f_t$, $\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_tQ_t^{-1}$, we have $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_te_t$ and $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_tQ_t\mathbf{A}_t'$. The induction is complete since by definition $(\boldsymbol{\theta}_0|D_0) \sim N[\mathbf{m}_0, \mathbf{C}_0]$. □

The vector $\mathbf{A}_t$ is known as the adaptive coefficient vector, since it changes the forecast errors e_t and $\mathbf{a}_t$ into the new estimate of the state vector $\mathbf{m}_t$

Now there are several stumbling blocks in the basic definition of the univariate DLM. Firstly, we do not know the observational variance, V_t, this problem will be solved in two parts. Section 3.4 will deal with constant observational variance case ($V_t = V$) and Section 3.6 will deal with a slowly varying observational variance V_t. Secondly, we do not know the evolutional variance $\mathbf{W}_t$, this problem will be dealt with in Section 3.5. Thirdly, we may not have any "reasonable" priors, this is dealt with in Section 3.7. Section 3.8 will describe the model assessment strategy and finally Section 3.9 will give two simple examples to illustrate the theory.

It should be noted that the model given in Definition 3.4 is closely related to the Kalman Filter (see Preistley [10] and references therein) with the exception of the initial prior distributional assumptions.

3.4 Unknown Constant Observational Variance V

This section generalizes the previous section to the case of constant unknown observational variance, that is we consider $V_t = V$ for all times. Consider modeling the precision $\phi = 1/V$, if for convenience we choose to apply a Gamma prior to ϕ, and combining it with the new information expressed in terms of normal densities, then the posterior distribution also turns out to be a Gamma distribution with a simple updating of the parameters. Also a the expected value of a Gamma with parameters n_0 and d_0 has expected value $n_0/d_0 = 1/S_0$ where S_0 is the prior estimate of the observational variance V. Using these facts we generalize the definition and theorem of the previous section to incorporate the unknown observational variance in the form of a new definition and theorem discussed presently.

Definition 3.5 *The General Univariate Dynamic Linear Model with unknown constant variance is defined by: observation equation:*

$$Y_t = \mathbf{F}_t \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V],$$

system equation:

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N[0, \mathbf{W}_t],$$

initial priors:

$$\begin{aligned}(\boldsymbol{\theta}_0 | D_0, \phi) &\sim N[\mathbf{m}_0, V\mathbf{C}_0^*], \\ (\phi | D_0) &\sim G[n_0/2, d_0/2],\end{aligned}$$

where $\phi = V^{-1}$.

Theorem 3.5 *In the univariate DLM of Definition , one-step forecast and posterior distributions are given, for each t, as follows:*

(a) Conditional on V:

$$\begin{aligned}(\boldsymbol{\theta}_{t-1} | D_{t-1}, V) &\sim N[\mathbf{m}_{t-1}, V\mathbf{C}_{t-1}^*], \\ (\boldsymbol{\theta}_t | D_{t-1}, V) &\sim N[\mathbf{a}_t, V\mathbf{R}_t^*], \\ (Y_t | D_{t-1}, V) &\sim N[f_t, VQ_t^*], \\ (\boldsymbol{\theta}_t | D_t, V) &\sim N[\mathbf{m}_t, V\mathbf{C}_t^*],\end{aligned}$$

with $\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1}$, *and* $\mathbf{R}_t^* = \mathbf{G}_t \mathbf{C}_{t-1}^* \mathbf{G}_t' + \mathbf{W}_t^*$, $f_t = \mathbf{F}_t' \mathbf{a}_t$ *and* $Q_t^* = 1 + \mathbf{F}_t' \mathbf{R}_t^* \mathbf{F}_t$, *and*

$$\begin{aligned}\mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t \\ \mathbf{C}_t^* &= \mathbf{R}_t^* - \mathbf{A}_t \mathbf{A}_t' Q_t^*\end{aligned}$$

where $e_t = Y_t - f_t$ *and* $\mathbf{A}_t = \mathbf{R}_t^* \mathbf{F}_t / Q_t^*$.

(b) For precision $\phi = V^{-1}$

$$\begin{aligned}(\phi|D_{t-1}) &\sim G[n_{t-1}/2, d_{t-1}/2],\\ (\phi|D_t) &\sim G[n_t/2, d_t/2],\end{aligned}$$

where $n_t = n_{t-1} + 1$ and $d_t = d_{t-1} + e_t^2/Q_t^*$

(c) Unconditional on V:

$$\begin{aligned}(\boldsymbol{\theta}_{t-1}|D_{t-1}) &\sim T_{n_{t-1}}[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}],\\ (\boldsymbol{\theta}_t|D_{t-1}) &\sim T_{n_{t-1}}[\mathbf{a}_t, \mathbf{R}_t],\\ (Y_t|D_{t-1}) &\sim T_{n_{t-1}}[f_t, Q_t],\\ (\boldsymbol{\theta}_t|D_t) &\sim T_{n_t}[\mathbf{m}_t, \mathbf{C}_t],\end{aligned}$$

where $\mathbf{C}_{t-1} = S_{t-1}\mathbf{C}_{t-1}^*$, $\mathbf{R}_t = S_{t-1}\mathbf{R}_t^*$, $Q_t = S_{t-1}Q_t^*$ and $\mathbf{C}_t = S_t\mathbf{C}_t^*$, with $S_{t-1} = d_{t-1}/n_{t-1}$ and $S_t = d_t/n_t$.

(d) Operational definition of updating equations:

$$\begin{aligned}\mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t,\\ \mathbf{C}_t &= (S_t/S_{t-1})[\mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t'Q_t],\\ S_t &= d_t/n_t,\end{aligned}$$

and

$$\begin{aligned}n_t &= n_{t-1} + 1\\ d_t &= d_{t-1} + S_{t-1}e_t^2/Q_t\end{aligned}$$

where $Q_t = S_{t-1} + \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t$ and $\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t/Q_t$.

Proof: Part (a) follows directly from Theorem 3.4. The proof of remainder is again by induction on t. In order to prove part (b), assume

$$(\phi|D_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2],$$

and from (a) we have

$$(Y_t|D_{t-1}, \phi) \sim N[f_t, Q_t^*/\phi],$$

Hence, by Theorem 3.3 (a) we have

$$(\phi|D_{t-1}, Y_t) = (\phi|D_t) \sim G[n_t/2, d_t/2] \quad \text{since } D_t = \{D_{t-1}, Y_t\}$$

where

$$n_t = n_{t-1} + 1 \quad \text{since } Y_t \text{ is univariate } i.e.\ p = 1$$

and

$$\begin{aligned} d_t &= d_{t-1} + (Y_t - f_t)^2/Q_t^* \\ &= d_{t-1} + e_t^2/Q_t^* \end{aligned}$$

and (b) is proved.

The results of (c) can be proved in the following manner. For example, in order to prove $(Y_t|D_{t-1}) \sim T_{n_{t-1}}[f_t, Q_t]$ we proceed as follows:

we know from part (a) that

$$(Y_t|D_{t-1}, \phi) \sim N[f_t, VQ_t^*]$$

and from (b) we have

$$(\phi|D_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2]$$

thus by Theorem 3.3 (b) we have

$$(Y_t|D_{t-1}) \sim T_{n_{t-1}}[f_t, Q_t]$$

where $Q_t = S_{t-1}Q_t^*$ with $S_{t-1} = d_{t-1}/n_{t-1}$ the other results can be shown similarly. Finally (d) follows from (c) as follows: for example,

$$\mathbf{C}_t = S_t\mathbf{C}_t^*$$

$$
\begin{aligned}
&= S_t[\mathbf{R}_t^* - \mathbf{A}_t\mathbf{A}_t'Q_t^*] \\
&= S_t[\mathbf{R}_t/S_{t-1} - \mathbf{A}_t\mathbf{A}_t'Q_t/S_{t-1}] \\
&= (S_t/S_{t-1})[\mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t'Q_t]
\end{aligned}
$$

The theorem is now proved by induction since the results hold for at $t = 0$ and the initial priors given in Definition 3.5. □

Note that $S_t = d_t/n_t$ estimates $V = 1/\phi$.

3.5 Discount Factors

This section will deal with the problem of specifying the state evolution variance $\mathbf{W}_t$, which controls the amount of stochastic variation in the state vector and hence controls the model stability over time. Consider the following,

$$\mathrm{Var}(\boldsymbol{\theta}_{t-1}|D_{t-1}) = \mathbf{C}_{t-1}$$

i.e. the prior variance of the current state vector. Now using the updating equations, the posterior variance is given by

$$\mathrm{Var}(\boldsymbol{\theta}_{t-1}|D_{t-1}) = \mathbf{P}_t + \mathbf{W}_t = \mathbf{R}_t.$$

where $\mathbf{P}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$. Thus, we see that $\mathbf{W}_t$ has the effect of increasing the variance of the state vector, which can also be seen as a loss of information about the state vector, $\boldsymbol{\theta}_t$, from time $t-1$ to time t.

Ameen and Harrison [1], suggest that a multiplicative rate of information decay is appropriate, since the relative magnitudes of $\mathbf{W}_t$ and $\mathbf{P}_t$ are important. That is, consider modeling, the information decay by

$$\mathbf{R}_t = \mathbf{P}_t/\delta_t.$$

where $0 < \delta_t \leq 1$. The dependence of δ on time allows the possibility of a different rate of information decay at different times. Thus, from time $t-1$ to time t the loss of information is $100(1-\delta_t)/\delta_t\%$ and

$$\mathbf{R}_t = \mathbf{P}_t + \mathbf{W}_t \quad \text{and} \quad \mathbf{W}_t = \mathbf{P}_t(1-\delta_t)/\delta_t$$

Note that this implies that for all components of the state vector, the information decays at the same rate.

A more general approach to discounting is the method of component discounting discussed in West and Harrison [13]. This approach involves partitioning the state vector $\boldsymbol{\theta}_t$ into h components, such that $p = p_1 + p_2 + \ldots + p_h$ as follows:

$$\boldsymbol{\theta}_t' = (\boldsymbol{\theta}_{1t}', \ldots, \boldsymbol{\theta}_{ht}'),$$

with

$$\mathbf{F}_t' = (\mathbf{F}_{1t}', \ldots, \mathbf{F}_{ht}'),$$

$$\begin{aligned} \mathbf{G}_t &= \text{block diag}[\mathbf{G}_{1t}, \ldots, \mathbf{G}_{ht}] \\ &= \begin{bmatrix} \mathbf{G}_{1t} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{2t} & \ldots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{G}_{ht} \end{bmatrix}, \end{aligned}$$

and

$$\mathbf{W}_t = \text{block diag}[\mathbf{W}_{1t}, \ldots, \mathbf{W}_{ht}]$$

$$
= \begin{bmatrix} \mathbf{W}_{1t} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{2t} & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{W}_{ht} \end{bmatrix}.
$$

Now, as before $\mathbf{P}_t$ represents the variance before the addition of the evolution noise term. Form a block diagonal matrix corresponding to the partition of $\boldsymbol{\theta}_t$ as follows:

$$
\mathbf{P}_{it} = \mathrm{Var}(\mathbf{G}_{it}\boldsymbol{\theta}_{it}|D_{t-1}) \quad (i = 1, \dots, h).
$$

Now, if we discount each component separately, we need discount factors $0 < \delta_{t1}, \dots, \delta_{th} \leq 1$, which correspond to the partitions of $\boldsymbol{\theta}_t$. Then, the evolution variance, $\mathbf{W}_t$ is given as follows:

$$
\mathbf{W}_t = \begin{bmatrix} \mathbf{P}_{1t}(1-\delta_{t1})/\delta_{t1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{2t}(1-\delta_{t2})/\delta_{t2} & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}_{ht}(1-\delta_{th})/\delta_{th} \end{bmatrix}
$$

and finally,

$$
\mathbf{R}_t = \mathbf{P}_t + \mathbf{W}_t.
$$

This approach allows the various components of $\boldsymbol{\theta}_t$ to suffer information loses at differing rates. For example, some components of the state vector may be more robust to information loss and would hence have a discount factor near 1, and other components may change quite rapidly and hence would have a smaller discount factor.

