AN INVESTIGATION INTO THE EFFICIENT MARKET HYPOTHESIS: A CANONICAL CORRELATION ANALYSIS APPROACH

by

DAREN MCCROSSAN SMITH

B.Sc., University of Western Ontario, 1993

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

October 1995

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.

Department of Statistics

The University of British Columbia
Vancouver, Canada

Date Oct 13/95

ABSTRACT

In this thesis we will consider the Efficient Market Hypothesis (EMH). Fama (1970) defined three levels in which to test market efficiency: weak, semi-strong, and strong, each level depending on the particular set of information being used to assess efficiency. We will mainly address weak level efficiency in which the information set is past security data. Before the mid 1980's it was widely believed that the EMH was true at the weak and semi-strong levels. It was not until the pioneering work of Shiller (1984) and Summers (1986) that some doubt was cast on the EMH. They proposed an inefficient model in which prices consist of a sum of a random walk component and a stationary (predictable) component which represents the market valuation error. Since their initial conjecture about a stationary component in stock prices much effort has been spent in trying to determine if it exists and if it does, determining how much of the variations in stock prices it accounts for. To investigate this problem we will use a combination of data filtering, canonical correlation analysis, simulations and bootstrapping. Using industry price data obtained from the Toronto Stock Exchange over the period January 1956 to June 1995, we find some evidence against the EMH.

Table of Contents

List of Tables

List of Figures

List of Nomenclature

$\mathrm{vec}(\cdot)$ – The column stacking operator.

∇ – Extracts the upper diagonal portion of the matrix it is applied to

B – is the backshift operator, for example $BY_t = Y_t - Y_{t-1}$.

$\sim$ – this is read "distributed as".

$N_p(\mu, \Sigma)$ – this represents a p-dimensional normal vector.

$\mathcal{L}(\cdot)$ – this denotes a linear portfolio filter.

$\mathcal{F}(\cdot)$ – this denotes an information filter.

$\xrightarrow{d}$ – this is read "converges in distribution".

Acknowledgements

I would like to thank my supervisor, Dr. Jian Liu, for providing me with a very interesting topic to investigate in this thesis. I would also like to thank Dr. Ruben Zamar for his comments and suggestions on improving the manuscript. Thanks go out also to Wade Blanchard, who provided much of the data that I worked with and also valuable latex scripts. I would like to thank the Statistics Department as a whole for making my two years in Vancouver enjoyable and rewarding.

Chapter 1

Introduction

1.1 Investigations into the Efficient Market Hypothesis (EMH)

The capital market's primary role is the allocation of ownership of the economy's capital stock. Its secondary role is to determine the current price of a stock. The ideal market is one in which the prices determined by the market are accurate indicators of the real value of the stock price. Such an ideal market in which prices always fully reflect all available information is called efficient.

It has been debated for many years whether the capital markets are efficient or not. Generally speaking a believer in the EMH presumes the following two statements to be valid. 1) The market has a very efficient pricing mechanism which is capable of interpreting fundamental economic realities and arrive at a price which comes close to reflecting the true worth of the security. It is able to do this because the market investors are acting rationally. 2) Except for a very small number of highly skilled professionals the time and effort spent in trying to locate mis-priced stocks is not worth the reward. This means that the market may not reflect the "true" price, but the difference is so small that almost no gain can be achieved by revealing the discrepancy. On the other hand, a non-EMH believer would say the following: 1) The prices determined by the market are not reflective of fundamental economic realities, since they are determined in part by irrational investors. 2) It is highly probable that with a little effort abnormally high returns can be earned. This is because the market price and the "true" price may differ by a large amount and revealing the difference can lead to a large financial gain.

As can be seen, these two views are diametrically opposed to one another. At the heart of the difference is the EMH belief in rational investors and little or no potential for abnormally high returns (market exploitability), and the non-EMH belief that some investors act irrationally and that this can lead to prices which deviate from "true" values enough that revealing the difference can lead to large financial rewards. Rationality of investors and non-exploitability of the market have in the past been considered as equivalent, but in fact they should be considered as separate entities. For example, irrational investing is a necessary condition for market exploitability but it is not sufficient. This implies that even if the market value is substantially different from the "true" value it may not be possible for investors to tell the difference (for example West (1988)). In the early stages of research into the EMH most of the effort was spent addressing the possibility of earning abnormally high returns and only recently have researchers considered the rationality of investors. If either irrationality in market prices or exploitability of the market can be shown than the EMH has been proven invalid.

The implications of whether the EMH can be considered as a reasonable depiction of the market are very far-reaching. Much of the modern theory of finance and economics has been built up on the assumption that the EMH is valid, in particular the assumption of a rational investor (e.g the Capital Asset Pricing Model (CAPM)). This debate also concerns the average investor, since if the market can be considered as efficient than the optimal portfolio investing strategy in light of transaction costs is to buy and hold, whereas if the market is not efficient than it would be more profitable to pursue an active investment strategy.

Much of the early work on the EMH debate is summarized by Fama (1970) in a survey paper entitled "Efficient Capital Markets: A review of theory and empirical work" [7].

1.1.1 The EMH universally accepted

Fama begins by dividing "information" into three subsets, which allows the EMH to be tested at three levels of efficiency.

• The first level is what he calls weak efficiency, in which the information set is simple historical prices. At this level the market is efficient if tomorrow's price moves independently of any past prices. This implies that any "trading rules" or attempts to determine the future price by studying charts of past price data is useless.

• The second level is semi-strong efficiency, where the information set is all publicly available information (e.g., quarterly reports, announcements of annual earnings, stock splits, etc.). At this level it is of interest to determine if prices respond instantaneously and in an unbiased manner to new information. Efficiency at this level would imply that the market is able to immediately arrive at the best interpretation of new information.

• The third level is strong efficiency, which is concerned with whether certain investors or groups who have monopolistic access to information can use this knowledge to predict future prices. In other words this level is concerned with whether the market incorporates all published and unpublished information in its determination of the price. If the market is efficient at this level it implies that insider trading would not be successful.

Each of the three levels of efficiency has its own collection of statistical tests, called weak form level tests, semi-strong form tests, and strong form tests. It should be noted that the above three levels of efficiency are not independent. If the market is inefficient at the weak level than it will be inefficient at both of the other two levels. As well, inefficiency at the semi-strong level implies inefficiency at the strong level.

Fama perceptively notes that almost all of the empirical literature up to 1970 is based on the assumption that the conditions of market equilibrium can be stated in terms of expected returns. Notationally speaking, this can be summarized in the following equation:

$$E(\tilde{p}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{jt}$$

where E is the expected value operator: p_{jt} is the price of security j at time t; $p_{j,t+1}$ is its price at t+1; $r_{j,t+1}$ is the one-period percentage return $(p_{j,t+1} - p_{jt})/p_{jt}$; Φ_t is a general symbol for whatever set of information is assumed to be "fully reflected" in the price at time t; and the tildes indicate that $p_{j,t+1}$ and $r_{j,t+1}$ are random variables at t (because they are unknown at time t and will not be known until they are realized at time t+1).

For all three sets of "information", theories were developed and tested. In the class of weak form tests almost all of the literature up to 1970 supported the EMH. In fact, so much support was found for the EMH at this level that most researchers stopped testing for it and switched to testing at the semi-strong form level, assuming that all work at the weak forms level was completed. At the semi-strong level as well, the consensus was that the market was efficient. Finally, at the strong form test level a few instances where individuals were able to exploit inside information for profit were documented. The first example concerned specialists on major security exchanges who had monopolistic access to information on unexecuted limit orders and who were able to utilize their information to generate trading profits. The second example concerned corporate insiders who had monopolistic access to information concerning their firms. Besides these two cases there was little or no evidence against the EMH at any of the three levels of "information".

Fama's paper was published in 1970 and the views expressed in it were wide spread among statisticians and econometricians alike. This is what the great economist John Maynard Keynes had to say about the subject.

> Professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to

> the average preferences of the competitors as a whole; so that each competitor has to pick not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgement, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion excepts the average opinion to be. And there are some, I believe, who practice the fourth, five and higher degrees. [13]

This may not be the case, but what it does indicate is a belief that investors are merely guessing at future price movements and therefore that the EMH is true. This was a belief that many people held.

1.1.2 A digression: Terminology in the EMH debate explained

In the sections that follow certain statistical concepts relating to time series analysis will be used in the discussion of approaches to resolving the EMH debate. Some of these concepts will be explained now. Let Y_t represent a general time series.

Definition 1.1 (A white noise time series)

$$\begin{aligned} Y_t &= \epsilon_t \quad where, \\ E(\epsilon_t) &= 0 \\ Cov(\epsilon_j \epsilon_k) &= 0 \quad i \neq j \\ &= \sigma^2 \quad i = j \end{aligned}$$

Definition 1.2 (A random walk (RW) time series)

$$\begin{aligned} Y_t &= \mu + Y_{t-1} + \epsilon_t \\ Var(Y_t) &= t\sigma^2 \end{aligned}$$

where ϵ_t follows a white noise process.

Definition 1.3 (An autoregressive process of order p, AR(p))

$$Y_t = \gamma + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \epsilon_t$$

where ϵ_t follows a white noise process.

When $p = 1$ we have an AR(1) time series which has the following variance:

$$Var(Y_t) = \frac{\sigma^2}{(1-\alpha_1^2)}$$

Definition 1.4 (Stationary time series) *A stationary time series Y_t must satisfy the following two conditions. If it does not satisfy them it is referred to as nonstationary.*

$$\begin{aligned} E(Y_t) = constant, &\qquad \textit{does not depend on } t \\ Cov(Y_t, Y_{t-s}) = \sigma_s^2, &\qquad \textit{does not depend on } t \end{aligned}$$

Of the time series we have considered above, the AR(1) process with $\alpha \leq 1$ and the white noise process are stationary. The random walk is nonstationary.

Definition 1.5 (Predictability of a time series) *This refers to the noise to signal variance ratio of a stationary time series given by:*

$$\frac{Var(\epsilon_t)}{Var(Y_t)}$$

Greater predictability implies a lower noise to signal variance ratio. This ratio is $1 - \alpha_1^2$ for an AR(1) process and 1 for a white noise process indicating that it is stationary but unpredictable. As α increases in the AR(1) process, but is less than 1, greater predictability is obtained.

Much of the early work in support of the EMH claimed that prices followed a RW. This implies that prices and their first differences returns (which are white noise if prices follow a RW), do not have any predictable components.

1.1.3 The EMH questioned

It was not until 1986 that someone seriously challenged the EMH. In his seminal paper "Does the Stock Market Rationally Reflect Fundamental Values" [19] Summers (1986) submitted that the current statistical rationale for regarding capital markets as efficient was somewhat flawed. He noted that most of the early work on market efficiency was a result of analysing the auto-correlations of daily and weekly stock returns. The common consensus of the studies was that the estimated auto-correlations were very close to zero. This implied that there was no significant predictability of returns and hence the capital markets were efficient. His main contribution was to demonstrate that the auto-correlations of short-horizon returns can give the impression that the predictable components of prices are of no aid in prediction when in fact they can actually explain a moderate amount of return variations. As Summers says,

> the existing evidence does not establish that financial markets are efficient in the sense of rationally reflecting fundamentals. It demonstrates that the types of statistical tests which have been used to date have essentially no power against at least one interesting hypothesis to market efficiency [19].

Summers hypothesized the following inefficient model for prices:

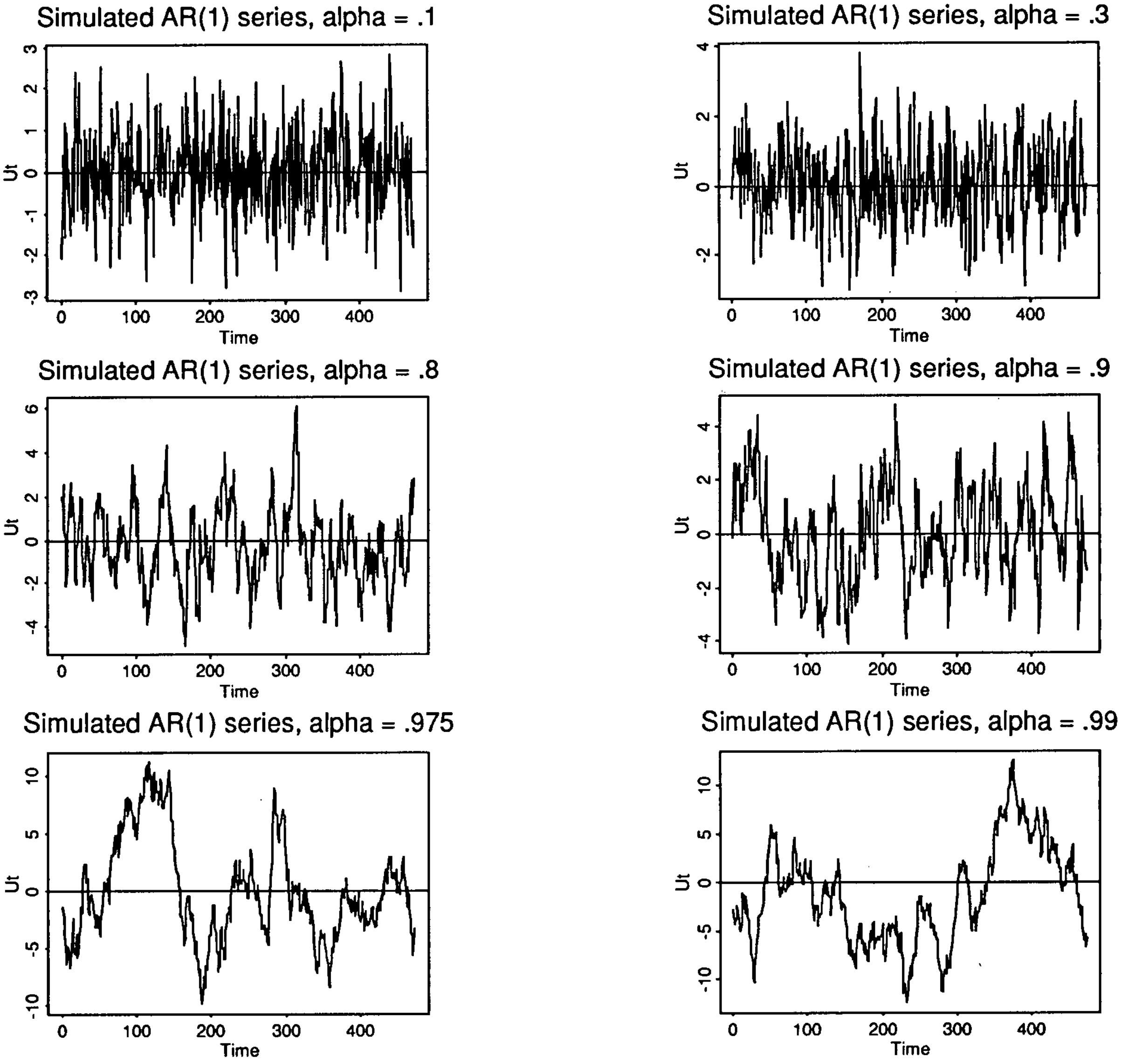
$$\begin{aligned} P_t &= P_t^* + u_t \\ u_t &= \alpha u_{t-1} + v_t \end{aligned} \tag{1.1}$$

where P_t is the market price, P_t^* is the market efficient price which follows a random walk (RW), $\{u_t\}$ are the valuation errors which follow a first-order autoregressive process (AR(1)), with errors given by $\{v_t\}$.

This model is commonly referred to as the "simple fads" model. The reasoning behind this name is explained well by Shiller (1984), who says:

> Investing in speculative assets is a social activity. Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others' successes or failures in investing. It is thus plausible that investors' behavior (and hence prices of speculative assests) would be influenced by social movements.

Model (1.1) is inefficient and partially predictable because the valuation errors $\{u_t\}$ are predictable as they represent persistent errors with a defined structure to them. If the $\{u_t\}$ were white noise than Model (1.1) would still be inefficient but there would be no predictable component to it. The random walk component P_t^* is commonly referred to as the permanent component of prices and u_t as the temporary component. The AR(1) component is called temporary because it is assumed that it is stationary and that even though it may differ from its constant expected value it will eventually revert to its mean. The larger the AR(1) coefficient the longer on average the process will take to return to its expected value. This is exemplified by the plots in Figure 1.1. Six AR(1) processes were simulated all with expected value zero. The process with the smallest α is highly variable and returns to its expected value very quickly; as α increases the $\{u_t\}$ become less variable but take longer swings away from their mean. Summers assumed an α of 0.98 for the $\{u_t\}$ process, which allows prices to take long temporary (since it will return to its mean eventually unlike a nonstationary series) swings away from fundamental values given by P_t^*. This has led some authors to refer to the investigation of a temporary component in prices as a search for mean reversion in prices.



Simulated AR(1) series, alpha = .1

Simulated AR(1) series, alpha = .3

Simulated AR(1) series, alpha = .8

Simulated AR(1) series, alpha = .9

Simulated AR(1) series, alpha = .975

Simulated AR(1) series, alpha = .99

Figure 1.1: Simulated AR(1) series with varying coefficients

Summers argued along the lines of Shiller (1984) that Model (1.1) is a reasonable hypothesis and also that it allows many instances of a market failing to rationally reflect fundamental values. Using Model (1.1) and a combination of assumed and empirically derived parameter values he conducted a weak form efficiency test. It is important to note that he is not saying his price model is the true one, only that it is reasonable and that the current statistical techniques would not be able to identify it if it was true. This basis for this is a belief that there may be many alternative models to efficiency which can also not be identified. In his analysis using monthly stock market returns over a 50-year period he showed that the standard statistical model could not reject the hypothesis of market efficiency. This was despite the fact that the market price frequently differed from the market efficient price by more than 30 percent.

He also conducted a semi-strong form test of market efficiency. The main conclusions of which were again that the common statistical tests of semi-strong efficiency had very little power against detecting his inefficient model for prices.

Summers makes a point of noting that he is not implying that abnormally high returns should be easy to make if prices actually follow his model. Instead he argues that for the same reasons it is very difficult for statisticians to uncover deviations of market prices from fundamental values it will be even more so for the average investor. The main point of his paper is to say that yes it may be exceedingly difficult to make abnormally high returns but that this does not imply that market prices reflect rational assessments of fundamental values which many people heretofore had assumed.

Thus Summers was able to demonstrate that at the weak form test level and the semi-strong form test level the EMH did not necessarily hold.

1.1.4 Evidence against the validity of the EMH

The additive components model for prices proposed by Summers spawned a flurry of research to determine if prices actually followed such a model. The most famous study was by Fama and French (hereafter F&F) (1988) entitled "Permanent and Temporary Components of Stock Prices" [8]. They proposed to investigate Summer's model by analysing the auto-correlations of long-horizon returns. Their rationale is given below:

> A slowly decaying component of prices (the AR(1) component in Summers' model) induces negative auto-correlation in returns that is weak for the daily and weekly holding periods common in market efficiency tests. But such a temporary component of prices can induce strong negative auto-correlation in long-horizon returns. [8]

Most of their work centered around the continuously compounded return from time t to t+T, $r_{t,t+T}$, which can be obtained from prices as follows:

$$r_{t,t+T} = p_{t+T} - p_t$$

where T is the holding period in years and ranges from 1 to 8.

Their tests involved calculating the slopes β_T in the following regressions:

$$r_{t,t+T} = \alpha_T + \beta_T r_{t-T,t} + \epsilon_{t,t+T} \tag{1.2}$$

The slopes β_T had to be biased adjusted because of the use of overlapping price data to determine T period returns. Under the assumption that prices follow Summers' "fads model", F&F were able to show that the biased adjusted slopes β_T in (1.2) represented the fraction of variation of returns that could be predicted. Using this hypothesis they reached some startling conclusions. The data they used consisted of the ten size-based deciles (decile ten contains the largest firms) plus the value-weighted and equal-weighted portfolios from the New York Stock Exchange (NYSE). Examining the slopes of the regressions in (1.2) for varying amounts of holding periods (T) for the time interval 1926-1985, they found that the predictable component of industry portfolio variances ranged from 25% to 40% of total 3-5 year return variances. This conclusion was in stark contrast to the expected near 0% under the efficient market hypothesis.

1.1.5 A re-evaluation of the F&F (1988) Results

Following the work of F&F (1988) were Eckbo and Liu (1993) in a paper entitled "Temporary Components of Stock Prices: New Univariate Results" [6]. In this paper they showed that if the model of stock prices that F&F used was modified slightly, their results were no longer valid. The slight modification involved allowing the $\{z_t\}$ to be a more general stationary process than the AR(1) that F&F had assumed. This modification was reasonable as Daniel and Torous (1991) [4] and Eckbo and Liu (1993) had both demonstrated that the "simple fads" model does not hold for the decile price portfolios of the NYSE.

The use of a more general stationary process implied that exact results for the fraction of variation of returns that can be explained by the stationary component was no longer possible. With this in mind Eckbo and Liu proposed a finite-sample lower bound estimator of the predictable variance proportion. Using this estimator they found that their lower bounds of the predictable component of 3-5 year return variances ranged from 10% to 17%. Though this was a conservative estimate, it was much different from the 25% to 40% that F&F obtained. The reason for the large difference in results is that the optimistic estimates of F&F are sensitive to model specification, whereas the conservative estimates of Eckbo and Liu allow for all possible model specifications subject to additivity of the price components.

1.1.6 Mounting evidence that the EMH does not hold

Tsay (1990) attacked the EMH question from a different angle than the studies previously mentioned. Tsay reasoned that under the EMH, the price today contains all of the available information on the price tomorrow. Hence, all the expected information lies in the relationship between prices today and prices tomorrow. Using this observation led Tsay, in his paper entitled "Correlation Transformation and Components of Stock Prices" [21], to consider a canonical correlation analysis between prices today and prices tomorrow. The idea to use a canonical correlation analysis to determine the structure of a time series was an old one first proposed by Box and Tiao (1977) in their paper "A canonical analysis of multiple time series" [2]. In this paper they showed that a k-dimensional autoregressive process of order p could be transformed into a process whose components were ordered from least to most predictable. Tsay used the same NYSE portfolio data as F&F (1988). His results were that for the period 1926 to 1989 the stationary (predictable) components of each decile explained from 10-15% of the variability in prices. He also subdivided the period into a smaller segment, 1941 to 1989, and found that the stationary price components explained 0-1.2% of the variation in stock prices. This indicated that the importance of the stationary components may have diminished greatly over the last 50 years.

1.1.7 The EMH re-considered

For a time after the work by Summers (1986) and F&F (1988) many researchers strongly believed that the EMH had been disproved and that all that remained was to determine by how much the market deviated from efficiency . Recently however, some disbelievers have appeared. Kim, Nelson, and Startz (1991) in their paper "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence" [14] discuss the results by F&F (1988).

They take issue with among other things, the claim issed by F&F that up to 30% of stock returns are predictable. This is an important consideration since what F&F demonstrated was a high degree of in-sample predictability, they mentioned nothing about an out-of-sample forecast. Kim et al considered such forecasting and showed that estimating the β_T coefficients of F&F in real-time, leds to very poor correlations between predicted returns and actual returns. Another disenter is Richardson (1993) with his paper "Temporary Components of Stock Prices: A Skeptic's View" [18]. He demonstrates that the large coefficients of β_T that F&F found could be consistent with prices that follow a RW. His work considers the distribution of the largest β_T , i.e the first order statistic, in the work by F&F. He shows via Monte Carlo simulation and bootstrapping that conditional on the largest value of β_T the results obtained by F&F can be explained by prices which follow a RW.