Note that in most practical situations we can remove the dependence of the discount factors on t.

3.6 Discounted Variance Learning

In the previous sections we have assumed that the unknown observational variance, V_t, is constant, namely V. Ameen and Harrison [1] and West and Harrison [13] discuss a relaxation in this assumption by allowing for stochastic changes in V_t. The key idea is to introduce a random walk component to the precision estimate $\phi_t = 1/V_t$. Now, at time $t-1$, the precision has posterior,

$$(\phi_{t-1}|D_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2].$$

Now modeling the stochastic variation as a random walk, we have

$$\phi_t = \phi_{t-1} + \psi_t$$

where ψ_t is un-correlated with $\phi_{t-1}|D_{t-1}$. Let

$$\psi_t \sim [0, U_t]$$

denote a distribution for ψ_t, of form unspecified, with mean 0 and variance U_t. Now we know the following,

$$E[\phi_{t-1}|D_{t-1}] = n_{t-1}/d_{t-1} = 1/S_{t-1}$$

and

$$\mathrm{Var}[\phi_{t-1}|D_{t-1}] = 2n_{t-1}/d_{t-1}^2 = 2/(n_{t-1}S_{t-1}^2).$$

Now using the updating equation we see that the mean is unchanged and the variance of ϕ_t increases to

$$\mathrm{Var}[\phi_t|D_{t-1}] = U_t + 2/(n_{t-1}S_{t-1}^2).$$

However, as in the previous section (Section 3.5) on Discount factors, it is practically useful to think of variance increases in a multiplicative sense; thus

$$\mathrm{Var}[\phi_t|D_{t-1}] = 2/(\kappa_t n_{t-1}S_{t-1}^2).$$

Where κ_t, $(0 < \kappa_t \leq 1)$, is implicitly defined via

$$U_t = \text{Var}[\phi_{t-1}|D_{t-1}](\kappa_t^{-1} - 1).$$

Ameen and Harrison [1] recommend constraining κ_t to the range $(.95, 1)$. They also recommend removing the dependence on t, that is, assume that $\kappa_t = \kappa$ for all t. Thus as in the section on Discount factors, κ represents the amount of information loss about the precision estimate ϕ_t moving from time $t-1$ to t.

The following definition and theorem show how the discount factor is incorporated into the analysis.

Definition 3.6 *The General Univariate Dynamic Linear Model with unknown stochastic observational variance, V_t, is defined by: observation equation:*

$$Y_t = \mathbf{F}_t\boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V_t],$$

system equation:

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N[0, \mathbf{W}_t],$$

initial priors:

$$\begin{aligned} (\boldsymbol{\theta}_0|D_0, \phi_0) &\sim N[\mathbf{m}_0, V_0\mathbf{C}_0*], \\ (\phi_0|D_0) &\sim G[n_0/2, d_0/2], \end{aligned}$$

where $\phi_0 = V_0^{-1}$.

Theorem 3.6 *In the above definition (3.6) , one-step forecast and posterior distributions are given, for each t, as follows:*

$$\begin{aligned} (\phi_{t-1}|D_{t-1}) &\sim G[n_{t-1}/2, d_{t-1}/2], \\ (\phi_t|D_{t-1}) &\sim G[\kappa_t n_{t-1}/2, \kappa_t d_{t-1}/2], \\ (\phi_t|D_t) &\sim G[n_t/2, d_t/2], \end{aligned}$$

$$
\begin{aligned}
(\boldsymbol{\theta}_{t-1}|D_{t-1}) &\sim T_{n_{t-1}}[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}], \\
(\boldsymbol{\theta}_t|D_{t-1}) &\sim T_{n_{t-1}}[\mathbf{a}_t, \mathbf{R}_t], \\
(Y_t|D_{t-1}) &\sim T_{\kappa_t n_{t-1}}[f_t, Q_t], \\
(\boldsymbol{\theta}_t|D_t) &\sim T_{n_t}[\mathbf{m}_t, \mathbf{C}_t],
\end{aligned}
$$

where,

$$
\begin{aligned}
\mathbf{a}_t &= \mathbf{G}_t\mathbf{m}_{t-1}, \\
\mathbf{R}_t &= \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}'_t + \mathbf{W}_t, \\
f_t &= \mathbf{F}'_t\mathbf{a}_t, \\
Q_t &= S_{t-1} + \mathbf{F}'_t\mathbf{R}_t\mathbf{F}_t, \\
\mathbf{A}_t &= \mathbf{R}_t\mathbf{F}_t/Q_t \\
e_t &= Y_t - f_t, \\
n_t &= \kappa_t n_{t-1} + 1, \\
d_t &= \kappa_t d_{t-1} + S_{t-1}e_t^2/Q_t, \\
S_t &= d_t/n_t, \\
\mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t, \\
\mathbf{C}_t &= (S_t/S_{t-1})[\mathbf{R}_t - \mathbf{A}_t\mathbf{A}'_t Q_t].
\end{aligned}
$$

Proof: The proof proceeds along the same lines as the proof of Theorem 3.5 and will not be presented here. □

Ameen and Harrison [1] note that we can in most practical situations we can remove the dependence of κ on t.

3.7 Reference Analysis

What if the forecaster is uncertain about an initial prior for an analysis or if one wishes to have a baseline for comparison of a given prior with the "non–informative" prior. This section gives some results based on the "non–informative" prior distribution of the parameters $\boldsymbol{\theta}_0$ and ϕ. Thus if the forecaster is unable or unwilling to specify the initial priors then reference analysis gives a data based alternative.

Theorem 3.7 *For the model defined in (3.5) let the initial prior information be represented by*

$$f(\boldsymbol{\theta}_t, \phi | D_{t-1}) \propto V^{-1}$$

Then the joint prior and posterior distributions of the state vector and the observation variance at time $t = 1, 2, \ldots$ are given by

$$\begin{aligned} f(\boldsymbol{\theta}_t, V | D_{t-1}) &\propto V^{-1+\gamma_{t-1}/2} \exp\{-1/2V^{-1}(\boldsymbol{\theta}_t'\mathbf{H}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}_t'\mathbf{h}_t + \lambda_t)\} \\ f(\boldsymbol{\theta}_t, V | D_{t-1}) &\propto V^{-1+\lambda_t/2} \exp\{-1/2V^{-1}(\boldsymbol{\theta}_t'\mathbf{K}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}_t'\mathbf{k}_t + \delta_t)\} \end{aligned}$$

where

$$\begin{aligned} \mathbf{H}_t &= \mathbf{W}_t^{*-1} - \mathbf{W}_t^{*-1}\mathbf{G}_t\mathbf{P}_t^{-1}\mathbf{G}_t'\mathbf{W}_t^{*-1} \\ \mathbf{P}_t &= \mathbf{G}_t'\mathbf{W}_t^{*-1}\mathbf{G}_t + \mathbf{K}_{t-1} \\ \mathbf{h}_t &= \mathbf{W}_t^{*-1}\mathbf{G}_t\mathbf{P}_t^{-1}\mathbf{k}_t \\ \mathbf{K}_t &= \mathbf{H}_t + \mathbf{F}_t\mathbf{F}_t' \\ \mathbf{k}_t &= \mathbf{h}_t + \mathbf{F}_t Y_t \\ \gamma_t &= \gamma_{t-1} + 1 \\ \lambda_t &= \delta_{t-1} - \mathbf{k}_{t-1}'\mathbf{P}_t^{-1}\mathbf{k}_{t-1} \\ \delta_t &= \lambda_t + Y_t^2 \end{aligned}$$

with initial values $\mathbf{H}_1 = \mathbf{0}$, $\mathbf{h}_1 = \mathbf{0}$, $\lambda_1 = 0$ *and* $\gamma_1 = 0$. *Provided that* $\mathbf{W}_t^*$ *are non–singular and known for each time.*

Proof: See Pole and West [9].

Now, following Pole and West [9], we revert to the usual updating equations once sufficient observations have been processed to give rise to a proper posterior distribution. In the general dynamic linear model (3.5) this happens after $p+1$ observations have been processed, where p is the dimension of the parameter vector $\boldsymbol{\theta}$. The following theorem from Pole and West proves this result.

Theorem 3.8 *For* $t = p + 1$ *the posterior distribution is*

$$\begin{aligned}(\boldsymbol{\theta}_t|D_t) &\sim T_{n_t}[\mathbf{m}_t, \mathbf{C}_t] \\ (V^{-1}|D_t) &\sim G[n_t/2, d_t/2]\end{aligned}$$

with

$$\mathbf{C}_t = S_t\mathbf{K}_t^{-1} \quad and \quad \mathbf{m}_t = \mathbf{K}_t^{-1}\mathbf{k}_t$$

where $n_t = 1$, *and* $S_t = d_t = e_t^2/Q_t^*$.

Proof: See Pole and West [9].

3.7.1 Case of $\mathbf{W}_t$ Unknown

This section will present a result which frees us from the assumption of known $\mathbf{W}_t$, the system equation variance. Ameen and Harrison [1] have a avoided the problem of specification of $\mathbf{W}_t$, by the use of discount techniques. However, these methods do not apply in the reference analysis for $t < p + 1$ because the posterior covariances do not yet exist. The proposed method is to assume that $\mathbf{W}_t = \mathbf{0}$ for $t = 1, 2, \ldots, p + 1$

The rationale behind this approach is as follows. In a reference analysis with $p + 1$ parameters, we need $p + 1$ observations to obtain a fully specified proper joint posterior distribution for $\boldsymbol{\theta}_t$ and V. At time $p+1$ we have one observation per parameter. Now $\mathbf{W}$ allows for changes in the parameter estimates, however, it is not possible to estimate any changes over the first $p+1$ time points so setting $\mathbf{W}_t = \mathbf{0}$ for $t = 1, 2, \ldots, p+1$, results in no loss of information. At time $p+1$ we can revert to the usual updating as was shown in Theorem (3.8).

Theorem 3.9 *In the framework of Theorem (3.7) with $\mathbf{W}_t = \mathbf{0}$ the prior and posterior distributions of $\boldsymbol{\theta}_t$ and V have the forms of Theorem (3.7) with the recursions modified as follows:*

$$
\begin{aligned}
\mathbf{H}_t &= \mathbf{G}_t^{-1\prime}\mathbf{K}_{t-1}\mathbf{G}_t^{-1} \\
\mathbf{h}_t &= \mathbf{G}_t^{-1\prime}\mathbf{k}_{t-1} \\
\mathbf{K}_t &= \mathbf{H}_t + \mathbf{F}_t\mathbf{F}_t' \\
\mathbf{k}_t &= \mathbf{h}_t + \mathbf{F}_t Y_t \\
\gamma_t &= \gamma_{t-1} + 1 \\
\lambda_t &= \delta_{t-1} \\
\delta_t &= \lambda_t + Y_t^2
\end{aligned}
$$

with initial values $\mathbf{H}_1 = \mathbf{0}$, $\mathbf{h}_1 = \mathbf{0}$, $\lambda_1 = 0$ and $\gamma_1 = 0$.

Proof: See Pole and West [9].

3.8 Model Assessment

How do we determine whether our model does an adequate job in terms of forecast performance. We concentrate on overall measures of forecast performance as opposed to the approaches of West [14] and West and Harrison [15]. Their approach is as follows: Assess the performance of the model at each time point and determine if (a) a change in the model is necessary or (b) an observation should be considered an outlier. Their approach is based on cumulative Bayes factors.

We will instead concentrate on overall model forecast performance, recognizing that the continual assessment techniques mentioned above will provide better results. The rationale behind our approach is to have an objective method of comparison with the standard regression methods discussed in chapter 2.

The measures considered are the following:

Definition 3.7 *The Mean Square Prediction Error, denoted by MSE, is*

$$MSE = \sum_{t=s}^{N} e_t^2 / N$$

Definition 3.8 *The Mean Absolute Prediction Error, denoted by MAD, is*

$$MAD = \sum_{t=s}^{N} |e_t| / N$$

Definition 3.9 *The observed predictive density*

$$\begin{aligned} PD = f(Y_N, Y_{N-1}, \ldots, Y_s | D_0) &= \prod_{t=s}^{t=N} f(Y_t|_{t-1}) \\ &= \prod_{t=s}^{t=N} \frac{df^{df/2}\Gamma((df+1)/2)}{\pi^{1/2}\Gamma(df/2)Q_t^{1/2}} (df + (Y_t - f_t)^2/Q_t)^{-(df+1)/2} \end{aligned}$$

where $df = \kappa n_{t-1}$. Thus PD is the product of the sequence of one–step forecast densities evaluated at the actual observation. Note that all products start at s, if a reference analysis was done, then $s = (p+1)+1$, where p is the dimension of $\boldsymbol{\theta}$, otherwise $s = 1$.

Since the discount factors, $\delta_1, \ldots, \delta_h$ and κ are assumed to be part of the initial information set D_0. We can consider the predictive density PD as a likelihood for the discount factors. That is, we have the following definition,

Definition 3.10 *The Log Likelihood for a parameter* $\boldsymbol{\eta} = (\delta_1, \ldots, \delta_h, \kappa)$*, denoted by* $LL(\boldsymbol{\eta})$*, is*

$$LL(\boldsymbol{\eta}) = \log PD = \sum_{t=s}^{N} \log(f(Y_t|D_{t-1}))$$

This suggests a method of finding, "optimal" values for the discount factors in the model. Evaluate the predictive density on a grid of discount $(\delta_1, \ldots, \delta_h, \kappa)$ values and choose the combination of values which maximizes the predictive density or likelihood.

3.9 Examples

This section will present two examples: the Constant DLM and second the Simple Regression DLM. These models are simple, yet illustrate all of the basic concepts of the DLM.

3.9.1 Constant DLM

The constant model is obtained from 3.6 by making the following simplifications,

$$\begin{aligned} \mathbf{F}_t &= 1 \text{ for all } t, \\ \boldsymbol{\theta}_t &= \mu_t, \\ \mathbf{G}_t &= 1. \end{aligned}$$

Thus the observation equation becomes:

$$Y_t = \mu_t + \nu_t, \quad \nu_t \sim N[0, V_t],$$

and the system equation is:

$$\mu_t = \mu_{t-1} + \omega_t \quad \omega_t \sim N[0, W_t]$$

with initial priors:

$$(\mu_0|D_0, \phi_0) \sim N[0, V_0C_0^*],$$
$$(\phi_0|D_0) \sim G[n_0/2, d_0/2],$$

where $\phi_0 = V_0^{-1}$.

Since there is only one element to the state vector we only need one discount factor, δ_μ and if we use variance learning (section 3.6) we need a second discount factor, κ in order to specify the system variance W_t.

3.9.2 Simple Regression

The simple regression model is obtained from 3.6 by making the following simplifications,

$$\begin{aligned}
\mathbf{F}_t' &= (1, x_t), \\
\boldsymbol{\theta}_t' &= (\alpha_t, \beta_t), \\
\mathbf{G}_t &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\end{aligned}$$

where, x_t is the independent variable observed at the same time as Y_t. That is, given the value of x_t we hope to be able to say something about the value of Y_t.

Thus the observation equation becomes:

$$Y_t = \alpha_t + \beta_t x_t + \nu_t, \quad \nu_t \sim N[0, V_t],$$

and the system equation is:

$$\begin{pmatrix} \alpha_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \omega_{1t} \\ \omega_{2t} \end{pmatrix}$$

where

$$\begin{pmatrix} \omega_{1t} \\ \omega_{2t} \end{pmatrix} \sim N[\mathbf{0}, \mathbf{W}_{t,T}]$$

with initial priors:

$$\begin{aligned} (\boldsymbol{\theta}_0 | D_0, \phi_0) &\sim N[\mathrm{m}_0, V_0 \mathbf{C}_0^*], \\ (\phi_0 | D_0) &\sim G[n_0/2, d_0/2], \end{aligned}$$

where $\phi_0 = V_0^{-1}$.

Now if we wish to perform component discounting on the simple regression model, we will need two discount factors δ_α and δ_β to apply to the constant term and the slope respectively to specify the system variance $\mathbf{W}_t$. Since we are using variance learning (Section 3.6) we will need a third discount factor, namely κ.

In chapter 4 the Constant DLM and the Simple Regression DLM will be applied to the the problem of forecasting value-weighted real returns of TSE portfolios discussed in Chapter 2.

3.10 Computer Implementation

This section briefly describes the computer implementation of the Unvariate DLM. The author programmed the dynamic linear model in Fortran 77 using NAG subroutines (Numerical Algorithms Group). We then used the dyn.load features of SPLUS 3.1 to provide a user friendly interface.

Chapter 4

Empirical Results of Applying DLM to Value–Weighted Real Returns

In this chapter we apply the methods of Bayesian forecasting using Dynamic Linear Models to the problem of forecasting value–weighted real returns, $R_{t,T}$, of TSE portfolios using dividend yields, DY_t, for different return horizons, T. Specifically, we will use the two models discussed in Section 3.9, namely the Constant DLM and the Simple Regression DLM.

4.1 Constant DLM

In this section we deal with the problem of specifying a model for the mean level of the real return series $R_{t,T}$. That is, are the returns made up of a slowly varying mean component plus observational noise with a time varying variance? Specifically, the model for return horizon T is given by:

observation equation:

$$R_{t,T} = \mu_{t,T} + \nu_{t,T}, \quad \nu_t \sim N[0, V_{t,T}],$$

and system equation:

$$\mu_{t,T} = \mu_{(t-1),T} + \omega_{t,T} \quad \omega_t \sim N[0, W_{t,T}]$$

with initial priors:

$$\begin{aligned} (\mu_{0,T}|D_0, \phi_{0,T}) &\sim N[\mathbf{m}_{0,T}, V_{0,T}C^*_{0,T}], \\ (\phi_{0,T}|D_0) &\sim G[n_{0,T}/2, d_{0,T}/2], \end{aligned}$$

where $\phi_{0,T} = V_{0,T}^{-1}$.

Notice that we have different parameter values for each return horizon T. Now, in order to specify the system variance W_{tT} we need a discount factor δ_{μ_T}. We shall also employ the methods of variance learning, thus we need a second discount factor κ. Are we justified in assuming that the precision ϕ_t follows a random walk? Figure 2.4 sheds some light on this assumption; the estimated residual variance doesn't seem to varying randomly about constant value. This graphical evidence seems to suggest that some time varying variance is appropriate. However, there is no guarantee that the random walk approach of variance learning is the "best" approach to take.

The initial priors are based on the "non-informative" priors and are discussed in the section on reference analysis (see Section 3.7).

We employ the methods of Section 3.8 to choose the "optimal" values of the discount factors. That is, we evaluate the predictive density at a grid of values for $\delta_{\mu_T} = 0.01, 0.02, \ldots, 1.00$ and $\kappa_T = 0.95, 0.96, \ldots, 1.00$ and choose the pair that maximizes the predictive density or equivalently the log likelihood, namely $\widehat{\delta}_{\mu_T}$ and $\widehat{\kappa}_T$. Table 4.1 presents the maximizing values.

Table 4.1: Maximum Likelihood estimates of δ_{μ_T} and κ_T for the constant DLM for return horizons T.

	Monthly		Quarterly		Yearly		2 Year		3 Year		4 Year	
Portfolio	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$	$\widehat{\delta}_{\alpha_T}$	$\widehat{\kappa}_T$
Composite	1.00	0.95	1.00	0.95	1.00	1.00	1.00	0.95	0.24	1.00	0.38	0.95
Metals	1.00	0.95	1.00	0.95	1.00	1.00	1.00	0.97	0.37	0.96	0.41	1.00
Gold	1.00	0.95	1.00	0.95	1.00	0.96	1.00	0.95	0.01	1.00	0.19	0.95
Oil	1.00	0.95	1.00	0.95	1.00	1.00	0.02	1.00	0.08	0.95	0.20	1.00
Paper	1.00	0.95	1.00	0.95	1.00	1.00	1.00	0.99	0.15	1.00	0.32	0.95
Consumer	1.00	0.95	1.00	0.97	1.00	1.00	0.68	0.96	0.06	0.95	0.25	1.00
Industrial	1.00	0.96	1.00	1.00	1.00	1.00	0.91	1.00	0.07	1.00	0.50	1.00
Real Estate	0.92	0.95	0.84	0.95	0.58	1.00	0.36	1.00	0.27	1.00	0.21	1.00
Transportation	1.00	0.97	1.00	0.97	1.00	1.00	0.16	1.00	0.12	1.00	0.18	0.95
Pipelines	1.00	0.95	1.00	0.95	1.00	0.96	1.00	0.95	0.22	1.00	0.35	0.95
Utilities	1.00	0.95	1.00	0.95	1.00	1.00	0.23	1.00	0.23	1.00	0.26	1.00
Communications	1.00	0.95	1.00	0.98	1.00	1.00	0.95	0.95	0.08	0.95	0.25	1.00
Merchandising	1.00	0.95	1.00	1.00	1.00	1.00	0.93	1.00	0.01	1.00	0.33	1.00
Financial	1.00	0.95	1.00	0.98	1.00	1.00	1.00	0.95	0.22	1.00	0.38	1.00
Conglomerates	1.00	0.95	1.00	1.00	1.00	0.99	0.16	1.00	0.01	0.97	0.05	1.00

Table 4.1 indicates that for most portfolios and return horizons the static model is not appropriate, that is, either the variance is changing indicated by κ being less than one or the mean is time varying indicated by δ_{μ_T} being less than one. There appears to be a shift from models with changing observational variance and static mean component to models with a highly adaptive mean component and constant observational variance as the return horizon increases. That is, the models are changing from becoming less adaptive with changing observational variance to models where the mean component is changing rapidly and the observational variance is more or less constant. This change in model character as we move from monthly to four year returns may be due to the fact that the constant model does not apply, hence the rapidly changing mean estimates.

Figure 4.1 shows plots of the real returns, $\{R_{t,T}\}$, forecasted returns and %95 forecast limits for the TSE composite portfolio using the constant DLM. The figure reflects the pattern shown in Table 4.1 for the Composite series. For example, for the monthly return horizon we have higher predictive density by not varying the level of the forecast, but we suffer a time varying forecast variance. On the other hand, for the 4 year return horizon, we maximize the predictive density by having the forecasts adapt or change very quickly in response to the changing returns, and the forecast variance is relatively constant.

The reason for discussing the Constant DLM model is to have a reference model in which to compare the forecasts generated with the Simple Regression DLM to be discussed in the following section. That is, if the dividend yields, DY_t, are to be useful in forecasting the value-weighted real returns of the TSE portfolios then the Simple Regression DLM should do better than the Constant DLM.