1.2 Our proposal

Obviously the EMH debate has not been settled yet. We propose to investigate it at the weak level, via price filtering and canonical correlation analysis. If inefficiency is found at this level it implies inefficiency at the other two levels. One of the implications of an inefficient market is that prices today can no longer be assumed to have digested all of the information available up to today. At best, prices today can provide a subset of information for price movement tomorrow. Consequently, we are led to find a better indicator for the expected price tomorrow based on all available price information up to today than merely the price today. This will be accomplished by searching for a (relatively) optimal predictive information filter of stock prices from the class of linear, quadratic and square-root filters. Once a predictive information filter has been found, its association with tomorrow's price will be investigated. This will be done using a canonical correlation analysis. The resulting transformed series will be analysed and used to infer results about the original stock price series.

The primary goal of this thesis is to find the predictive filter that produces a stationary (predictable) component in prices which can explain the largest (if any) percentage of variations in stock prices, and to determine if this percentage is statistically different from those generated by a RW model for prices. One reason why this is important will be explained by an example. Suppose that we are a fund manager and have several techniques at our disposal with which we can measure the amount of stock price variations which can be explained by a stationary (predictable) component of the stock price. We are primarily interested in finding the technique which offers the lowest amount of risk. In this case lower risk is equivalent to better prediction of future price movements. Thus we would like to obtain the method that allows us to predict the largest possible amount of variation in stock prices. Since the predictive filters are simply a re-arrangement of the past data, we are interested in finding the filter which allows us to explain the largest possible amount (if any) of variation in stock prices.

1.3 Layout of the thesis

Chapter 2 contains a description of our techniques and results. Chapter 3 contains a summary of the findings of this thesis, interpretations and indicates future directions for development.

1.4 The Data

The data consists of 14 monthly industry index values taken from the Toronto Stock Exchange, see Table 1.1 for a description. Each of the Toronto Stock Exchange indices measures 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. Notationally this corresponds to:

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

The data was obtained from the Toronto Stock Exchange Review [20].
The Appendix contains a detailed description of the method used to calculate the index.

Time series plots of the industry index values are displayed in Figures 1.2, 1.3 and 1.4 on pages 18,19 and 20. The indexes exhibit the same generals trends. They all started low and steadily rose to a local maximum around 1980 after which time there was a short period of decline. Following this decline they generally rose sharply and had another local maximum just before 1989, after this there was a period of decline and then a gradual increase until the present.

Table 1.1: The 14 industries of the TSE.

Industry	First Date	Last Date	Percentage of TSE market value as of June, 1995
Communications and Media	Jan, 1956	June, 1995	3.50
Financial Services	Jan, 1956	June, 1995	15.42
Gold and Silver	Jan, 1956	June, 1995	11.20
Industrial Products	Jan, 1956	June, 1995	15.95
Conglomerates	Jan, 1956	June, 1995	4.25
Merchandising	Jan, 1956	June, 1995	3.89
Metals and Minerals	Jan, 1956	June, 1995	8.36
Oil and Gas	Jan, 1956	June, 1995	11.20
Paper and Forest Products	Jan, 1956	June, 1995	5.47
Pipelines	Jan, 1956	June, 1995	2.32
Trans and Env. Services	Jan, 1956	June, 1995	1.70
Utilities	Jan, 1956	June, 1995	8.83
Consumer Products	Jan, 1956	June, 1995	7.67
Real Estate and Construction	Jan, 1968	June, 1995	0.26

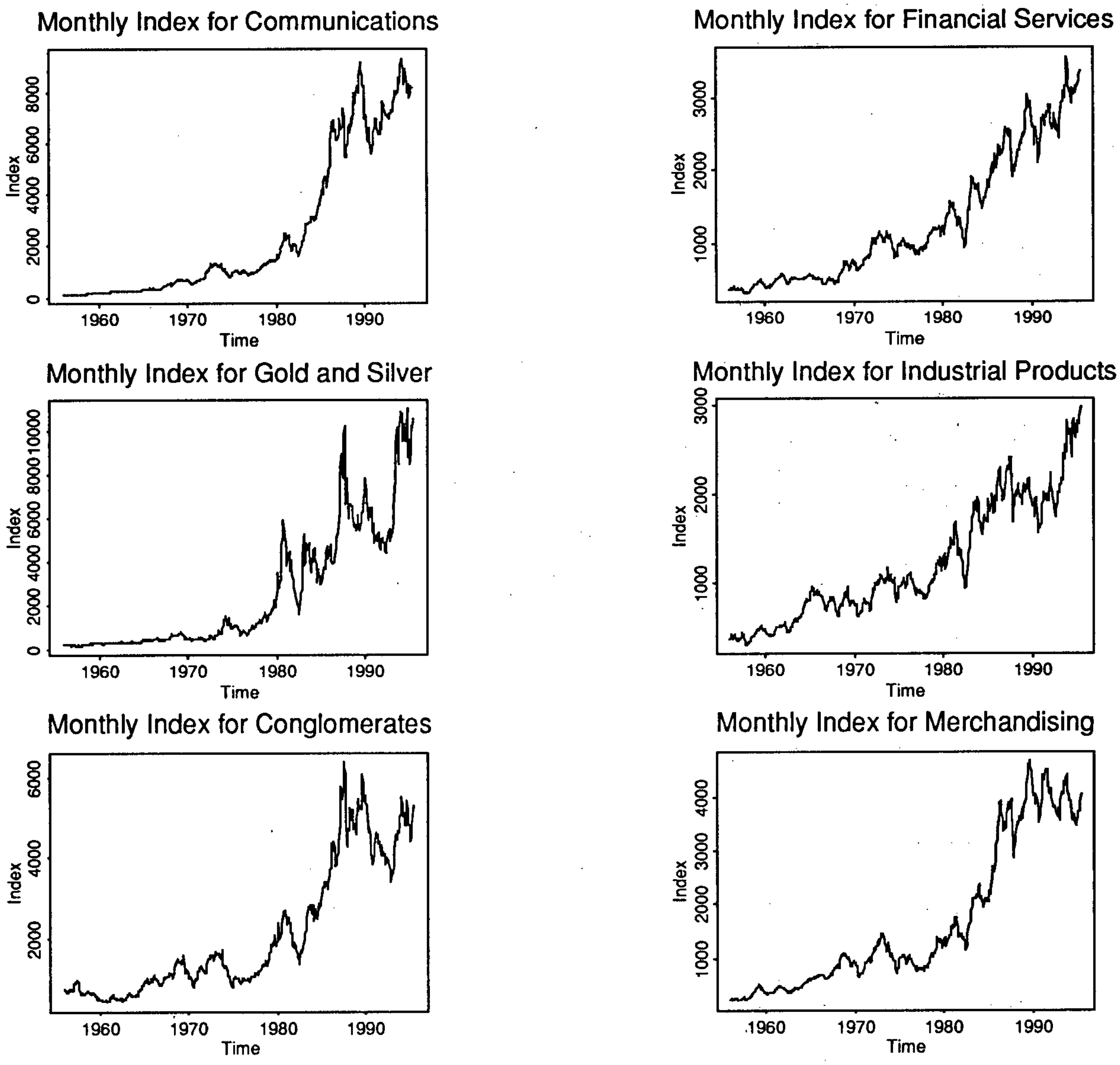


Figure 1.2: Plots of monthly TSE industry portfolio indexes

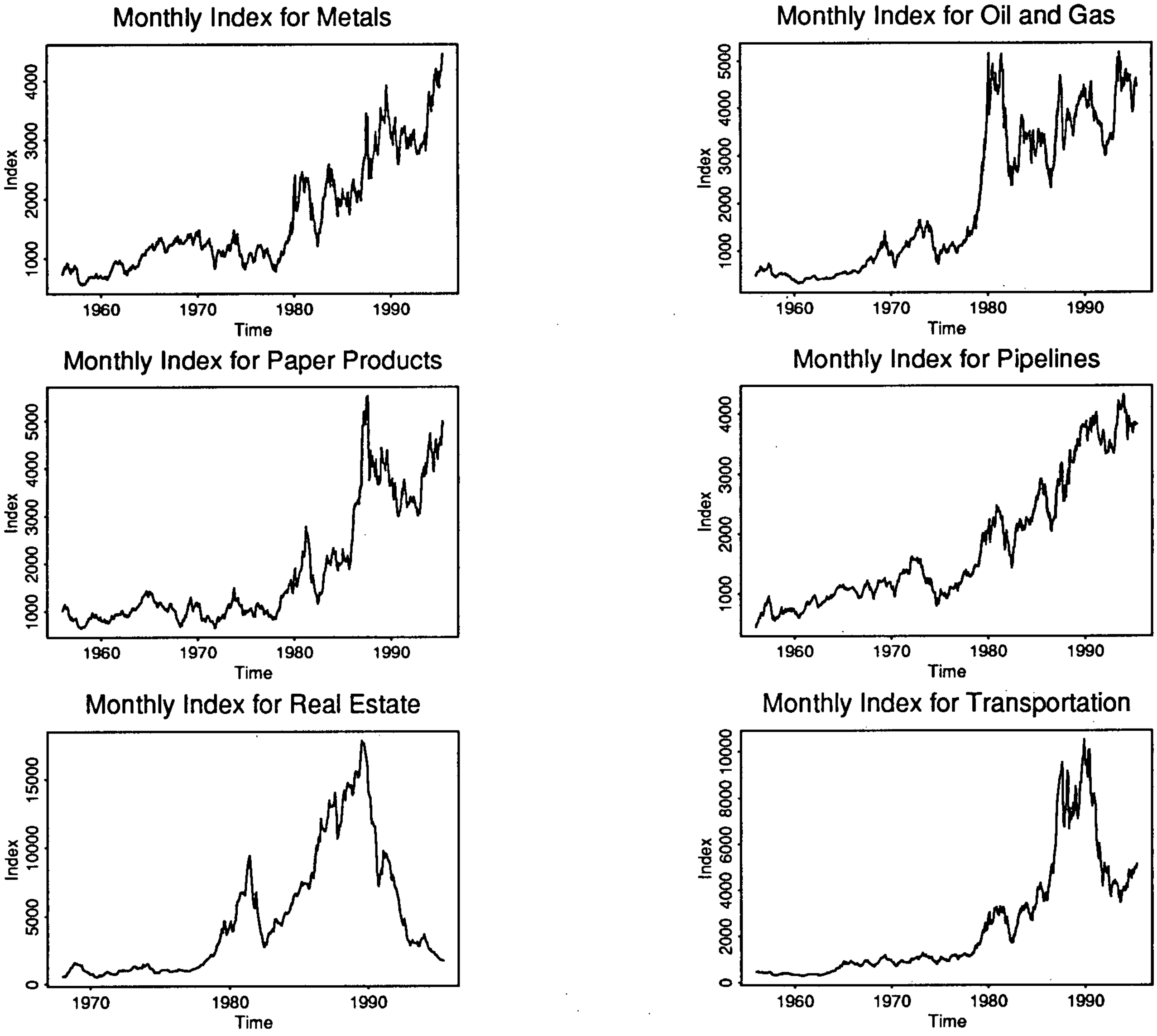


Figure 1.3: Plots of monthly TSE industry portfolio indexes

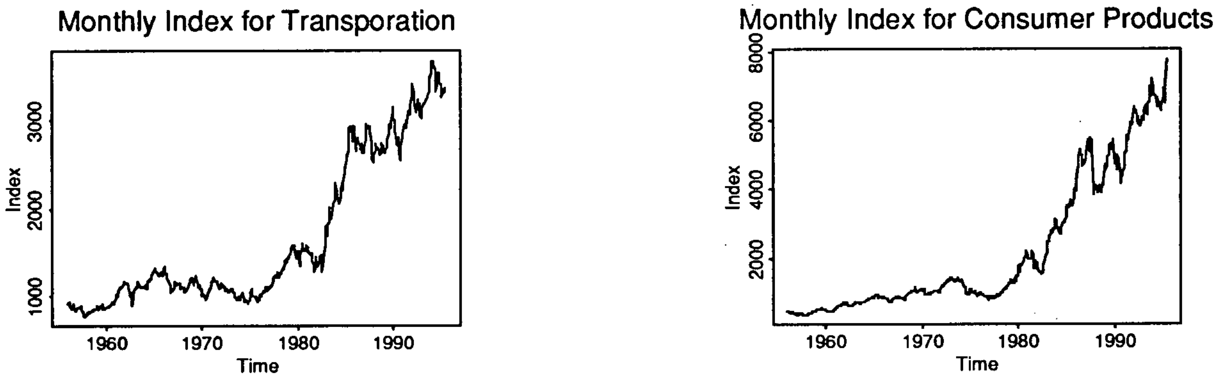


Figure 1.4: Plots of monthly TSE industry portfolio indexes

Chapter 2

The Mean-Reverting Price Component

2.1 Our Method and Rationale

We will consider price models of the following form,

$$p_t = q_t + z_t$$

where p_t is the log stock price at time t, z_t is the stationary or mean-reverting component of the stock price at time t, and q_t is the nonstationary component of the stock price at time t. Often $\{q_t\}$ and $\{z_t\}$ are assumed to be independent or at least uncorrelated to avoid trivial identifiability problems.