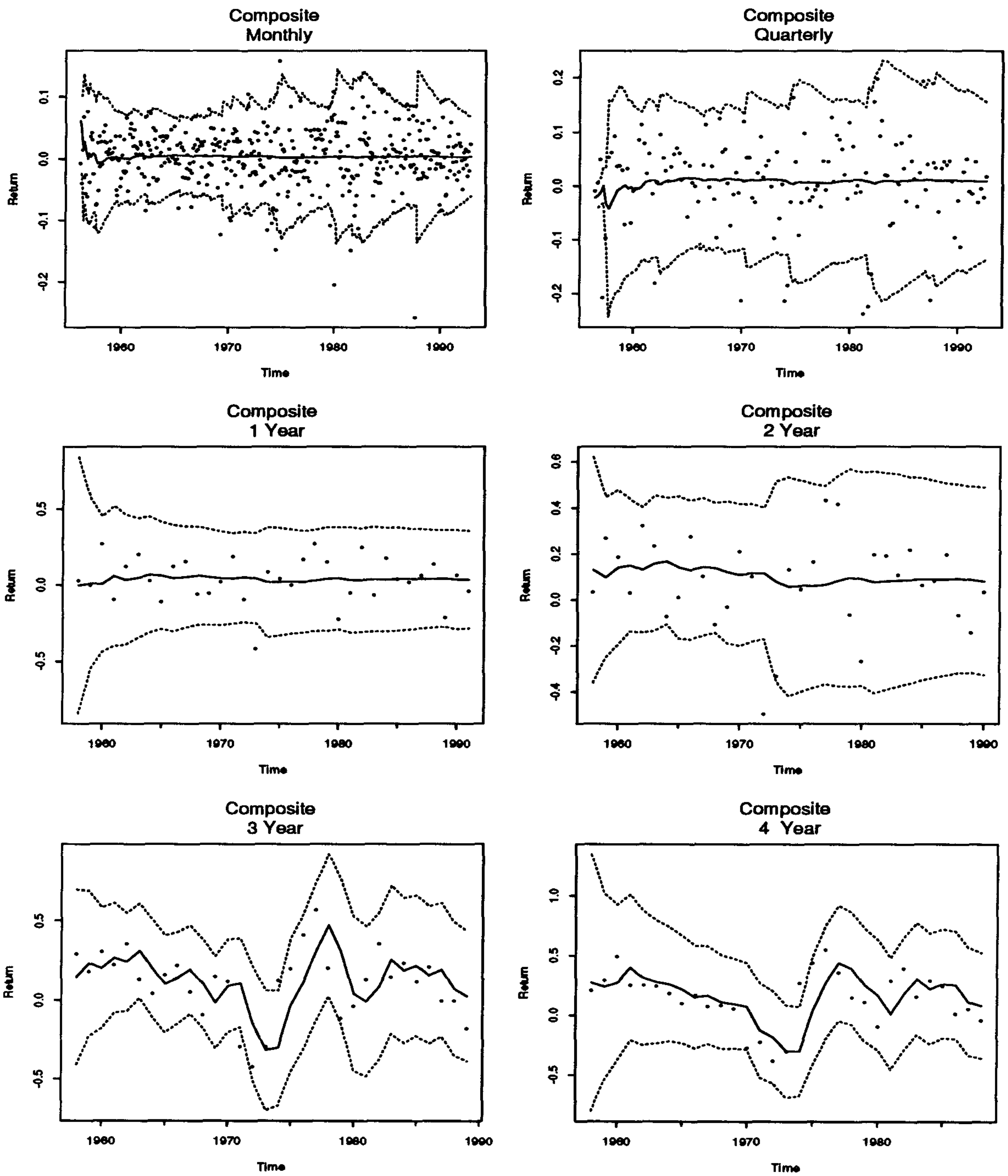

Figure 4.1: Plots of the actual returns $\{R_{t,T}\}$ (points), the forecasted returns at time t (solid lines) and 95% forecast limits (dashed lines) for each of the six return horizons for the TSE Composite portfolio for the Constant DLM.

4.2 Simple Regression DLM

This section applies the simple regression DLM (see section 3.9.2), to problem of forecasting the value-weighted real returns of the TSE portfolios, $R_{t,T}$, using dividend yields, DY_t for varying return horizons, T. The advantage of using the DLM approach over (2.1) is that the the parameters of the model can be time varying. Specifically, the Simple Regression DLM is given by:
observation equation:

$$R_{t,T} = \alpha_{t,T} + \beta_{t,T} DY_t + \nu_{t,T}, \quad \nu_t \sim N[0, V_{t,T}],$$

system equation:

$$\begin{pmatrix} \alpha_{t,T} \\ \beta_{t,T} \end{pmatrix} = \begin{pmatrix} \alpha_{(t-1),T} \\ \beta_{(t-1),T} \end{pmatrix} + \begin{pmatrix} \omega^1_{t,T} \\ \omega^2_{t,T} \end{pmatrix}$$

where

$$\begin{pmatrix} \omega^1_{t,T} \\ \omega^2_{t,T} \end{pmatrix} \sim N[\mathbf{0}, \mathbf{W}_{t,T}]$$

with initial priors:

$$\begin{aligned} (\boldsymbol{\theta}_{0,T} | D_0, \phi_{0,T}) &\sim N[\mathbf{m}_{0,T}, V_{0,T}\mathbf{C}^*_0], \\ (\phi_{0,T} | D_0) &\sim G[n_{0,T}/2, d_{0,T}/2], \end{aligned}$$

where $\phi_{0,T} = V^{-1}_{0,T}$.

As in the previous section we employ reference analysis to choose the initial priors $(\boldsymbol{\theta}_{0,T} | D_0, \phi_{0,T})$ and $(\phi_{0,T} | D_0)$. In addition, we must specify the the evolution variance matrix $\mathbf{W}_{t,T}$. To accomplish this we employ the method of component discounting. That is, we must choose discount factors δ_{α_T} for α_T and δ_{β_T} for β_T, which represent the amount of information decay in the parameter estimates in the time period $t-1$ to t. We also employ the method of variance learning, which allows the variance to be a slowly varying function of time, thus we need a third discount factor κ_T to account for the changing variance estimates. A graphical justification for assuming a random walk for the parameter estimates is discussed in the previous section (see Figure 2.4).

As for the Constant DLM we evaluate the predictive density or equivalently the log likelihood over a grid of the discount values in order to find the "optimal" or maximum likelihood values of the discount factors. Specifically, the grid consists of $\delta_{\alpha_T} = 0.01, 0.02, \ldots, 1.00$, $\delta_{\beta_T} 0.01, 0.02, \ldots, 1.00$, and $\kappa_T = 0.95, 0.96, \ldots, 1.00$. The maximum likelihood estimates of the discount factors, are given in Table 4.2 for the six return horizons, T, and for each of the fifteen portfolios of the TSE.

Table 4.2: Maximum Likelihood estimates of δ_{α_T}, δ_{β_T} and κ_T for the Simple Regression DLM for return horizons T.

	Monthly			Quarterly			Yearly		
Portfolio	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$
Composite	1.00	1.00	0.95	0.96	0.89	0.95	1.00	1.00	1.00
Metals	1.00	1.00	0.95	1.00	1.00	0.95	1.00	1.00	1.00
Gold	1.00	1.00	0.95	1.00	1.00	0.95	1.00	1.00	0.95
Oil	1.00	0.99	0.95	1.00	1.00	0.95	1.00	1.00	0.95
Paper	1.00	1.00	0.95	1.00	1.00	0.95	1.00	1.00	1.00
Consumer	1.00	1.00	0.95	0.99	1.00	0.96	1.00	1.00	1.00
Industrial	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.99	1.00
Real Estate	0.99	0.96	0.95	1.00	0.95	0.95	0.93	0.89	1.00
Transportation	1.00	1.00	0.97	1.00	1.00	0.97	1.00	1.00	0.95
Pipelines	1.00	0.99	0.95	1.00	0.97	0.95	1.00	0.93	0.97
Utilities	1.00	1.00	0.95	1.00	1.00	0.95	0.97	0.95	1.00
Communications	1.00	1.00	0.95	1.00	1.00	0.98	1.00	1.00	1.00
Merchandising	0.97	0.99	0.95	1.00	1.00	0.95	1.00	0.99	1.00
Financial	1.00	1.00	0.95	0.94	0.89	1.00	1.00	0.97	1.00
Conglomerates	1.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00
	2 Year			3 Year			4 Year		
Portfolio	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$	$\widehat{\delta_{\alpha_T}}$	$\widehat{\delta_{\beta_T}}$	$\widehat{\kappa_T}$
Composite	0.68	0.96	1.00	0.80	0.95	1.00	0.80	1.00	1.00
Metals	0.98	1.00	0.95	0.90	1.00	0.95	0.90	0.75	1.00
Gold	0.55	0.72	0.95	0.45	0.75	1.00	0.60	0.65	1.00
Oil	1.00	1.00	0.95	1.00	0.95	0.95	0.95	1.00	0.97
Paper	1.00	1.00	0.97	0.85	0.50	1.00	0.95	1.00	0.95
Consumer	0.87	0.93	1.00	0.80	0.90	1.00	0.90	0.95	1.00
Industrial	1.00	0.96	1.00	0.60	1.00	0.97	0.65	1.00	1.00
Real Estate	0.74	0.81	1.00	0.70	0.70	1.00	0.65	0.75	1.00
Transportation	0.86	0.55	1.00	0.90	0.45	0.95	0.90	0.45	0.95
Pipelines	0.77	1.00	1.00	0.85	1.00	1.00	0.80	1.00	0.95
Utilities	0.95	1.00	1.00	0.95	1.00	0.95	1.00	0.95	1.00
Communications	0.95	0.55	0.95	0.35	0.95	0.95	0.85	0.80	1.00
Merchandising	0.84	0.83	1.00	0.55	0.95	0.95	0.75	0.95	1.00
Financial	0.97	0.57	1.00	0.90	0.95	1.00	0.95	0.90	0.98
Conglomerates	0.55	0.91	1.00	0.95	0.25	0.95	0.55	0.85	1.00

It is apparent from examining Table 4.2 that the assumption of non–time varying parameters is suspect for the majority of the portfolios and return horizons. The parameters α and β for Fama and French model (2.1) are time varying for the majority of the portfolios for the two to four year return horizons as evidenced by the maximum likelihood estimates for δ_{α_T} and δ_{β_T} being less than one. However, in most instances for the two to four year return horizons the observational variance is constant as indicated by the κ_T estimate being one. For the monthly and quarterly return horizons the opposite appears to happen, that is the parameters α and β are constant and the observational variance is time varying. And finally, for the yearly returns, the model with no time varying parameters is "best" for most return horizons.

Figure 4.1 shows plots of the real returns, $\{R_{t,T}\}$, forecasted returns and %95 forecast limits for the TSE composite portfolio using the simple regression DLM. Figure 4.2 depicts a similar pattern to that of Figure 4.1, that is, for short return horizons the dominate feature is the changing forecast variance and for longer return horizons the adaption is more pronounced, indicating that dividend yields actually forecast real returns. Of course, this refers only to the Composite portfolio.

In the following two sections we compare the simple regression DLM with the constant regression DLM to determine if dividend yields contribute exhibit any forecast power. Secondly, we compare the simple regression DLM with the classical regression approach discussed in chapter 2.

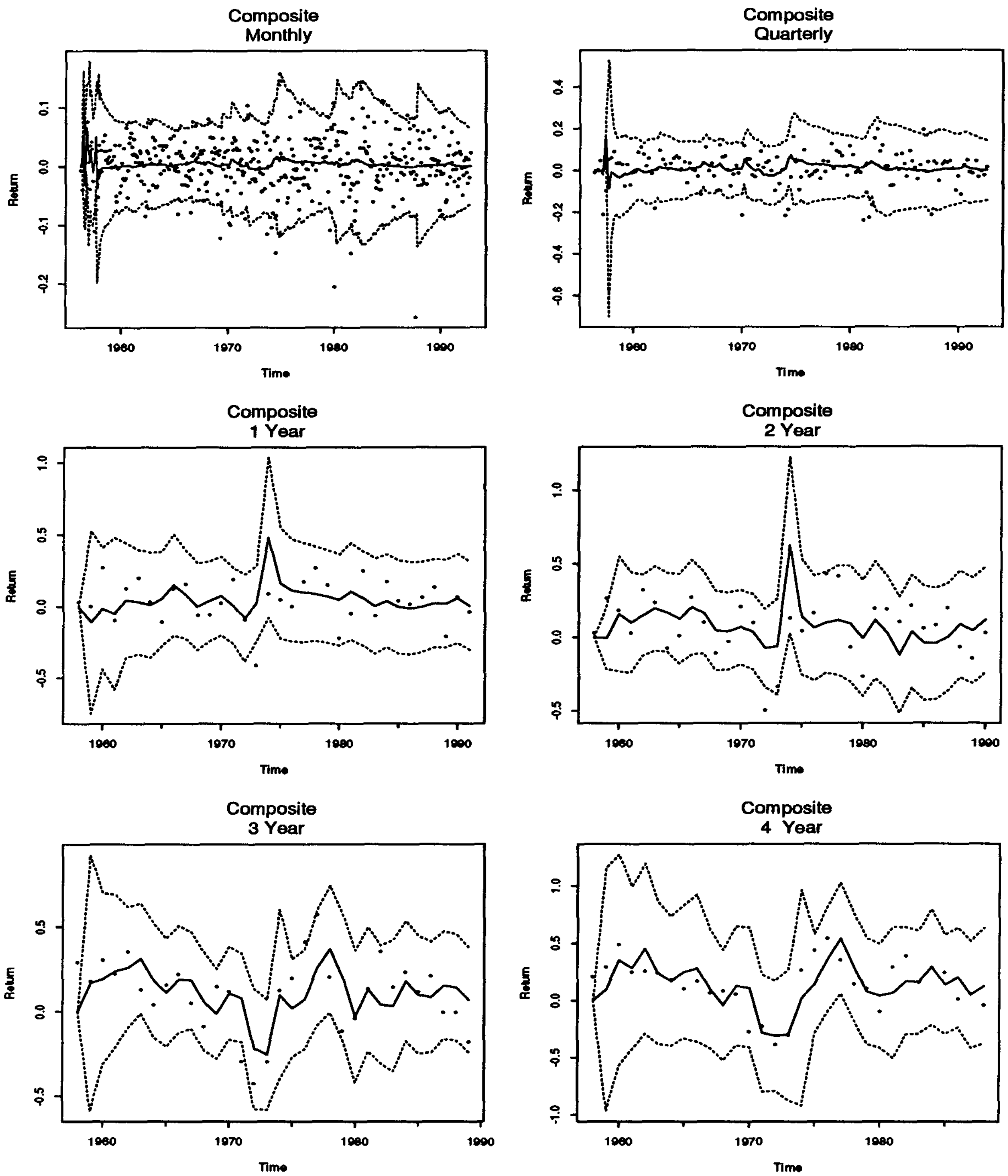

Figure 4.2: Plots of the actual returns $\{R_{t,T}\}$ (points), the forecasted returns at time t (solid lines) and 95% forecast limits (dashed lines) for each of the six return horizons for the TSE Composite portfolio for the Regression DLM.

4.3 Constant vs Regression DLM

We compare the predictive densities of the Constant DLM and the Simple Regression DLM to determine if dividend yields, DY_t, have any ability to forecast value-weighted real returns, $R_{t,T}$, of TSE portfolios. A model is better if it has a higher predictive density or higher log likelihood. We define the following difference of log likelihoods:

$$\nabla LL = LL_{Regression} - LL_{Constant}.$$

The following table gives values of ∇LL for each of the six return horizons and for each of the TSE portfolios. Note that there is no guarantee that if we add a regression variable that the Predictive Density will increase.

Table 4.3: Comparison of Constant DLM to the Regression DLM for each of the six return horizons and each of the TSE portfolios using ∇LL.

Portfolio	Monthly	Quarterly	1 Year	2 Year	3 Year	4 Year
Composite	-3.481	-4.422	-0.283	4.614	3.082	7.287
Metals and Minerals	-5.601	-5.832	-1.931	-1.612	-2.435	0.430
Gold and Silver	-4.078	-3.731	-0.940	6.249	4.930	5.662
Oil and Gas	2.306	1.833	0.715	5.918	8.507	9.377
Paper and Forest	-5.690	-4.043	-2.568	-3.024	-2.910	-4.776
Consumer	-3.787	-1.174	-0.667	7.190	1.773	6.196
Industrial	-5.287	-4.627	-2.284	1.052	2.019	5.466
Real Estate	-2.624	-3.286	-3.041	-1.031	3.075	2.542
Transportation	-5.833	-6.025	-2.299	-0.434	0.745	2.234
Pipelines	-2.377	0.715	1.247	6.754	7.669	13.680
Utilities	-6.321	-5.711	-0.444	7.784	8.659	10.050
Communications	-6.184	-4.204	-0.568	4.370	-0.308	3.272
Merchandising	-0.257	-2.570	0.199	9.747	5.064	9.814
Financial Services	-5.293	-5.712	-0.505	8.586	4.621	6.867
Conglomerates	-7.678	-2.905	0.518	2.088	-1.541	-1.626

Examining Table 4.3 we can make the following observations:

- Monthly return horizon – constant DLM outperforms regression DLM for all portfolios except for the Oil and Gas portfolio;

- Quarterly return horizon – constant DLM outperforms regression DLM for all portfolios except the Oil and Gas portfolio and the Pipelines portfolio;
- One Year return horizon – constant DLM outperforms regression DLM for all portfolios except for the Oil and Gas portfolio, the Pipelines portfolio, the Merchandising portfolio and the Conglomerates portfolio;
- Two Year return horizon – regression DLM outperforms constant DLM for all portfolios except for the Metals and Minerals portfolio, the Paper and Forest portfolio, the Real Estate and Construction portfolio, and the Transportation portfolio.
- Three Year return horizon – regression DLM outperforms constant DLM for all portfolios except for the Metals and Minerals portfolio, the Paper and Forest portfolio, and the Communications portfolio.
- Four Year return horizon – regression DLM outperforms constant DLM for all portfolios except for the Paper and Forest portfolio, and the Conglomerates portfolio.

It is interesting to note that, generally speaking, all portfolios seem to follow the pattern of increased predictability of real returns as the return horizon increases using dividend yields. The only portfolios, where this is not the case are the following: Paper and Forest, Communications, and Conglomerates. This lends support to the hypothesis of Fama and French [4], that is, as the return horizon increases the predictive ability dividend yields increases.

The above results should be interpreted with some caution, since we do not have a method of determining how large a difference in predictive density is meaningful. Also, even when there is a change in predictive density with return horizon the increase in not always monotonic in nature.

4.3.1 Classical vs DLM

In order to make a fair comparison between the classical regression approach and the regression DLM approach we would like to have a measure which behaves in the following manner. It should penalize a forecast which is far from the observed value whenever our forecast variance is small *i.e.* we have a lot of information available. It should not penalize as heavily a forecast which is far from the observed value if the forecast variance is large *i.e.* we don't have a lot of information about the next observation. So, we can summarize the aspects of the measure in the following table:

forecast error	forecast variance	measure
small	small	acceptable
small	large	acceptable
large	small	unacceptable
large	large	acceptable

With these critera in mind we define the following measure for a forecast a time t

$$E_t = \text{forecast error}/\sqrt{\text{forecast variance}}$$

which is then used as follows:

$$SDMSE = \sum_t E_t^2$$

In order to make a fair comparison we compare the out–of–sample forecasts (see Section 2.2) generated by Fama and French [4] model with the DLM approach forecasts for the same time period. We then examine the ratio of the two standardized mean squared errors as follows:

$$\text{RATIO} = MSE_{\text{DLM}}/MSE_{\text{Classical}}$$

Note that we did not carry out the comparison for the Real Estate and Construction portfolio due to the limited number of observations. It should also be stressed that for

Table 4.4: Comparison of Classical Regression to the Regression DLM for each of the six return horizons and each of the TSE portfolios using RATIO.

Portfolio	Monthly	Quarterly	1 Year	2 Year	3 Year	4 Year
Composite	0.918	0.905	0.945	1.479	1.806	0.521
Metals and Minerals	0.716	0.893	0.943	0.486	0.438	0.395
Gold and Silver	0.795	0.756	1.077	0.942	0.908	0.960
Oil and Gas	0.905	0.949	1.579	2.585	1.936	0.941
Paper and Forest	0.924	0.959	1.187	1.351	1.161	1.354
Consumer	0.863	0.910	1.000	0.046	0.042	0.006
Industrial	1.013	1.035	0.916	1.508	1.861	1.813
Transportation	0.751	2.401	0.874	1.154	1.262	1.153
Pipelines	0.900	0.897	0.868	1.046	0.313	0.481
Utilities	0.882	1.051	0.946	0.013	0.009	0.041
Communications	1.088	1.382	1.086	1.180	1.223	1.448
Merchandising	1.097	0.993	1.260	0.805	0.423	0.156
Financial Services	0.978	0.875	1.040	1.032	2.448	2.018
Conglomerates	1.116	1.320	1.192	0.043	0.045	0.008

the one to four year return horizons, the out of sample mean squared errors for the Classical approach are based on a small number of observations.

After examining Table 4.4 we make the following observations:

- Monthly Return Horizon – Regression DLM beats classical regression in 10 of 14 cases;
- Quarterly Return Horizon – Regression DLM beats classical regression in 9 of 14 cases;
- One Year Return Horizon – Regression DLM beats classical regression in 6 of 14 cases;
- Two Year Return Horizon – Regression DLM beats classical regression in 6 of 14 cases;
- Three Year Return Horizon – Regression DLM beats classical regression in 7 of 14 cases;

- Four Year Return Horizon – Regression DLM beats classical regression in 9 of 14 cases;

It is interesting to note that when the classical regression beats the Regression DLM the greatest margin of victory is 2.6, however, in the some cases the regression DLM the margin of victor is as much as 170 times better than the classical.

4.4 Conclusions

In all but two cases the Fama and French [4] results hold true in that the regression DLM with dividend yields does better than the constant DLM for all return portfolios except for the Paper and Forest, Communications and Conglomerates as measured by the change in predictive density from the regression DLM to the constant DLM.

When comparing the regression DLM approach with the classical approach discussed in chapter 2, there is not clear cut winner. However, the regression DLM approach does beat the classical case in the majority of cases, 55% in fact.

Chapter 5

Conclusions

This concluding chapter summarizes the findings of the thesis.

5.1 Fama and French

The results of Fama and French [4] which state that dividend yields show increased forecast power to predict real returns for increasing return horizons do not extend to all portfolios of the Toronto Stock Exchange. The Fama and French results only apply to the Composite portfolio and the Oil and Gas portfolio where there is forecast power for all return horizons as shown in Fama and French. However, the result of increasing forecast performance also apply to the following portfolios: Gold and Silver, Consumer Products, Industrial Products, Real Estate and Construction, Transportation, Pipelines, Utilities, Communications, and Merchandising.

However, the model suffers from some problems, most notably:

- residual auto–correlation
- dramatic decreases in out–of–sample R^2 indicating lack of stationarity
- Model 2.1 may not apply for all time periods as suggested in Fama and French [4] as evidenced by the changing residual variance results given in Section 2.4.3. It should be noted however, that Fama and French [4] mention the fact that the return variances are not constant throughout their sampling period. They present

results for various sub periods of interest. We did not pursue this option since we had only a limited amount of data available.

5.2 Dynamic Linear Model

In all but two cases the Fama and French [4] results hold true in that the regression DLM with dividend yields does better than the constant DLM for all return portfolios except for the Paper and Forest, Communications and Conglomerates as measured by the change in predictive density from the regression DLM to the constant DLM.

When comparing the regression DLM approach with the classical approach discussed in chapter 2, there is not clear cut winner. However, the regression DLM approach does beat the classical case in the majority of cases 55%.

Bibliography

[1] Ameen, J.R.M. and Harrison, P.J. (1985), "Normal Discount Bayesian Models", *Bayesian Statistics 2*, eds Bernardo, J.M., De Groot, M.H.,, Lindley, D.V. and Smith, A.F.M., North-Holland,Amsterdam, 271-298.

[2] Box,G.E.P. and Jenkins,G.M., (1976), *Time Series Analysis: Forecasting and Control*, Rev. Ed., Holden–Day, Oakland

[3] Draper, N.R., and Smith, H. (1981), *Applied Regression Analysis*, 2d ed., Wiley, New York.