This is a generalization of the Summers (1986) "fads" model for prices, in which q_t and z_t can be any arbitrary nonstationary and stationary series. We will propose a method to obtain this decomposition. We will then then test via simulations and bootstrapping whether the the percentage of price variation explained by the stationary components is significantly different from results obtained under the EMH view that prices follow soley a RW.

Many early studies approached the EMH problem by simply using a linear regression of prices at time t on prices at time t-1, time t-2, and so on. This approach failed to uncover any relationship beyond a random walk for prices. The reason we believe it failed is because of a confounding effect of the stationary and nonstationary components of prices. The stationary component of prices represents the mean-reverting price component, and it is this component which may allow some of the variation in stock prices to be predicted. One approach to alleviating this problem of confounding is to first separate the stationary and nonstationary components of prices. We propose to do this through the use of predictive information filtering and canonical correlation analysis. Once this has been done a regression with stock prices at time t as the dependent variable and the stationary components of the canonical variate at time t as the independent variable can be performed. The R^2 of this regression represents the percentage of stock price variation that can be accounted for by the stationary z_t component.

2.2 On the use of returns to avoid the confounding problem

At this point it seems natural to ask if returns can be used to avoid the issue of confounding. To answer this question it is appropriate to recall the main goal of this thesis: to determine the amount of variation in stock prices that is accounted for by the stationary (predictable) component z_t in (2.1). The following notation will help us out. For simplicity, we assume that $\{q_t - q_{t-1}\}$ is stationary. This would be the case if $\{q_t\}$ followed a random walk.

$$p_t = q_t + z_t \tag{2.3}$$

$$var(p_t) = \underbrace{var(q_t)}_{\text{varies in t}} + \underbrace{var(z_t)}_{\text{const in t}} = \underbrace{\sigma_t^2}_{\text{varies in t}} \tag{2.4}$$

$$r_t = p_t - p_{t-1} = (q_t - q_{t-1}) + (z_t - z_{t-1}) \tag{2.5}$$

$$var(r_t) = \underbrace{var(q_t - q_{t-1})}_{\text{const in t}} + \underbrace{var(z_t - z_{t-1})}_{\text{const in t}} = \underbrace{\sigma^2}_{\text{const in t}} \tag{2.6}$$

Model (2.3) represents a general decomposition of prices into a stationary z_t component and a nonstationary q_t component. The price variance is given in Equation (2.4). The stationary component has a variance which is constant in time and the nonstationary component has a time-varying variance, which together imply that prices have a time-varying variance. Equation (2.5) represents the decomposition of returns. If returns are stationary then the variance of returns will be time independent and we have an identifiability problem. The reason for this is that we will not be able to separate the amount of variance attributed to $(q_t - q_{t-1})$ and $(z_t - z_{t-1})$ since both of these variances would also be constant. To avoid such problems we will use prices to resolve the confounding problem.

In the following three sections we will discuss the technical aspects of the canonical correlation analysis, Filtering, Auto-regressive integrated moving average (ARIMA) modeling and diagnostics, and how to determine the amount of variation in prices that can be attributed to stationary components.

2.3 Details of the Canonical Correlation Analysis

Before we carry out the canonical correlation analysis the theoretical and sample results will be discussed:

Notation: Let $\underbrace{\underline{X}}_{k\times 1} = (X_1 X_2 \ldots X_k)'$ represent a $(k \times 1)$ vector random variable. The sample realizations of the jth component of this vector will be denoted by $\{x_{jt}\}$.

2.3.1 Theoretical Canonical Correlation Analysis

Canonical Correlation analysis seeks a measure of association between two groups of variables. Let the first group of p variables be represented by the (p×1) random vector $\underline{X}^{(1)}$ and the second group of q variables be represented by the (q×1) random vector $\underline{X}^{(2)}$. In the development which follows we assume that $\underline{X}^{(1)}$ represents the smaller set, so that p≤q. The following theorem will be used to calculate the theoretical canonical variates $\underline{Y}$ and $\underline{Z}$, both (p×1) random vectors:

Theorem 2.1 (Canonical Variates and Correlation) *Let the random vectors* $\underline{X}^{(1)}$ *and* $\underline{X}^{(2)}$ *have* $Cov(\underline{X}^{(1)}) = \Sigma_{11}$, $Cov(\underline{X}^{(2)}) = \Sigma_{22}$, *and* $Cov(\underline{X}^{(1)}, \underline{X}^{(2)}) = \Sigma_{12}$, *where* Σ *has full rank. For coefficient vectors* $\underline{a}$ *and* $\underline{b}$, *form the linear combinations* $Y_i = \underline{a}_i'\underline{X}^{(1)}$ *and* $Z_i = \underline{b}_i'\underline{X}^{(2)}$. *Then the maximum* $Corr(Y_i, Z_i) = \rho_1^*$ *is attained by the following linear combinations (called the first canonical variate pair):*

$$Y_1 = \underline{e}_1'\Sigma_{11}^{-1/2}\underline{X}^{(1)} \quad \textit{and} \quad Z_1 = \underline{f}_1'\Sigma_{22}^{-1/2}\underline{X}^{(2)}$$

The kth pair of canonical variates j=2,3, ..., p is given by:

$$Y_k = \underline{e}_k'\Sigma_{11}^{-1/2}\underline{X}^{(1)} \quad \textit{and} \quad Z_k = \underline{f}_k'\Sigma_{22}^{-1/2}\underline{X}^{(2)}$$

maximizes,

$$Corr(Y_k, Z_k) = \rho_k^*$$

among those linear combinations uncorrelated with the preceding 1,2, ..., k-1 canonical variables. Here $\rho_1^{*2} \geq \rho_2^{*2} \geq \ldots \rho_p^{*2}$ *are the p ordered eigenvalues of* $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ *and* $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_p$ *are the associated (p× 1) orthonormal eigenvectors. (The quantities* $\rho_1^{*2}, \rho_2^{*2} \ldots \rho_p^{*2}$ *are also the p largest eigenvalues of the matrix* $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$ *with corresponding (q× 1) orthonormal eigenvectors* $\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_q$*)*

2.3.2 Sample Canonical Correlation Analysis

The following theorem will be used to obtain the sample canonical variates $\{y_{jt}\}$ and $\{z_{jt}\}$ $j = 1, 2, \ldots, p$. The notation $\{y_{jt}\}$ denotes the vector of observations for the jth canonical variate. The two groups in which we are interested in assessing the association between are represented by $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ (bold faces are used to indicate matrices). Each group is assumed to have n observations per variable.

Definition 2.1 (Notation for the two groups)

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^{(1)} \\ \dots \\ \mathbf{x}^{(2)} \end{bmatrix} = \begin{bmatrix} x_{11}^{(1)} & x_{12}^{(1)} & \dots & x_{1n}^{(1)} \\ x_{21}^{(1)} & x_{22}^{(1)} & \dots & x_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ x_{p1}^{(1)} & x_{p2}^{(1)} & \dots & x_{pn}^{(1)} \\ \dots & \dots & \dots & \dots \\ x_{11}^{(2)} & x_{12}^{(2)} & \dots & x_{1n}^{(2)} \\ x_{21}^{(2)} & x_{22}^{(2)} & \dots & x_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ x_{q1}^{(2)} & x_{q2}^{(2)} & \dots & x_{qn}^{(2)} \end{bmatrix} = [\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n] \qquad \textit{where,} \quad \underline{x}_j = \begin{bmatrix} \underline{x}_j^{(1)} \\ \dots \\ \underline{x}_j^{(2)} \end{bmatrix}$$

Definition 2.2 (Sample Covariance)

$$S_{kl} = \tfrac{1}{n-1}\Sigma_{j=1}^{n}(\underline{x}_j^{(k)} - \underline{\bar{x}}^{(k)})(\underline{x}_j^{(l)} - \underline{\bar{x}}^{(l)})', \quad k,l = 1,2 \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}$$

Now that we have defined the notation above we can use the following theorem to obtain the p sample canonical variate pairs $\{y_{jt}\}$ and $\{z_{jt}\}$.

Theorem 2.2 (Sample Canonical Variates and Correlation)

Let $\hat{\rho}_1^{*2} \geq \rho_2^{*2} \geq \dots \rho_p^{*2}$ be the p ordered eigenvalues of $S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}$ with corresponding orthonormal eigenvectors $\hat{\underline{e}}_1, \hat{\underline{e}}_2, \dots, \hat{\underline{e}}_p$, where the S_{kl} are defined in Definition (2.2). Let $\hat{\underline{f}}_1, \hat{\underline{f}}_2, \dots, \hat{\underline{f}}_p$ be the p largest orthonormal eigenvectors of $S_{22}^{-1/2}S_{21}S_{11}^{-1}S_{12}S_{22}^{-1/2}$. The kth canonical variate pair is:

$$\{y_{kt}\} = \hat{\underline{e}}_k' S_{11}^{-1/2}\mathbf{x}^{(1)} \quad \textit{and} \quad \{z_{kt}\} = \hat{\underline{f}}_k' S_{22}^{-1/2}\mathbf{x}^{(2)}$$

where $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are as in Definition (2.1).

See Johnson and Wichern [11] for a more detailed treatment of canonical correlation analysis.

2.4 A New Look at Canonical Correlation Analysis

2.4.1 Historical Overview

The classical canonical correlation analysis was developed by Hotelling (1936) primarily in the context of analyzing I.I.D. (independent identically distributed) data sampled from populations with multivariate normal distributions. Recently it has also been used to partition a general multivariate time series into stationary and nonstationary components (Tsay (1990)).

One of the drawbacks of canonical correlation analysis is that it only uses the first two moments of the empirical distribution of the data. Since the normal distribution is typically characterized by its first two moments and is well known and easy to use, such an analysis is usually confined to the class of population distributions which are at least approximately normal. However, since most data is only approximately normal at best we are led to consider the use of Filters which allow the use of higher order moment information.

2.4.2 Filters

Let $\underline{X}^{(1)}$ be $k_1 \times 1$ and $\underline{X}^{(2)}$ be $k_2 \times 1$ random vectors.

Definition 2.3 (Portfolio Filter) *A portfolio filter $\mathcal{L}(\cdot)$ is a (measurable) transformation given by $\mathcal{L}(\cdot) : \Re^{k_1} \longrightarrow \Re^{k_1}$.*

Two examples of a portfolio filter are considered below.

Example #1: A linear portfolio filter $\mathcal{L}(\cdot)$ is given by:

$$\mathcal{L}(\underline{X}^{(1)}) = A\underline{X}^{(1)}$$

where A is a $k_1 \times k_1$ matrix of constants.