[4] Fama, E.F. and French, K.R. (1988), "Dividend Yields and Expected Returns", *Journal of Financial Economics*, **22**, 3-25.

[5] Fama, E.F. and Schwert, R.F. (1977), "Asset Returns and Inflation", *Journal of Financial Economics*, **5**, 115-146.

[6] French,K.R., Schwert,G.W. and Stambaugh,R.F. (1987), "Expected Stock Returns and Volatility", *Journal of Financial Economics*, **19**, 3-29.

[7] Hansen, L.R. and Hodrick, R.J. (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis", *Journal of the Political Economy*, **88**, no 5, 829-853.

[8] Keim, D.B. and Stambaugh, R.F. (1986), "Predicting Returns in the Stock and Bond Markets", *Journal of Financial Economics*, **17**, 357-390.

[9] Pole, A. and West,M. (1989), "Reference Analysis of the Dynamic Linear Model", *Journal of Time Series Analysis*, **10**,no 2,131-147.

[10] Priestley, M.B. (1981), "Spectral Analysis and Time Series", Academic Press, Toronto.

[11] Traynor, P., ed *The Toronto Stock Exchange Review*, The Toronto Stock Exchange.

[12] Weisberg,S (1985), *Applied Linear Regression*, 2d ed., Wiley, New York.

[13] West, M. and Harrison, P.J. (1989), *Bayesian Forecasting and Dynamic Models*, Springer-Verlag, New York.

[14] West, M. (1986), "Bayesian Model Monitoring", *Journal of the Royal Statistical Society Series B* , **48**, no 1,70-78.

[15] West, M. and Harrison, P.J. (1986), "Monitoring and Adaptation in Bayesian Forecasting Models", *Journal of the American Statistical Association*, **81**, no 395,741-750.

Appendix 1

1 Index Formula and Rules

Each of the Toronto Stock Exchange indices measure the current aggregate market value (i.e. number of presently outstanding shares × current price) of the stocks included in the index as a proportion of an average base aggregate market value (number of base outstanding shares × average base price ± changes proportional to changes made in the current aggregate market value figure) for such stocks. The starting level of the base value has been set equal to 1000. Expressed more briefly this is:

$$INDEX = \frac{Current\ aggregate\ market\ value}{Adjusted\ average\ base\ aggregate\ market\ value} \times 1000$$

Essentially, there are two stages in the production of indices: (1) establishment of an initial base and initial calculation of the indices; and (2) subsequent calculation of the indices taking into account recurring shifts of the market. Following is a detailed description of how The Toronto Stock Exchange indices are produced.

The following formula is the basis for initial calculation of each of the indices of The Toronto Stock Exchange:

$$INDEX = \frac{(P_A \times Q_A) + (P_B \times Q_B) + \ldots + (P_N \times Q_N)}{(\overline{P}_{A_B} \times Q_{A_B}) + (\overline{P}_{B_B} \times Q_{B_B}) + \ldots + (\overline{P}_{N_B} \times Q_{N_B})} \times 1000$$

A,B, ...N: the various stocks in the index portfolio.

P_A,P_B, ... P_N: the current board–lot market prices of each stock in the index.

Q_A,Q_B, ... Q_N: the numbers of currently outstanding shares of each stock in the index less any individual and/or related control blocks of 20% or more.

$\overline{P}_{A_B}$,$\overline{P}_{B_B}$, ... $\overline{P}_{N_B}$: the trade weighted average board–lot prices of each stock in the index during the base period

Q_{A_B},Q_{B_B}, ... Q_{N_B}: the number of shares of each stock in the index outstanding in the base period less any individual and/or related control blocks of 20% or less.

The base period is 1975. Calculation of the 1975 average base aggregate market value i.e.

$$(\overline{P}_{A_B} \times Q_{A_B}) + (\overline{P}_{B_B} \times Q_{B_B}) + \ldots + (\overline{P}_{N_B} \times Q_{N_B})$$

for each index was accomplished by multiplying the trade-weighted average board–lot price for each stock for the 1975 base period by the number of shares (share weight) of each stock outstanding at the beginning of the base period i.e. January 1, 1975 less any individual and/or control blocks of 20% or more. The current aggregate market value is determined using closing prices for each period for which the index is calculated multiplied by the number of shares then outstanding, less any individual and/or related control blocks of 20% or more, as at that period.

As an example of these calculations, assume there are only two stocks in a hypothetical index. The problem is to calculate the level of the index as of January 31, 1975.

Company 1

The current price (January 31, 1975) is \$10 and the number of shares currently outstanding is 18,000. The average base aggregate market value in 1975 is \$162,000.

Company 2 The current price (January 31, 1975) is \$25 and the number of shares currently outstanding is 30,000. The average base aggregate market value in 1975 is \$690,000.

Computation of the index would be as follows:

$$INDEX = \frac{10 \times 18,000) + (25 \times 30,000)}{162,000 + 690,000} \times 1000$$

$$INDEX = \frac{930,000}{852,000} \times 1000$$

$$INDEX = 1091.55$$

ADJUSTMENT TO INDEX

To calculate the indices subsequent to the establishment of the average base aggregate market value, recurring capital changes must be taken into account. **Adjustments to the indices resulting from these changes**

must normally be introduced without altering the level of the index(see Bankruptcy Rule (7) for exception). In other words, continuity of the index must be preserved. To accomplish this, certain procedures are followed. These vary according to whether the adjustments result from: (1) the issuance of additional shares of a stock in the indices; or the addition to, withdrawal from, or substitution of stocks in the indices; (2) stock rights; (3) stock dividends and stock splits; (4) a liquidation of the company; (5) an asset spin-off; (6) takeover bid, amalgamation or merger; (7) a bankruptcy; or (8) a control block adjustment.

1.1 Addition or Withdrawal of Shares or Changes in Number of Stocks

Two steps are necessary to make adjustments for additions or withdrawals of shares to or from the index calculations:

(1) **Updating the current aggregate market value of the index.** If additional shares of an index stock are issued, the current aggregate market value of the stocks in that index will be accordingly higher. Likewise, if a new stock is added to the index, or if a stock is removed, the current aggregate market value of that stock will be added to, or subtracted from the current aggregate market value of the other stocks in that index.

(2) **Adjusting the average base aggregate market value** of the index proportional to the change in the current aggregate market value so that the index level will remain the same.

The first step, therefore, towards making an adjustment is to calculate the new current aggregate market value as indicated in (1) above.

The second step is to calculate the new average base aggregate market value. Expressed as a formula the second step would be as follows:

Let the old average base aggregate market value = A. Let the un–adjusted current aggregate market value = C. Let the current aggregate market value of the capital to be added or withdrawn = D. The current adjusted aggregate market value will equal $C \pm D$.

Therefore, to establish a new average base aggregate market value (B) for an index that formula is:

$$B = A \times \frac{(C \pm D)}{C}$$

To calculate the index on the new base, the formula for the hypothetical example given above would be:

$$INDEX = \frac{(C \pm D)}{B} \times 1000$$

Continuing the example above, assume that Company 1 issued 2,000 new shares. This required an addition of \$20,000 (\$10 $\times$ 2,000) to the aggregate market value of the stocks in the index and therefore the new current aggregate market value resulting from the change is: 930,000 + 20,000 = 950,000.

The average base aggregate market value of the index also has to be changed proportionately.

Here the formula $B = A \times \frac{(C \pm D)}{C}$ is used.

$$B = 852,000 \times \frac{(930,000 + 20,000)}{930,00}$$

$$B = 870,323$$

The index level remains unchanged as shown below:

$$INDEX = \frac{950,000}{870,323} \times 1000$$

$$INDEX = 1091.55$$

1.2 Stock Rights

The day the stock sells ex–rights, the additional shares resulting from the rights are included in the calculations to establish the current aggregate market value of the indices. The average base aggregate market value, however, is adjusted by taking into account both the market price and the subscription price because on ex–rights day the current market price, and accordingly aggregate market value, discounts the rights.

The formula to calculate the new base aggregate market value following subscription to stock rights would be:

$$B = S \times \frac{C + C}{C + D - S}$$

where S = the total capital subscribed for the newly issued shares.

A concrete example of how a stock rights issue is incorporated into the index is the December 5, 1975 Bank of Nova Scotia offer. The Bank, with an outstanding capital of 18,562,500 shares, offered the shareholders of record at the close of business on December 5, 1975 rights to buy one new share at $36 per share of each 9 shares held. As a result, 2,062,500 new shares were issued. Ex–rights date was December 3, 1975 and from that date additional capitalization for the Bank of Nova Scotia used in the bank index was 2,062,500 shares times the current price (theoretically, at this opening on the "ex" date, $\$41 frac38$ adjusted for the value of the right) amounting to $85,335,938. Actual subscription price was 2,062,500 shares times $36, amounting to $74,250,000. Calculations for the proportionately adjusting the base were as follows:

Bank Index

Un–adjusted current aggregate market value:$_*$ $4,193,109,375 Un–adjusted base aggregate market value:$_*$ $1,337,840,000 New current aggregate market value after allowing for rights: (4,194,109,375 + 85,335,938) = $4,278,445,313

New base aggregate market value after allowing for rights offering:

$$1,337,840,000 \times \frac{4,278,445,313}{4,278,445,313 - 74,250,000} = \$1,361,467,499$$

$_*$ As at the close on the day prior to the ex–date. Adjustments are made after the close and before the market opens the following day. Bank of Nova Scotia closed at $42 on December 2, 1975.

1.3 Stock Dividends, Splits, and Consolidations

On the ex–dividend day the outstanding share total is increased by the number of shares issued in the form of dividends. Theoretically, the price of the stock should drop by the extent of the worth of the dividend. The current aggregate market value, therefore, will not change. Hence the base figure is not adjusted. Similarly, in the case of share splits, the increased number of shares times the lower price should equal the old number of shares times the higher price. Thus, the current aggregate market value is theoretically unchanged, and the base figure is not adjusted. The same reasoning holds in the case of stock consolidations, except that the higher price time the smaller number of shares leaves the current aggregate market value unchanged.

1.4 Liquidation of A Company

Effective January, 1979, where a capital distribution is announced as being a liquidation of a company whose stock is included in the index, that stock will be removed from the index effective the ex–distribution date.

1.5 Asset Spin–off

Effective January, 1979, adjustments necessary to leave the level of an index unchanged when a stock in that index has its per share value decreased through an asset spin–off are made at the opening of the ex–distribution day or as soon thereafter as the value of the asset being spun–off is known by the Exchange staff. Thus the staff may have to recalculate index values if a stock trades "ex–asset spin–off" without the index being stabilized.

1.6 Takeover Bid, Amalgamation or Merger

Effective January, 1979, changes in share weight or control blocks resulting from takeover bids, amalgamations or mergers are incorporated into the index as soon as is administratively possible after the fact. This procedure replaces the former procedure of incorporating such changes at the next quarterly update made just after the end of the calendar quarter to which they relate.

1.7 Bankruptcy of Stock in Index System

If and when any company, whose stock is included within the TSE "300" indices, has made an assignment in bankruptcy or been placed in receivership, its stock will be removed as soon as possible at the lowest possible price per share (one–half cent under the present computer programmes) rather than at the last board–lot price before trading was suspended. If, as, and when the company recovers in any form, it will only be eligible to be included in the index system again after fully complying with and meeting all criteria; that is, after qualifying in the normal fashion.

1.8 Control Blocks

(a) All known individual and related control blocks equal to 20% or more of the share capital of any stock included in the indices is removed in order to reflect, as nearly as may be practical, the market float or stock normally available to portfolio investors.

(b) If at any time **more than 90% of the outstanding shares** which are included in the TSE 300 index is held by a controlling group; as defined by the methods of computing control group holdings for index weighting purposes, or if the shares in public hands of the same class are so reduced that the value calculated by multiplying the most recent share price by the number of shares held by parties other than the control group is insufficient to meet the market capitalization criterion for admission to the index, then each such class of equity security shall be removed from the index as soon as is conveniently practicable.