Example #2: A quadratic portfolio filter $\mathcal{L}(\cdot)$ is given by:

$$\mathcal{L}(\underline{X}^{(1)}) = A\underline{X}^{(1)} + B\text{vec}(\nabla \underline{X}^{(1)}\underline{X}^{(1)'})$$

where A is a $k_1 \times k_1$ matrix of constants, B is a $k_1 \times \frac{k_1(k_1+1)}{2}$ matrix of constants, vec$(\cdot)$ is the column stacking operator and ∇ is an operator that returns the upper diagonal portion of a matrix. We need to use the ∇ operator because the matrix it operates on is symmetric and we only want the unique elements.

Definition 2.4 (Information Filter) *An information filter $\mathcal{F}(\cdot)$ is a (measurable) transformation given by $\mathcal{F}(\cdot) : \Re^{k_1+k_2} \longrightarrow \Re^{k_1}$.*

There are two main uses of an information filter. The first is as a smoother or recalibration technique in which both $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ are observed. For example, let $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ represent repeated observations from the same phenomena in which we would like to estimate a parameter. Information filtering may lead to a more accurate measurement of this parameter than either of them could provide individually. Two examples of an information filter are considered below.

Example #3: A Linear Smoothing Information Filter $\mathcal{F}(\cdot)$ is given by:

$$\mathcal{F}(\underline{X}^{(1)}, \underline{X}^{(2)}) = A_1\underline{X}^{(1)} + A_2\underline{X}^{(2)}$$

where A_1 is a $k_1 \times k_1$ matrix of constants and A_2 is a $k_1 \times k_2$ matrix of constants.

The second use of information filtering is as a prediction scheme. In this scenario it is assumed that $\underline{X}^{(1)}$ is unobserved and $\underline{X}^{(2)}$ is observed. We also assume that there is some association between $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ that we are interested in exploiting to predict $\underline{X}^{(1)}$. In the analysis which follows in the rest of the thesis $\underline{X}^{(1)}$ represents the vector of tomorrow's prices, $\underline{X}^{(2)}$ represents the vector of all available past price information and $\mathcal{F}(\underline{X}^{(2)})$ represents our prediction of tomorrow's prices, which we will call $\widehat{\underline{X}}_t$.

Example #4: A Linear Predictive Information Filter $\mathcal{F}(\cdot)$ is given by:

$$\mathcal{F}(\underline{X}^{(2)}) = A_3\underline{X}^{(2)}$$

where A_3 is a $k_1 \times k_2$ matrix of constants.

More generally, for the purpose of predicting $\underline{X}^{(1)}$ using $\underline{X}^{(2)}$, the choice of the Predictive Information Filter is often given by $\mathcal{F}(\underline{X}^{(2)}) = E(\underline{X}^{(1)}|\underline{X}^{(2)})$.

2.4.3 Uses of Filtering

Our goal in using filtering is to restructure the components of $\underline{X}^{(1)}$ and $\mathcal{F}(\underline{X}^{(1)})$ such that the new portfolios $\mathcal{L}_1(\underline{X}^{(1)})$ and $\mathcal{L}_2 \circ \mathcal{F}(X^{(2)})$ have their respective pairwise correlations in decreasing order of magnitude and all pairs of the new portfolios are uncorrelated with previous pairs.

When we restrict our attention to the class of linear portfolio filters, we are led to examine the canonical correlation analysis between $\underline{X}^{(1)}$ and the current information $\mathcal{F}(\underline{X}^{(2)})$. From the point of view of portfolio prediction, $\mathcal{L}_2 \circ \mathcal{F}(\underline{X}^{(2)})$ can be viewed as the best linear predictor of tomorrow's portfolios $\mathcal{L}_1(\underline{X}^{(1)})$.

The utility of using information filtering for predictive purposes will be illustrated in the following example:

Example #5: Let $X^{(2)} \sim N(0,1)$ and $X^{(1)} = X^{(2)^2}$. Since $\text{corr}(X^{(1)}, X^{(2)}) = 0$ the only canonical variates possible in the usual canonical correlation analysis setting are $X^{(1)}$ and $X^{(2)}$. We must then conclude falsely that the canonical variates $X^{(1)}$ and $X^{(2)}$ are unrelated due to their zero correlation and hence that $X^{(2)}$ has no predictive power for

$X^{(1)}$. However, with the use of information filtering it is possible to achieve maximum correlation if one uses the canonical variates $X^{(1)}$ and $\mathcal{F}(X^{(2)})$, where $\mathcal{F}(X^{(2)}) = X^{(2)^2}$.

2.4.4 Filtering in a time-series setting

As mentioned previously the other main use of canonical correlation analysis is to separate the stationary and nonstationary components of a general multivariate time-series. Since most time-series are only approximately normally distributed the usual canonical correlation analysis will give sub-optimal results, as measured by the regression R^2 in a regression of the time-series on its stationary components. However, if we utilize higher order moment information contained in the data by selecting appropriate portfolio and information filters we will obtain a more precise R^2.

In general, the use of information filtering and canonical correlation analysis go hand in hand. The first step is to determine a (relatively) optimal predictive information filter within a class of filters $\{\mathcal{F}\}$. In this thesis we will consider the class of linear, quadratic and square-root predictive information filters. More generally, we may consider $\mathcal{F}(\underline{X}^{(2)}) = E(\underline{X}^{(1)}|\underline{X}^{(2)})$. The next step is to choose an optimal portfolio filter by maximizing the correlation between $\mathcal{L}_1(\underline{X}^{(1)})$ and $\mathcal{L}_2 \circ \mathcal{F}(\underline{X}^{(1)})$, subject to both $\mathcal{L}_1(\cdot)$ and $\mathcal{L}_2(\cdot)$ being in a certain class of filters. Using this procedure will result in a constantly changing portfolio given by $\mathcal{L}_1(\underline{X}^{(1)})$. It changes because as each month goes by and another data point is added to the data the entire analysis is re-done and a new $\mathcal{L}_1(\underline{X}^{(1)})$ results. In this thesis we will consider the class of linear filters which is equivalent to a canonical correlation analysis. It is the $\mathcal{L}_1(\underline{X}^{(1)})$ portfolio that we are interested in. Using it we will be able to infer properties of $\underline{X}^{(1)}$. For ease of notation the jth series of the canonical variate $\mathcal{L}_1(\underline{X}^{(1)})$ will be denoted by $\{y_{jt}\}$.

2.4.5 Filtering in an investment setting

If we were to consider long term investments with relatively unchanging portfolio's, we would require our portfolio filter $\mathcal{L}_1$ to be the same each time we repeated the analysis.

To simplify our discussion, we shall in the following consider only linear portfolio filters. The following equations will be discussed:

$$\begin{aligned}
\underbrace{\underline{X}^{(1)}}_{p\times 1} &= \mathcal{F}(\underline{X}^{(2)}) + \underline{\epsilon} && (2.7)\\
y_{\underline{a}} &= \underline{a}'\underline{X}^{(1)}, \qquad \underline{a} \in \Re^p && (2.8)\\
y_{\underline{a}} &= \underline{a}'\mathcal{F}(\underline{X}^{(2)}) + \underline{a}'\underline{\epsilon} && (2.9)\\
\underbrace{Var(y_{\underline{a}})}_{signal} &= Var(\underline{a}'\mathcal{F}(\underline{X}^{(2)})) + \underbrace{Var(\underline{a}'\underline{\epsilon})}_{noise} && (2.10)\\
c(\underline{a}) &= \frac{Var(\underline{a}'\underline{\epsilon})}{Var(y_{\underline{a}})} = 1 - \frac{Var(\underline{a}'\mathcal{F}(\underline{X}^{(2)}))}{Var(y_{\underline{a}})} && (2.11)\\
c^*(\underline{a}) &= \frac{\underline{a}'Var(\mathcal{F}(\underline{X}^{(2)}))\underline{a}}{\underline{a}'Var(\underline{X}^{(1)})\underline{a}} && (2.12)
\end{aligned}$$

Equation (2.7) represents the prediction equation for $\underline{X}^{(1)}$. Equation (2.8) represents a portfolio of stocks. If we substitute Equation (2.7) into (2.8) we obtain Equation (2.9), the variance of which is given in Equation (2.10). We now define the noise-to-signal variance ratio as Equation (2.11). We want to minimize this value, which is equivalent to maximizing Equation (2.12).

Now let $\rho_1 \geq \rho_2 \geq \ldots \rho_p$ and $\underline{a}_1, \underline{a}_2 \ldots \underline{a}_p$ be the eigenvalues and linearly independent right eigenvectors of $[Var(\underline{X}^{(1)})]^{-1} \cdot Var[\mathcal{F}(\underline{X}^{(2)})]$. Then the portfolio's $\underline{a}_1'\underline{X}^{(1)}, \ldots, \underline{a}_p'\underline{X}^{(1)}$ are arranged from the least predictable to most predictable, as measured by the noise-to-signal variance ratio. However, we will not consider such an analysis, the interested reader is referred to Box and Tiao (1977) [2] for a detailed discussion in the linear time series setting.

2.4.6 Filtering used to solve an identifiability problem

Sometimes, an intelligent choice of a portfolio filter will help to solve the problem of identifiability of the additive stationary and nonstationary price components.

For example, suppose prices follow a k-dimensional vector RW model driven by an I.I.D. noise sequence given by:

$$\underline{p}_t = \underline{p}_{t-1} + \underline{\epsilon}_t, \qquad t = 1, 2, \ldots \qquad \{\underline{\epsilon}_t\} \ I.I.D. \ N_k(\underline{0}, \Sigma) \tag{2.13}$$

If the components of $\underline{\epsilon}_t$ are highly correlated, so are those of $\underline{p}_t$. This means that some portfolio filters $\mathcal{L}$, could result in components of $\mathcal{L}(\underline{p}_t)$ that are approximately stationary. For the stationary series of $\mathcal{L}(\underline{p}_t)$ it is no longer possible to to separate the stationary components from the nonstationary ones.

However, if we left-multiply $\underline{p}_t$ in Equation (2.13) by $\Sigma^{-1/2}$ to create $\underline{p}_t^*$, which has independent components, we will be free of this particular identifiability problem. This is illustrated in the following equations:

$$\begin{aligned} \mathcal{L}(\underline{p}_t^*) &= A\underline{p}_t^* \\ &= \begin{pmatrix} \sum_{i=1}^k a_{1i} p_{it}^* \\ \sum_{i=1}^k a_{2i} p_{it}^* \\ \vdots \\ \sum_{i=1}^k a_{ki} p_{it}^* \end{pmatrix} \end{aligned}$$

Since the components of $\underline{p}_t^*$ are independent RW's then the variance of the bth component of $\mathcal{L}(\underline{p}_t^*)$ is given by:

$$Var(\sum_{i=1}^k a_{bi} p_{it}^*) = \sum_{i=1}^k a_{bi}^2 t \sigma_b^2$$

This variance depends on t for all b, hence all of the components of $\mathcal{L}(\underline{p}_t^*)$ are nonstationary. This means that it will still be possible to separate the stationary and nonstationary components.

2.5 ARIMA modeling and diagnostics

In the following two sections we will discuss the techniques used to fit the auto-regressive integrated moving average (ARIMA) models and also the diagnostic tools that were used to evaluate the goodness of fit.

2.5.1 ARIMA modeling

The basic procedure followed was as in Box and Jenkins [3].

Suppose our series is represented by $\{y_{jt}\}$. The procedure is as follows:

1. Difference $\{y_{jt}\}$ until it is stationary.
2. Identify the appropriate auto-regressive moving average (ARMA) process.

The main tools used in step 2 are the auto-correlation function (acf) and the partial auto-correlation functions (pacf) as defined in Box and Jenkins. The behavior of these functions will indicate which class of ARMA models is appropriate for the series under consideration.

The model(s) indicated by the acf and pacf were fitted using the S-plus arima.mle function. If more than one model was appropriate the one with the better goodness of fit based on the Ljung-Box statistic was chosen. The Ljung-Box statistic is discussed in the next section.

2.5.2 ARIMA diagnostics

In this section we will discuss the diagnostic tools that were used to check the adequacy of the fitted ARIMA models.

2.5.3 Using the acf to check for serial dependence in residuals

To check for serial dependence in the residuals, we define the following:

Definition 2.5 *The auto–covariance function γ_τ of a series , $\{z_t\}_1^N$, where N is the number of observations, is given by:*

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

and the auto–correlation function (acf) ρ_τ is given by:

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

Definition 2.6 *The sample analogue of γ_τ, denoted by $\widehat{\gamma}_\tau$, is given by:*

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

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

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

For more details on these functions see Box and Jenkins [3].

We will also need the following result due to Bartlett:

Definition 2.7

$$\widehat{Var}(\widehat{\rho}_\tau) \approx \frac{(1 + 2(\widehat{\rho}_1^2 + \widehat{\rho}_2^2 + \ldots + \widehat{\rho}_q^2))}{N} \qquad \tau > q.$$

We are sometimes interested in testing for zero auto-correlation after a specified lag q (for example, a moving average process of order 2 (MA(2)) should have an acf which is zero for any lag greater than 2, in Definition (2.7) τ would represent a general lag > 2). To do this we first compute the estimated variance of the sample auto-correlation given by Definition (2.7). We can then use the result due to Andersen that for reasonably large N, $\widehat{\rho}_\tau$ is approximately distributed as a standard normal random variable. The last step is to make a plot of the sample auto-correlation function and determine how many points lie outside the 95% confidence interval of zero auto-correlation. If many points are outside this limit it is an indication that the residuals are serially correlated and the ARIMA model may be inadequate.

2.5.4 Using a runs test to check for serial dependence in residuals

Another method of checking for serial dependence is given by using a runs test [15]. A runs test is usually used to test data to see if the order is "random". A run is one or more observations in a row greater than the median or one or more observations in a row less than or equal to the median. Wald and Wolfowitz (1940) showed that as m and n tend infinity with m/n tending to γ then:

Theorem 2.3

$$\frac{R - 2m/(1+\gamma)}{\sqrt{4\gamma m/(1+\gamma)^3}} \xrightarrow{d} N(0,1)$$

where R is the total number of runs, m is the number of observations greater than the median, and γ is the ratio of m and the number of observations less than or equal to the median.

A simple check for independence is given by calculating the above ratio and comparing it to the appropriate critical value from the standard normal distribution.

2.5.5 The Ljung-Box test for lack of fit

Ljung and Box [16] have suggested the following measure of lack of fit for time series models:

Definition 2.8

$$Q(m) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\hat{\rho}_k^2,$$

where m is the number of lags used, n is the length of the series, and $\hat{\rho}_k$ is the fitted residual auto-correlation lag-k. For large n, $Q(m)$ is distributed as χ^2 with m-p-q degrees of freedom, where p is the number of AR components in the model and q is the number of MA components in the model.

A check on the adequacy of any ARIMA model is given by computing $Q(m)$ in Definition (2.8) and comparing it with the 95% critical value from the appropriate χ^2_{m-p-q} distribution.

2.6 The Dickey Fuller Unit Root Test

We will be interested in determining whether a given series is stationary or nonstationary. Due to the fact that existing studies have primarily centered around the RW, we will consider testing for this type of nonstationarity. We use the test developed by Dickey and Fuller (1979) [5]. Consider the following autoregressive model:

$$Y_t = \rho Y_{t-1} + \epsilon_t, \qquad t = 1, 2, \ldots, n$$

It is assumed $Y_0 = 0$, ρ is a real number and $\{\epsilon_t\}$ is a sequence of I.I.D. random variables with mean zero, variance σ^2 and finite fourth moment.

If $\rho = 1$ the time series is a random walk (RW), and it has a unit root. Dickey and Fuller developed tests for a unit root based on the following two statistics.

$$\begin{aligned} \hat{\rho} &= (\sum_{t=1}^{n} Y_t^2)^{-1} \sum_{t=2}^{n} Y_t Y_{t-1} && (2.14) \\ \hat{\tau} &= (\hat{\rho} - 1) S_e^{-1} (\sum_{t=2}^{n} Y_{t-1}^2)^{\frac{1}{2}} \quad \text{where,} && (2.15) \\ S_e^2 &= (n-2)^{-1} \sum_{t=2}^{n} (Y_t - \hat{\rho} Y_{t-1})^2 \end{aligned}$$

The empirical distribution of these two statistics is given in Fuller [9] (1976, pp. 371,373). To test for a unit root one simply calculates either of the above two statistics and compares it to a critical value. We choose to use $\hat{\rho}$ to classify the series. The series is nonstationary if $n(\hat{\rho} - 1)$ is between the critical values of -10.4 and 1.61 (this corresponds to a two sided test region with probability of falsely rejecting the null of a unit root equal to 5%).

A more recent unit root test given by Phillips (1987) [17] was also used in our subsequent analysis. The initial results obtained by both methods were very similar but since the Phillips test took much longer computationally to evaluate it was decided to use only the Dickey Fuller unit root test.

2.7 Variation in prices explained by stationary components

Each of the k series of the canonical variate $\mathcal{L}_1(\underline{X}^{(1)})$, which we have denoted by $\{y_{jt}\}$ $j = 1, 2 \ldots k$, were analysed to determine if they were stationary or nonstationary. When the stationary $\{y_{jt}\}$ were identified we performed k separate regressions. For each regression one of the original logarithm of stock price series was the dependent variable and the stationary $\{y_{jt}\}$ series were the independent variables. To estimate the percentage of stock price variation that is attributed to stationary (predicatable) components we used the R^2 from the following regression:

Definition 2.9 *The percentage of stock price variation in each industry that is attributed to stationary (predictable) components is given by the R^2 in the following regression:*

$$x_{jt} = \gamma_0 + \gamma_{v+1} y_{v+1,t} + \ldots + \gamma_k y_{k,t} + \varepsilon_{jt}$$

where $\{x_{jt}\}$ is the jth original series, $\{y_{jt}\}$ is the jth stationary canonical variate series, k is the number of stock price series in the data, v is the number of nonstationary $\{y_{jt}\}$ series and $\{e_{jt}\}$ is white noise.

2.8 Procedure that was followed

We started out with k logarithm of stock price series. We first considered filters within the class of linear, quadratic and square root predictive information filters. We then performed a canonical correlation analysis between the original logarithm of stock price series $\underline{X}^{(1)}$ and a predictive information filter $\mathcal{F}(\underline{X}^{(1)})$. The resulting k canonical variate series $\{y_{jt}\}$ $j = 1, 2, \ldots k$ were modeled using the standard ARIMA modeling techniques discussed in Section 2.3. We were primarily interested in determining which of the k $\{y_{jt}\}$ series were stationary and which were nonstationary. We then performed k separate regressions with each of the original logarithm of stock price series as the dependent variable and the stationary components of $\{y_{jt}\}$ as the independent variables, as discussed in Section 2.4. The R^2's of these regressions represent the amount of variation in the stock price series that can be explained using the particular predictive information filter under consideration. Finally, we consider the question of whether the R^2's obtained in these regressions are statistically different than what is expected under a RW model for prices.

2.9 The Data Revisited

The data that we used are the industry index values taken from the Toronto Stock Exchange (TSE), see Table 1.1 for a description. As can be seen from this table all of the series except real-estate start on Jan, 1956. Real-estate commences on Jan, 1968. In our analysis we performed a canonical correlation analysis using these industry index values. To do such an analysis, all of the series had to be the same length. We could have either truncated all of the series so that they started on Jan, 1968, or simply deleted the real-estate index. Since the real-estate industry only consists of 0.26% of the total TSE market value, we chose to proceed by deleting the real-estate index from the data. Thus the data that we used consisted of 13 industries of the TSE each with 474 monthly observations from Jan, 1956 to June, 1995.

2.10 Results

We present results for the original data set and also a standardized version. The standardized set is considered because of the possible confounding problem of serial correlation and correlation between variables.

We standardized the price data using the following formulas:

$$\begin{aligned} \underline{X}_t - \underline{X}_{t-1} &= \underline{\epsilon}_t, \quad \text{if prices follow a RW} && (2.16) \\ Var(\underline{\epsilon}_t) &= \Sigma && (2.17) \\ \Sigma &= AA' && (2.18) \\ \widetilde{\underline{X}}_t &= A^{-1}\underline{X}_t && (2.19) \end{aligned}$$

where $\underline{X}_t$ is the vector of pries at time t, $\underline{\epsilon}_t$ is the error if prices follow a RW and Σ is the variance-covariance matrix of $\underline{\epsilon}_t$.

The original industry data is highly correlated, to lessen this problem we can simply form a new standardized version of $\underline{X}_t$ which is given by $\widetilde{\underline{X}}_t$. As can be seen by Equation (2.18) the variance of the original industry data can be broken into two parts , A and A'. When A^{-1} premultiplies the original data the new variable $\widetilde{\underline{X}}_t$ is created. The new data $\widetilde{\underline{X}}_t$ is a linear combination of the industries from the TSE, and is referred to as a portfolio. The weights of the original industries in each portfolio are given in Table 2.1.

Canonical correlation analysis seeks a linear combination of variables from one group that have maximum correlation with a linear combination of another set of variables such that each combination is independent of all others. Since standardization merely involves linear combinations of the original data this does not change the resulting canonical variates.

Table 2.1: The industry weights of each portfolio

#	Industry	P1	P2	P3	P4	P5	P6	P7
1	Communications & Media	28.09	-1.78	-0.28	-1.51	-2.20	-4.25	-0.29
2	Financial Services	-1.78	34.79	-0.03	-4.65	-1.06	-3.72	-0.69
3	Gold & Silver	-0.28	-0.03	11.83	0.40	-0.09	0.18	-2.53
4	Industrial Products	-1.51	-4.65	0.40	38.15	-1.35	-5.24	-4.40
5	Conglomerates	-2.20	-1.06	-0.09	-1.35	24.86	-1.69	-2.08
6	Merchandising	-4.25	-3.72	0.18	-5.24	-1.69	33.59	-0.88
7	Metals & Minerals	-0.29	-0.69	-2.53	-4.40	-2.08	-0.88	25.19
8	Oil & Gas	0.35	0.38	-1.60	-1.21	-2.82	-1.90	-2.26
9	Paper & Forest Products	-1.50	-2.10	-0.84	-3.90	-2.74	-1.48	-4.88
10	Pipelines	-1.04	-3.38	-0.51	-2.03	-1.12	-0.84	-1.66
11	Trans. & Environmental	-0.48	-0.87	-0.21	-2.76	-2.48	-0.58	-1.37
12	Utilities	-2.89	-6.04	-0.20	-2.00	-0.34	-0.44	1.67
13	Consumer Products	-4.20	-3.55	0.18	-5.05	-1.61	-4.20	-1.58

#	Industry	P8	P9	P10	P11	P12	P13	
1	Communications & Media	0.35	-1.50	-1.04	-0.48	-2.89	-4.20	
2	Financial Services	0.38	-2.10	-3.38	-0.87	-6.04	-3.55	
3	Gold & Silver	-1.60	-0.84	-0.51	-0.21	-0.20	0.18	
4	Industrial Products	-1.21	-3.90	-2.03	-2.76	-2.00	-5.05	
5	Conglomerates	-2.82	-2.74	-1.12	-2.48	-0.34	-1.61	
6	Merchandising	-1.90	-1.48	-0.84	-0.58	-0.44	-4.20	
7	Metals & Minerals	-2.26	-4.88	-1.66	-1.37	1.67	-1.58	
8	Oil and Gas	20.57	0.61	-5.11	-1.02	0.01	-2.64	
9	Paper & Forest Products	0.61	25.29	0.46	-2.33	0.77	-3.11	
10	Pipelines	-5.11	0.46	27.94	-1.22	-3.90	0.07	
11	Trans. & Environmental	-1.02	-2.33	-1.22	20.76	-1.29	-1.52	
12	Utilities	0.01	0.77	-3.90	-1.29	42.15	-4.46	
13	Consumer Products	-2.64	-3.11	0.07	-1.52	-4.46	40.37	

where P indicates a standardized portfolio, for example P1 is standardized portfolio 1. Note that portfolio i is primarily composed of industry i. The negative weights for many industries may be taken to represent certain buying or selling strategys such as selling short.