(c) If an individual control block of 20% or more, or a related group of control blocks which in aggregate total 20% or more of the relevant shares outstanding, are initially removed from the total of such shares then outstanding for purposes of computing the share weight of the stock in the index portfolio, and (1) the holder or holders of such stock subsequently sell stock from their position to reduce the amount of such stock holding(s) below 20%, then the holding(s) will be added back to the float at the first practical time subsequent to such sale; (2) if the 20% or more block(s) subsequently falls below 20% as a result of an increase or increases in the total of such share capital outstanding, then such block(s) will not be added back to the share weight until such time as the holding falls or is reduced to 15% or less and as soon thereafter as is practical for it to be added back.

1.9 Frequency of Adjusting the Index

Stock rights, stock dividends, splits, consolidations, and liquidations are reflected in the calculations of the indices immediately as they become affective, i.e. on the "ex" date. Asset spin–offs are reflected effective the "ex" date or as soon thereafter as the value of the asset being spun-off is known by the Exchange staff. Takeovers, amalgamations and mergers are reflected as soon as possible after the fact. Bankruptcy and receivership situations are reflected as soon as possible after they are announced. Any changes resulting from the annual post–year–end revision as noted in the section entitled "Stock Eligibility Criteria" are made at the end of the first calendar quarter. Other changes (such as those related to control blocks or to addition or withdrawal of shares) are usually made on a quarterly basis. Additions or deletions of stocks are usually made on a quarterly basis but may be necessary at other times due to delistings caused by takeovers, amalgamations, or mergers or to normal delistings.

2 Dividends, Current Indicated Annual Yields, and Dividends Adjusted to Index

2.1 General

Each time an index value or price is computed, a "current indicated annual dividend yield" (in percent terms) and a corresponding "dividends adjusted to index amount" (in dollars and cents) may also be produced for that index. The Exchange has computed and included both of these dividend values in this book for each of the historical index values contained in the tables whether on a closing monthly, weekly, or daily basis.

2.2 Compilation Assumptions

2.2.1 Historical Series (1956 through 1976)

In general, the historical dividend per share series for each stock in the index system have been compiled on the following basis:

- the original dividend payments per share against payment date of the dividend;
- if it could be determined that the payment in question formed part of a regular annual dividend policy of a company (whether paid quarterly, semi-annually, or annually) then the payment was multiplied by the

appropriate factor e.g. 4 in the case of a quarterly amount to forecast the current indicated annual rate at that point in time. In addition, any extra paid within the last twelve months was then added to the regular annual rate to become the current annual rate for that stock at that point in time.

- In cases where the regular periodic amount was not level, as in the case where a company has a policy of paying $1 per share per annum payable 20 cents in the first quarter, 30 cents in the second quarter, 20 cents in the third quarter and 30 cents in the fourth quarter, no forecasting from the periodic amount was carried out; instead the annual amount was used directly in making payout calculations.

- Irregular payments in terms of either time or amount were treated as such and included on payment date and were carried forward in yield calculations for twelve months, after which time they were removed from the calculations.

2.2.2 Current Series (1977 forward)

Updating of the dividend per share data takes place daily as new dividend reports are received by the Exchange, subject to computer cut–off requirements;

In general, the same compilation rules as were used for the historical series are used for the daily up–dates, except that updating is on an "as reported" basis as noted above rather than on the payment date.

2.2.3 Translation of Foreign Currency Amounts

If dividends per share are stated in a foreign currency, then Exchange staff translate the amounts into Canadian currency on the following basis:

(a) Historical (January 1956 through December 1976):
Dividends originally expressed in U.S. funds were converted to Canadian funds using monthly average noon spot rate applicable to the month during which the dividend was paid. The annual rate for companies paying dividends in U.S. funds was the sum of all dividends (including extra dividends) paid within the past 12 months, as converted by using the monthly average noon spot rate applicable to the month during which each dividend was paid.

Therefore the annual rate in Canadian funds may included four different exchange rates.

(b) Current (January 1977 to present):
Dividends originally expressed in U.S. funds are converted to Canadian funds using the closing foreign exchange rate as reported in the current Globe and Mail Report on Business for the day prior to the date of receipt of the dividend declaration.

3 Calculation Procedure and Product

The computation procedure to produce a "current indicated dividend yield" on an index and the corresponding "dividends adjusted to index" amount in dollars and cents is straightforward and analogous to that on an individual stock. The current indicated annual dividend rate for each stock within an index is multiplied by the share weight of that stock within the index as at that computation date. The resulting dividend payout figures for the stocks within the index are then summed. This aggregate pool of dividend dollars is then divided by the aggregate current market value (numerator of the index) as at that same point in time for the stocks within that index. The resultant figure is the current indicated annual yield expressed as a decimal; it may be multiplied by 100 to express the amount as a percent.

This current indicated annual yield on an index, stated as a decimal may then be multiplied by the index for that same computation date to produce the "Dividends Adjusted to Index" amount in dollars and cents. This amount may then be used in market valuation formulae along with other assumptions to forecast the expected level of the index based on these assumptions.

Appendix 2

TSE 300 Composite Index

RELATIVE WEIGHTS

The following list (containing 300 stocks, fourteen Group Indices, and forty-three Sub-Group Indices) provides the weight which individual stocks, Group Indices, and Sub-Group Indices bear on the Toronto Stock Exchange 300 Composite Index. Computations are as of the close of December 31, 1991. Percentages in brackets indicate available float on which relative weights on Composite Index are calculated. Numbers in brackets after each Index name indicate the number of stocks within that Index. TOTALS MAY NOT ADD DUE TO ROUNDING. All stocks in the TSE 300 are common stock or inter-convertible pairs of common stocks (shown as A, B or A,B,C) unless otherwise noted.

Stock Symbol	Company Name	Relative Weight on Composite %
1.0 METALS & MINERALS (16)		7.70
1.1 Integrated Mines (6)		6.87
AL	Alcan Aluminum	3.00
BMS	Brunswick Mining & Smelt (15%)	0.02
CLT	Cominco (56%)	0.55
HBM.S	Hudson Bay Min. & Smelt S (52%)	0.02
N	Inco	2.15
NOR	Noranda (55 %)	1.13
1.2 Metal Mines (8)		0.33
CCH	Campbell Resources (51%)	0.01
COR	Cominco Resources (35%)	0.02
KER	Kerr Addison (51%)	0.08
MLM	Metal Mining (41%)	0.11
MVA	Minnova (50%)	0.07
NGX	Northgate Exploration (75%)	0.01
PMC	Princetown Mining	0.02
WMI	Estmin Resources (26%)	0.02
1.4 Non-Base Metal Mining (2)		0.49
POT	Potash Corp. of Saskatchewan (62%)	0.29
ROM	Rio Algom (49%)	0.20

2.0 GOLD & SILVER (28)		7.39
2.1 Gold & Silver Mines (25)		**7.30**
ABX	American Barrick (79%)	2.04
AGE	Agnico-Eagle	0.08
AUR	Aur Resources (79%)	0.05
BGO	Berma Gold	0.03
BWR	Breakwater Resources (73%)	0.01
CBJ	Cambior (79%)	0.13
DML.A	Dickenson Mines CL A (39%)	0.01
ECO	Echo Bay Mines	0.52
EN	Euro-Nevada (80%)	0.11
FN	Franco Nevada (77%)	0.12
GKR	Golden Knight Resources (56%)	0.04
GLC	Galactic Resources (66%)	0.01
GXL	Granges Inc (49%)	0.01
HEM	Hemlo (45%)	0.28
ICR	Int'l Corona Corp (70%)	0.18
LAC	LAC Minerals	0.70
MVG	Minven Gold (55%)	0.00
PDG	Placer Dome	1.72
PGU	Pegasus Gold	0.23
RAY	Rayrock Yellowknife (75%)	0.02
RYO	Royal Oak Mines	0.05
TEK.B	Teck Corporation CL B	0.80
TVX	TVX Gold (24%)	0.07
VOY	Viceroy Resources	0.05
WFR	Wharf Resources (49%)	0.03
2.2 Precious Metal Funds (3)		**0.10**
BPT.A	BGR Precious Metals CL A	0.03
CEF.A	Central Fund of Canada CL A	0.04
G	Goldcorp	0.03
3.0 OIL & GAS (38)		6.71
3.1 Integrated Oils (4)		**2.04**
CCT	Canadian Turbo	0.02
IMO	Imperial Oil (30%)	1.36
SHC	Shell Canada A (22%)	0.54
TPN	Total Petroleum N.A. (47%)	0.12

3.2 Oil & Gas Producers (34)		4.66
AEC	Alberta Energy (63%)	0.31
AXL	Anderson Exploration (33%)	0.04
BPC	B.P. Canada (43%)	0.15
BVI	Bow Valley Industries (67%)	0.27
CBE	Cabre Exploration	0.06
CEX.B	Conwest Exploration	0.09
CGH	Computalog Ltd	0.01
CHA	Chauvco Resources (38%)	0.10
CID	Chieftain International (45%)	0.04
CIR	Cimarron Petroleum	0.05
CNQ	Canadian Natural Resources (81%)	0.08
COH	Coho Resources (63%)	0.01
CSW	Canadian Southern Petroleum	0.03
CXY	Canadian Occidental (52%)	0.54
ECR	Encor Inc	0.02
GOU	Gulf Canada Resources (26%)	0.14
LMO	Lasmo Canada (66%)	0.06
MHI	Morgan Hydrocarbons	0.07
MKC	Mark Resources (36%)	0.04
MRP	Morrison Petroleums (77%)	0.08
NCN	Norcen Energy (63%)	0.24
NCN.A	Norcen Energy CL A (70%)	0.29
NCO	North Canadian Oils (48%)	0.11
NMC	Numac Oil & Gas (53%)	0.05
NWS	Nowsco Well Service	0.10
PCP	PanCanadian Petroleum (13%)	0.23
PNN	Pinnacle Resources	0.05
POC	Poco Petroleums (65%)	0.13
RES	Renaissance Energy	0.49
RGO	Ranger Oil	0.49
RRL	Ranchmen's Resources (44%)	0.02
SKO	Saskatchewan Oil and Gas (76%)	0.18
SRL	Sceptre Resources (55%)	0.05
ULP	Ulster Petroleums	0.05

4.0 PAPER & FOREST PRODUCTS (18)		2.23
4.1 Paper & Forest Products (18)		**2.23**
A	Abitibi-Price (18%)	0.10
CAS	Cascades (41%)	0.08
CFI	Crestbrook Forest (46%)	0.04
CFP	Canfor (55%)	0.21
DHC.B	Donohue CL B (46%)	0.10
DOM.B	Doman Industries CL B (74%) 0.03	
DTC	Domtar (57%)	0.25
FCC.A	Fletcher Challenge Canada CL A (28%)	0.16
IFP.A	International Forest Products CL A (76%)	0.07
MB	MacMillan Bloedel (51%)	0.61
NF	Noranda Forest (18%)	0.12
PFP	Canadian Pacific Forest (20%)	0.12
RPP	Repap Enterprises (62%)	0.04
SPL	Scott Paper (50%)	0.08
TBC.A	Tembec CL A (59%)	0.05
WFT	West Fraser Timber (58%)	0.08
WGL	Westar Group (70%)	0.04
WLW	Weldwood of Canada (15%)	0.03
5.0 CONSUMER & PRODUCTS (25)		9.47
5.1 Food Processing (5)		**0.69**
BCS.A	B.C. Sugar Refinery CL A,B	0.12
CFL	Corporate Foods (33%)	0.07
CMG	Canada Malting (60%)	0.10
FPL	FPI Ltd	0.06
MFL	Maple Leaf Foods (44%)	0.35
5.2 Tabacco (2)		**1.56**
IMS	Imasco (60%)	1.49
ROC	Rothmans (29%)	0.07
5.3 Distilleries (2)		**4.56**
CDL.A	Corby Distilleries CL A (49%)	0.08
VO	Seagram (62%)	4.48
5.4 Breweries & Beverages (4)		**1.65**
KOC	Coca–Cola Beverages (51%)	0.09
LBT	Labatt, John (62%)	0.71
MOL.A	Molson CL A	0.71
MOL.B	Molson CL B (50%)	0.13