Notation used in the results section:

k – refers to the dimension of the price vector, which in our case is 13.

vec($\cdot$) – this is the column stacking operator.

∇ – extracts the upper diagonal portion of the matrix it is applied to (this is needed for any matrix of the form $\underline{X}_t\underline{X}_t^{'}$ since it will be symmetric and when the vec($\cdot$) operator is used on it, $k^2 - \frac{k(k+1)}{2}$ elements in this new vector will be identical).

$\%R^2$ – this is simply $100\times R^2$.

R^2_w – this is a weighted average R^2, with weights given by the percentage of TSE market value for each industry as of June, 1995 (see Table 1.1). To calculate this statistic each industry R^2 is multiplied by its share of TSE market value given in Table 1.1 and the resulting values are added together. This statistic will be used for the original data set.

R^2_a – this is an unweighted average R^2. This statistic will be used for the standardized data where no direct correspondence with an industry is possible (each portfolio is a linear combination of industries).

B – this is the usual backshift operator, i.e $BY_t = Y_t - Y_{t-1}$. This operator is used in the ARIMA models in the Appendix.

Note: All Tables in the results section with an A preceeding a number are located in the Appendix. The critical values for the Dickey-Fuller unit root test statistic are -10.4 and 1.61. If the Dickey-Fuller statistic is between these two values the series is classified as nonstationary.

2.10.1 Results using a linear lag-1 prediction filter

The linear lag-1 predictor was obtained using the following regression model:

$$\underbrace{\underline{X}_t}_{(k\times 1)} = \underbrace{A}_{(k\times k)}\underline{X}_{t-1} + \underline{\epsilon}_t, \quad t = 2, 3, \ldots, 474. \tag{2.20}$$

Here $\underline{X}_t$ is the vector of stock prices at time t, A is a constant coefficient matrix, and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}_t}$ is obtained by estimating A and setting $\underline{\epsilon}_t$ to zero in (2.20).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}_t}$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2,...,13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series, and the Q(12)-statistics of Ljung-Box for residuals of the fitted model are given in Table A.1.

Referring to Table A.1 we can see that all of the transformed series are AR(1)'s, except series 1, 2 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.2. **Original Data:** The R^2's range from a low of 0.50% for industry 1 and 2 to a high of 9.49% for industry 7. R^2_w is 3.76% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.63% for portfolio 1 to a high of 61.66% for portfolio 10. R^2_a is 17.57% which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.10.2 Results using a linear lag-2 prediction filter

The linear predictor of lag-2 was obtained using the following regression model:

$$\underbrace{\underline{X}_t}_{(k\times 1)} = \underbrace{A}_{(k\times k)}\underline{X}_{t-1} + \underbrace{B}_{(k\times k)}\underline{X}_{t-2} + \underline{\epsilon}_t, \quad t = 3, 4, \ldots, 474. \tag{2.21}$$

Here $\underline{X}_t$ is the vector of stock prices at time t, $\underline{X}_{t-1}$ is the vector of stock prices at time $t-1$, $\underline{X}_{t-2}$ is the vector of stock prices at time $t-2$, A and B are constant coefficient matrices, and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}}_t$ is obtained by estimating A and B and setting $\underline{\epsilon}_t$ to zero in (2.21).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}}_t$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2...13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series, and the Q(12)-statistics of Ljung-Box for residuals of the fitted model are given in Table A.3.

Referring to Table A.3 we can see that all of the transformed series are AR(1)'s, except series 1, 2 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.4. **Original Data:** The R^2's range from a low of 0.48% for industry 1 to a high of 9.52% for industry 7. R^2_w is 3.77% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.60% for portfolio 1 to a high of 61.06% for portfolio 10 . R^2_a is 17.55% which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.10.3 Results using a quadratic lag-1 prediction filter

The quadratic predictor of lag-1 was obtained using the following regression model:

$$\underbrace{\underline{X}_t}_{(k\times 1)} = \underbrace{A}_{(k\times k)}\underline{X}_{t-1} + \underbrace{B}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla \underline{X}_{t-1}\underline{X}_{t-1}^{'}) + \underline{\epsilon}_t, \quad t = 2,3,\ldots,474. \tag{2.22}$$

Here $\underline{X}_t$ is the vector of prices at time t, $\underline{X}_{t-1}$ is the vector of prices at time t-1, A and B are constant coefficient matrices, vec$(\cdot)$ is the column stacking operator, ∇ extracts the upper diagonal portion of the matrix it is applied to (this is needed because the matrix $\underline{X}_{t-1}\underline{X}_{t-1}^{'}$ is symmetric and we only want the unique elements), and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}}_t$ is obtained by estimating A and B and setting $\underline{\epsilon}_t$ to zero in (2.22).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}}_t$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2...13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series, and the Q(12)-statistics of Ljung-Box for residuals of the fitted model are given in Table A.5.

Referring to Table A.5 we can see that all of the transformed series are AR(1)'s, except series 1, 2 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.6. **Original Data:** The R^2's range from a low of 0.43% for industry 6 to a high of 9.43% for industry 7. R_w^2 is 3.74% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.65% for portfolio 1 to a high of 64.22% for portfolio 10 . R_a^2 is 17.81% which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.10.4 Results using a quadratic lag-2 prediction filter

The quadratic predictor of lag-2 was obtained using the following regression model:

$$\begin{aligned}\underbrace{\underline{X}_t}_{(k\times 1)} &= \underbrace{A}_{(k\times k)}\underline{X}_{t-1} + \underbrace{B}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla \underline{X}_{t-1}\underline{X}'_{t-1}) + \underbrace{C}_{(k\times k^2)}\mathrm{vec}(\nabla \underline{X}_{t-1}\underline{X}'_{t-2}) + \\ &\quad \underbrace{D}_{(k\times k)}\underline{X}_{t-2} + \underbrace{E}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla \underline{X}_{t-2}\underline{X}'_{t-2}) + \underline{\epsilon}_t \quad t = 3, 4, \ldots 474 \qquad (2.23)\end{aligned}$$

Here $\underline{X}_t$ is the vector of prices at time t, $\underline{X}_{t-1}$ is the vector of prices at time $t-1$, $\underline{X}_{t-2}$ is the vector of prices at time $t-2$, A, B, C, D and E are constant coefficient matrices and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}_t}$ is obtained by estimating A, B, C, D and E and setting $\underline{\epsilon}_t$ to zero in (2.23).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}_t}$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2...13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series and the Q(12)-statistics of Ljung-Box for the residuals of the fitted model are given in Table A.7.

Referring to Table A.7 we can see that all of the transformed series are AR(1)'s, except series 1, 2,3 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.8. **Original Data:** The R^2's range from a low of 0.51% for industry 6 to a high of 9.84% for industry 7. R^2_w is 4.46% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.99% for portfolio 1 to a high of 63.36% for portfolio 10 . R^2_a is 18.68% which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.10.5 Results using a square Root lag-1 prediction filter

The square-root predictor of lag-1 was obtained using the following regression model:

$$\underbrace{\underline{X}_t}_{(k\times 1)} = \underbrace{A}_{(k\times k)} \underline{X}_{t-1} + \underbrace{B}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla\sqrt{\underline{X}_{t-1}\underline{X}'_{t-1}}) + \underline{\epsilon}_t, \quad t = 2, 3, \ldots, 474. \qquad (2.24)$$

Here $\underline{X}_t$ is the vector of prices at time t, $\underline{X}_{t-1}$ is the vector of prices at time $t-1$, A and B are coefficient matrices and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}}_t$ is obtained by estimating A and B and setting $\underline{\epsilon}_t$ to zero in (2.24).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}}_t$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2...13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series, and the Q(12)-statistics of Ljung-Box for the residuals of the fitted model are given in Table A.9.

Referring to Table A.9 we can see that all of the transformed series are AR(1)'s, except series 1,2 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.10. **Original Data:** The R^2's range from a low of 0.43% for industry 6 to a high of 9.50% for industry 7. R^2_w is 3.75% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.64% for portfolio 1 to a high of 64.32% for portfolio 10 . R^2_a is 17.80% which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.10.6 Results using a square Root lag-2 prediction filter

The square-root predictor of lag-2 was obtained using the following regression model:

$$\begin{aligned}\underbrace{\underline{X}_t}_{(k\times 1)} &= \underbrace{A}_{(k\times k)}\underline{X}_{t-1} + \underbrace{B}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla\sqrt{\underline{X}_{t-1}\underline{X}'_{t-1}}) + \underbrace{C}_{(k\times k^2)}\mathrm{vec}(\nabla\sqrt{\underline{X}_{t-1}\underline{X}'_{t-2}}) + \\ &\quad \underbrace{D}_{(k\times k)}\underline{X}_{t-2} + \underbrace{E}_{(k\times \frac{k(k+1)}{2})}\mathrm{vec}(\nabla\sqrt{\underline{X}_{t-2}\underline{X}'_{t-2}}) + \underline{\epsilon}_t \quad t = 3,4,\ldots,474 \qquad (2.25)\end{aligned}$$

Here $\underline{X}_t$ is the vector of prices at time t, $\underline{X}_{t-1}$ is the vector of prices at time $t-1$, $\underline{X}_{t-2}$ is the vector of prices at time $t-2$, and A, B, C, D and E are constant coefficient matrices, vec$(\cdot)$ is the column stacking operator, ∇ extracts the upper diagonal portion of the matrix it is applied to and $\{\underline{\epsilon}_t\}$ is vector white noise. Our predictor $\widehat{\underline{X}}_t$ is obtained by estimating A, B, C, D and E and setting $\underline{\epsilon}_t$ to zero in (2.25).

A canonical correlation analysis between $\underline{X}_t$ and $\widehat{\underline{X}}_t$ was performed, and this resulted in 13 transformed series, $\{y_{jt}\}$ j=1,2...13. The ARIMA models for the 13 transformed $\{y_{jt}\}$ series, and the Q(12)-statistics of Ljung-Box for the residuals of the fitted model are given in Table A.11.

Referring to Table A.11 we can see that all of the transformed series are AR(1)'s, except series 1, 2 and 4. The Dickey-Fuller unit root test was performed on the transformed series and indicated that the first 5 series were nonstationary.

Carrying out the regressions given in Definition (2.9) with the number of nonstationary series equal to 5, we obtained the results summarized in Table A.12. **Original Data:** The R^2's range from a low of 0.49% for industry 6 to a high of 9.77% for industry 7. R^2_w is 4.38% which indicates that very little variation in industry prices is accounted for by stationary components. **Standardized data:** The R^2's range from a low of 0.75% for portfolio 1 to a high of 65.43% for portfolio 10 . R^2_a is 18.58 which indicates that a large amount of variation in standardized prices is accounted for by stationary components.

2.11 Testing the significance of the R^2's

The last question we consider is whether the R^2's that we have obtained are statistically different from those that we could expect to obtain under a RW model for prices. To answer this question we used bootstrapping and simulation techniques with the linear lag-1 prediction filter. In all cases the sample size was 1000. The linear lag-1 prediction filter was used because it took the least amount of computational time. See Figures A.1, A.2, A.3 and A.4 in the Appendix for plots of the empirical R^2 distribution for all cases. The y-axis of these Figures represents the range of numerical quantiles (0,1) and the x-axis represents the corresponding empirical R^2 values.

2.12 Bootstrapping

Bootstrapping is a technique used when the exact values for parameters of a distribution are unknown. It involves randomly sampling elements from the original data set. In our case we will difference the price data to obtain returns and sample randomly from this set, then reconstruct prices using the fact that $P_{t+1} = P_t + R_t$ (the initial price was chosen to accord with the original data). If the prices actually follow a RW model then this should not make a difference in the R^2 values (if prices follow a RW, returns are independent and identically distributed and their ordering should not matter). For both the original data set and the standardized data set we will generate 1000 samples and calculate the R^2's for each industry or portfolio according to the outlined scheme. We will then calculate the empirical quantiles of the R^2 statistic and compare the R^2 from the real price data to the empirical quantiles. If the real price data R^2 is located outside of the empirical 95% quantile we can conclude that it is significantly different than would have occurred under a RW model for prices.