5.5 Household Goods (4)		0.28
COC	Camco Inc (29%)	0.03
CRW	Cinram Ltd (62%)	0.05
DTX	Dominion Textile	0.15
NMA.A	Noma industries CL A (71%)	0.05
5.6 Autos & Parts (4)		0.50
FMC	Ford Motor of Canada (6%)	0.04
HAY	Hayes-Dana (43%)	0.05
MG.A	Magna International CL A	0.36
UAP.A	UAP Inc CL A (68%) 0.05	
5.7 Packaging Products (4)		0.23
CCQ.B	CCL Industries CL B (73%)	0.12
CGC	Consumers Packaging (25%)	0.01
INO	International Innnopac (63%)	0.01
LMP.A	Lawson Mardon CL A	0.09
6.0 INDUSTRIAL & PRODUCTS (38)		11.05
6.1 Steel (5)		1.42
CEI	Co-Steel (84%)	0.17
DFS	Dofasco	0.75
ISP	Ipsco	0.19
IVA.A	Ivaco CL A	0.03
STE.A	Stelco CL A,B	0.29
6.2 Metal Fabricators (5)		0.32
DRE.A	Dreco Energy Services CI A	0.04
DRL	Derlan Industries	0.05
HLY	Haley Industries (20%)	0.01
SHL.A	Shaw Industries (73%)	0.09
SNU.A	SNC Group CL A	0.13
6.3 Machinery (1)		0.07
UDI	United Dominion Industries (44%)	0.07
6.4 Transportation Equipment (3)		1.00
BBD.A	Bombardier CL A (20%)	0.07
BBD.B	Bombardier CL B	0.88
HSC	Hawker Siddeley Canada (41%)	0.05
6.5 Electrical/Electronic (10)		4.29
AAZ	Archer Communications (63%)	0.01
CAE	CAE Industries	0.41
CMW	Canadian Marconi (48%)	0.10
GAN	Gandalf Technologies	0.02
KGL.B	Kaufel Group CL B	0.03

MLT	Mitel (49%)	0.02
NNC	Newbridge Networks (58%)	0.10
NTL	Northern Telecom (47%)	3.43
SHK	SHL Systemhouse (33%)	0.06
SPZ	Spar Aerospace	0.11
6.6 Cement/Concrete Products (3)		0.26
GYP	CGC Inc (24%)	0.03
LCI.PR.E	Lafarge Canada E	0.12
ST.A	St. Lawrence Cement CL A (63%)	0.11
6.7 Chemicals (6)		2.03
ANG	Alberta Natural Gas (50%)	0.08
CCL	Celanese Canada (44%)	0.14
DUP.A	Du Pont Canada CL A (23%)	0.17
NVA	Nova Corp	1.36
OIL	Ocelot Industries	0.12
SE	Sherritt Gordon	0.15
6.9 Business Services (5)		1.65
COS	Corel Systems (53%)	0.07
CSN	Cognos Inc. (61%)	0.06
IIT.A	Intera Information CL A (72%)	0.03
MCL	Moore Corp	1.40
XXC.B	Xerox Canada CL B	0.11
7.0 REAL ESTATE & CONSTRUCTION (11)		0.91
7.1 Developers & Contractors (4)		0.18
COT	Coscan (43%)	0.04
CXA.A	Consolidated HCI Holdings CL A (42%)	0.01
ITW	Intrawest Development	0.10
RLG	Royal Lepage (42%)	0.03
7.2 Property Managers (7)		0.74
BCD	Bramalea (30%)	0.08
CBG	Cambridge Shopping Centres (75%)	0.25
CDN	Carena Development (33%)	0.10
MKP	Markborough Properties (19%)	0.06
RPC	Revenue Properties (29%)	0.02
TZC.A	Trizec CL A (28%)	0.15
TZC.B	Trizec CL B (21%)	0.08

8.0 TRANSPORTATION & ENVIRONMENTAL SERVICES (8)		2.12
8.1 Transportation & Environmental Services (8)		2.12
AC	Air Canada	0.34
ALC	Algoma Central (59%)	0.01
GHL	Greyhound Lines of Canada (31%)	0.05
LDM.A	Laidlaw Inc CI A (53%)	0.15
LDM.B	Laidlaw Ind CI B	1.20
PEN	Philip Environmental (64%)	0.09
PWA	PWA Corporation	0.16
TMA	Trimac (70%)	0.12
9.0 PIPELINES (4)		2.06
9.1 Oil Pipelines (2)		0.29
IPL	Interprovincial Pipe Line (36%)	0.27
TMP	Transmountain Pipelines (33%)	0.03
9.2 Gas Pipelines (2)		1.77
TRP	TransCanada Pipelines (78%)	1.34
W	Westcoast Energy (63%)	0.43
10.0 UTILITIES (17)		13.72
10.1 Gas Utilities (3)		0.54
BCG	BC Gas Inc	0.38
GWT	GW Utilities (11%)	0.04
UEI	Union Energy (39%)	0.12
10.2 Electrical Utilities (5)		1.77
ACO.X	Atco CL I (80%)	0.14
CU	Canadian Utilities CL A (60%)	0.26
CU.X	Canadian Utilities CL B (35%)	0.10
FTS	Fortis Inc	0.14
TAU	TransAlta Utilities	1.12
10.3 Telephone Utilities (9)		11.41
AGT	Telus Corp	1.26
B	BCE Inc	8.51
BCT	British Columbia Telephone (50%)	0.70
BCX	BCE Mobile Communications (26%)	0.30
BRR	Bruncor (67%)	0.14
MTT	Maritime Telephone & Telegraph (66%)	0.22
NEL	Newtel Enterprises (44%)	0.08
QT	Quebec Telephone (49%)	0.09
TGO	Teleglobe Inc (49%)	0.12

11.0 COMMUNICATIONS & MEDIA (19)		4.74
11.1 Broadcasting (5)		0.31
BNB	Baton Broadcasting (47%)	0.05
CF	CFCF (71%)	0.03
CHM.B	CHUM CL B	0.12
TM.B	Tele-Metropole CL B (73%)	0.03
WIC.B	WIC Western CL B (67%)	0.08
11.2 Cable & Entertainment (6)		0.79
CPX	Cineplex Odeon (43%)	0.04
RCI.A	Rogers Communications CL A (9%)	0.04
RCI.B	Rogers Communications CL B (35%)	0.52
SAT	Canadian Satellite (26%)	0.02
SCL.B	Shaw Cablesystems CL B (78%)	0.10
VDO	Le Groupe Videotron (29%)	0.06
11.3 Publishing & Printing (8)		3.64
HLG	Hollinger (32%)	0.12
MHP	Maclean Hunter (80%)	1.01
QBR.A	Quebeccor CL A (44%)	0.09
QBR.B	Quebeccor CL B	0.14
STM	Southam (76%)	0.45
TOC	Thompson Corporation (30%)	1.57
TSP	Toronto Sun Publishing (36%)	0.07
TS.B	Torstar CL B (43%)	0.20
12.0 MERCHANDISING (31)		5.13
12.1 Wholesale Distributors (5)		0.37
ACK	Acklands (76%)	0.03
EML	Emco Ltd (51%)	0.02
FTT	Finning Ltd	0.26
UWB	United Westburne (28%)	0.04
WJX.A	Wajax CL A,B (46%)	0.02
12.2 Food Stores (6)		1.46
EMP.A	Empire CL A	0.12
L	Loblaw Companies (24%)	0.19
OSH.A	Oshawa Group CL A	0.47
PGI	Provigo (48%)	0.23
UGO.B	Unigesco CL B	0.03
WN	Weston, George (42%)	0.42
12.3 Department Stores (3)		0.49
HBC	Hudson's Bay Co (27%)	0.26
SCC	Sears Canada (37%)	0.20
WDS	Woodward's Ltd (49%)	0.03

12.4 Clothing Stores (3)		0.17
DLX.A	Dylex CL A (76%)	0.10
GFG.A	Grafton Group CL A	0.00
RET.A	Reitman's CL A	0.07
12.5 Specialty Stores (5)		1.38
CTR.A	Canadian Tire CL A	1.12
GDS.A	Gendis CL A,B (43%)	0.09
PCJ.A	Peoples Jewellers CL A	0.01
PJC.A	Jean Coutu Group CL A	0.14
TCG	TCG International (41%)	
12.6 Health & Hospitality (9)		1.27
BCH	Biochem (IAF) (61%)	0.30
CAO.A	Cara Operations CL A (63%)	0.12
CAO	Cara Operations (38%)	0.07
DEP	Deprenyl Research (68%)	0.15
FSH	Four Seasons Hotel	0.15
LWN	Loewen Group (77%)	0.23
MHG.A	MDS Health Group CL A (42%)	0.03
MHG.B	MDS Health Group CL B	0.18
QLT	Quadra Logic (67%)	0.04
13.0 FINANCIAL SERVICES (35)		20.98
13.1 Banks (7)		18.12
BMO	Bank of Montreal	3.01
BNS	Bank of Nova Scotia	2.51
CM	Cdn Imperial Bank of Commerce	3.62
LB	Laurentian Bank of Canada (38%)	0.07
NA	National Bank of Canada	0.86
RY	Royal Bank of Canada	4.88
TD	Toronto-Dominion Bank	3.17
13.2 Trust, Savings & Loan (6)		0.60
CEH	Central Capital (32%)	0.00
CEH.A	Central Capital CL A (78%)	0.00
CGA	Central Guaranty Trustco (12%)	0.00
NT	National Trustco (53%)	0.24
RYL	Royal Trustco (47%)	0.33
TTG	General Trustco (35%)	0.02
13.3 Investment Co's & Funds (6)		0.61
CGI	Canadian General Investments (52%)	0.05
CNN.UN	Canada Trust Income Investments	0.02
FMS.A	First Marathon CL A	0.12
IGI	Investors Group (25%)	0.15
MKF	Mackenzie Financial Corp	0.21
UNC	United Corporations (50%)	0.06

13.5 Insurance (6)		0.49
CRX	Crownx (22%)	0.00
CRX.A	Crownx CL A	0.01
ELF	E-L Financial (64%)	0.08
FFH	Fairfax Financial	0.05
GWO	Great West Lifeco (14%)	0.09
LON	London Insurance Group (43%)	
13.6 Financial Management Co's (10)		1.15
CXE	Consol. Canadian Express (57%)	0.01
CXS	Counsel Corp (53%)	0.03
DBC.A	Dundee Bancorp CL A	0.04
FC	FCA International (74%)	0.02
GLZ	Great Lakes Group (8%)	0.05
HIL	Hees International Bancorp (37%)	0.25
LGC.B	Laurentian Group CL B (31%)	0.03
PGC.A	Pagurian CL A (65%)	0.18
PWF	Power Financial Corp (31%)	0.30
TFC.A	Trilon Financial CL A (42%)	
14.0 CONGLOMERATES (12)		5.76
14.1 Conglomerates (12)		5.76
AGR.B	Agra Industries CL B (74%)	0.05
BL.A	Brascan CL A (49%)	0.45
CAM.A	Canam Manac CL A (31%)	0.01
CP	Canadian Pacific	3.31
FIL.A	Federal Industries CL A,B	0.14
HSM	Horsham	0.49
ISE	International Semi-Tech (65%)	0.07
JN	Jannock (73%) 0.19	
OCX	Onex Corporation (75%)	0.07
POW	Power Corp of Canada (61%)	0.60
SRC	Scott's Hospitality	0.33
SRC.C	Scott's Hospitality CL C	0.06
	TSE Composite	100.00