Table 2.2: R^2's under the RW Null using bootstrapping vs actual

	Empirical Quantiles				
Industry	0.85	0.90	0.95	Actual	Significant at 5% (Y/N)
Commercial	0.0197	0.0250	0.0376	0.0050	N
Financial	0.0772	0.0942	0.1444	0.0050	N
Gold	0.1185	0.1575	0.2364	0.0396	N
Industrial	0.0999	0.1409	0.2002	0.0585	N
Conglomerates	0.1411	0.1841	0.2633	0.0332	N
Merchandising	0.0451	0.0613	0.1070	0.0051	N
Metals	0.1642	0.2179	0.2984	0.0949	N
Oil	0.1460	0.1730	0.2523	0.0425	N
Paper	0.1706	0.2145	0.2927	0.0450	N
Pipeline	0.1009	0.1388	0.2197	0.0570	N
Transportation	0.1224	0.1669	0.2424	0.0196	N
Utilities	0.0929	0.1121	0.1503	0.0384	N
Consumer	0.0369	0.0475	0.0813	0.0146	N

where 0.85, 0.90 and 0.95 are the empirical R^2 quantiles based on the RW Null.

2.12.1 Bootstrapping without standardization

As Table 2.2 indicates, none of the R^2's obtained with the original data are significantly different from those generated under the hypothesis that prices follow a RW, at the 5% significance level.

Table 2.3: R^2's under the RW Null using bootstrapping vs actual

	Empirical Quantiles				
Portfolio	0.85	0.90	0.95	Actual	Significant at 5% (Y/N)
Portfolio #1	0.0285	0.0397	0.0661	0.0063	N
Portfolio #2	0.2282	0.2860	0.4228	0.1334	N
Portfolio #3	0.1421	0.1846	0.2734	0.0537	N
Portfolio #4	0.2803	0.3416	0.4487	0.2748	N
Portfolio #5	0.2372	0.2956	0.4110	0.3665	N
Portfolio #6	0.1455	0.1944	0.2727	0.0392	N
Portfolio #7	0.2476	0.3262	0.4460	0.2773	N
Portfolio #8	0.2607	0.3200	0.4276	0.1310	N
Portfolio #9	0.2430	0.3106	0.4160	0.1241	N
Portfolio #10	0.2211	0.2743	0.3967	0.6166	Y
Portfolio #11	0.2082	0.2684	0.3812	0.0429	N
Portfolio #12	0.1950	0.2306	0.2868	0.1842	N
Portfolio #13	0.0883	0.1321	0.1977	0.0340	N

where 0.85, 0.90 and 0.95 are the empirical R^2 quantiles based on the RW Null.

2.12.2 Bootstrapping with standardization

As Table 2.3 indicates, the R^2 obtained for portfolio #10 of the original standardized data is significantly different than the one generated under the hypothesis that standardized prices follow a random walk. We take this as evidence that this portfolio does not follow a RW. The other portfolio's are adequately explained by a RW model, at the 5% significance level.

2.13 Simulations

We also simulated 1000 vector random walk price data using normal errors with a mean and variance the same as the return data. The procedure used to simulate the data is briefly outlined below.

If prices follow a vector RW then the following two equations are valid for prices and returns:

$$\begin{aligned} \underline{P}_t &= \underline{\mu} + \underline{P}_{t-1} + \underline{\epsilon}_t \\ \underline{R}_t &= \underline{P}_t - \underline{P}_{t-1} = \underline{\epsilon}_t + \underline{\mu} \end{aligned}$$

If we assume $\underline{\epsilon}_t$ is distributed as a p-dimensional normal random vector, we can construct a simulated group of 13 prices which we can use to compare to the original price series. We do this by generating observations from a p-dimensional normal distribution with mean and variance given by the original return data. We then construct simulated prices by letting the initial simulated price series be the same as the original price series and then recursively generate subsequent prices using the fact that $\underline{P}_{t+1} = \underline{P}_t + \underline{R}_t$.

After each simulation has been completed it is subjected to the same canonical correlation analysis as the real data. Each simulation generates a series of R^2's. After all the simulations and canonical correlation analyses have been completed the empirical distribution of the simulated RW R^2's are compared to those obtained by the real data. If any of the real data R^2's are larger than the simulated RW R^2's we take this as evidence that the real price series does not follow a RW.

Table 2.4: R^2's under the RW Null using simulations vs actual

	Empirical Quantiles				
Industry	0.85	0.90	0.95	Actual	Significant at 5% (Y/N)
Commercial	0.0188	0.0243	0.0329	0.0050	N
Financial	0.0792	0.1035	0.1549	0.0050	N
Gold	0.1081	0.1434	0.2319	0.0396	N
Industrial	0.1006	0.1344	0.1826	0.0585	N
Conglomerates	0.1375	0.1969	0.2846	0.0332	N
Merchandising	0.0448	0.0570	0.0822	0.0051	N
Metals	0.1564	0.1952	0.2688	0.0949	N
Oil	0.1434	0.1863	0.2521	0.0425	N
Paper	0.1775	0.2326	0.3103	0.0450	N
Pipeline	0.1049	0.1572	0.2224	0.0570	N
Transportation	0.1263	0.1659	0.2279	0.0196	N
Utilities	0.0907	0.1174	0.1718	0.0384	N
Consumer	0.0352	0.0447	0.0663	0.0146	N

where 0.85, 0.90 and 0.95 are the empirical R^2 quantiles based on the RW Null.

2.13.1 Simulations with no standardization

The procedure outlined above was completed with 1000 samples and the results are presented in Table 2.4. The first column describes the series being analysed, the second through fourth columns are the upper tail quantiles of the R^2's obtained from the simulated vector random walk series, the fifth column contains the R^2 obtained with the real data and the last column indicates whether the real data R^2 is larger than the corresponding 95% quantile from the simulated vector RW series.

As Table 2.4 indicates, none of the original price data R^2's seems to be significantly different from the R^2's generated under the hypothesis that prices follow a vector RW, at the 5% significance level.

Table 2.5: R^2's under the RW Null using simulations vs actual

	Empirical Quantiles				
Portfolio	0.85	0.90	0.95	Actual	Significant at 5% (Y/N)
Portfolio #1	0.2751	0.3376	0.4566	0.0063	N
Portfolio #2	0.2726	0.3352	0.4366	0.1334	N
Portfolio #3	0.2735	0.3392	0.4425	0.0537	N
Portfolio #4	0.2735	0.3462	0.4821	0.2748	N
Portfolio #5	0.2680	0.3327	0.4350	0.3665	N
Portfolio #6	0.2488	0.3089	0.4002	0.0392	N
Portfolio #7	0.2754	0.3321	0.4669	0.2773	N
Portfolio #8	0.2708	0.3188	0.4093	0.1310	N
Portfolio #9	0.2635	0.3124	0.3945	0.1241	N
Portfolio #10	0.2739	0.3394	0.4591	0.6166	Y
Portfolio #11	0.2641	0.3242	0.4133	0.0429	N
Portfolio #12	0.2518	0.2931	0.4027	0.1842	N
Portfolio #13	0.2546	0.3247	0.4268	0.0340	N

where 0.85, 0.90 and 0.95 are the empirical R^2 quantiles based on the RW Null.

2.13.2 Simulations with standardization

All of the portfolio R^2's except one are not significantly different from those expected if prices followed a vector RW process, at the 5% significance level. However, as Table 2.5 indicates, the R^2 obtained by Portfolio #10 is significantly different than we would expect if it followed a RW process. We take this as evidence that this portfolio does not follow a RW process.

Chapter 3

Discussion

3.1 Comparison of the three predictive information filters

In this thesis we have considered three predictive information filters, linear, quadratic and square-root with lags 1 and 2, as well as standardized and unstandardized data. A summary of the performance of each predictive filter relative to the linear lag-1 predictive information filter for each data set is given in Tables 3.1 and 3.2.

Table 3.1: Summary of predictive information filter R^2's without standardization

	L 1	**L 2**		**Q 1**		**Q 2**		**SR 1**		**SR 2**	
I	$\%R^2_{l1}$	$\%R^2_{l2}$	$\%\Delta$	$\%R^2_{q1}$	$\%\Delta$	$\%R^2_{q2}$	$\%\Delta$	$\%R^2_{sr1}$	$\%\Delta$	$\%R^2_{sr2}$	$\%\Delta$
1	0.50	0.48	-4.23	0.51	2.34	0.88	74.74	0.51	2.56	0.82	63.23
2	0.50	0.52	3.94	0.58	17.42	1.07	114.45	0.59	18.33	1.02	104.60
3	3.96	3.83	-3.28	3.94	-0.42	5.34	34.87	3.94	-0.40	5.34	34.98
4	5.85	5.84	-0.25	5.86	0.06	6.79	16.05	5.88	0.37	6.69	14.27
5	3.32	3.31	-0.22	3.16	-4.83	3.61	8.86	3.15	-5.02	3.62	9.23
6	0.51	0.49	-4.02	0.43	-16.04	0.51	-0.69	0.43	-16.07	0.49	-4.01
7	9.49	9.52	0.37	9.43	-0.63	9.84	3.77	9.50	0.18	9.77	3.00
8	4.25	4.35	2.23	4.20	-1.33	3.56	-16.43	4.21	-1.03	3.51	-17.60
9	4.50	4.75	5.46	4.53	0.64	7.82	73.71	4.57	1.44	7.76	72.32
10	5.70	5.61	-1.55	6.03	5.85	5.88	3.08	6.06	6.32	5.82	2.02
11	1.96	2.04	4.06	1.68	-14.19	2.01	2.66	1.72	-12.12	2.12	8.26
12	3.84	3.79	-1.42	3.59	-6.70	4.93	28.37	3.56	-7.38	4.58	19.16
13	1.46	1.49	1.85	1.47	0.54	2.04	39.89	1.49	1.88	1.95	33.39
R^2_w	3.76	3.77		3.74		4.46		3.75		4.38	

I is industry (see Table 2.1 for a description) , L 1 is linear lag-1 prediction, L 2 is linear lag-2 prediction, Q 1 is quadratic lag-1 prediction, Q 2 is quadratic lag-2 prediction, SR 1 is Square-root lag-1 prediction, SR 2 is Square-root lag-2 prediction, $\%\Delta$ is the relative percentage change in R^2 using each of the prediction methods as compared to the R^2 obtained with linear lag-1 prediction. For example the $\%\Delta$ for Quadratic lag-1 was calculated using $\frac{R^2_{q1}-R^2_{l1}}{R^2_{l1}}$. R^2_w is the weighted R^2.

Examining Table 3.1 we make the following observations:

- The R^2's increase with lag for quadratic and square-root prediction for most industries but did not for linear prediction.

- The weighted R^2 for quadratic lag-2 prediction is the largest of any predictive filter, it is 18.6% larger than the weighted R^2 for linear lag-1 prediction.

- Based on the quadratic lag-2 predictor, the following industries have the most variance explained by stationary components: 1. Metals & Minerals 2. Paper 3. Industrial . Each of these industries has an $R^2 > 6\%$.

- Based on the quadratic lag-2 predictor, the following industries have the least amount of variance explained by stationary components: 1. Merchandising 2. Media & Communications 3. Financial. Each of these industries has an $R^2 < 2\%$.

Examining Table 3.2 we make the following observations:

- The R^2's increase with lag for quadratic and square-root prediction for most industries but the same does not hold for linear prediction.

- The average R^2 for quadratic lag-2 prediction is the largest of any predictive filter, it is 6.3% larger than linear lag-1 prediction.

- Based on the quadratic lag-2 predictor, the following portfolios have the largest amount of variance explained by stationary components: 1. portfolio 10, 2. portfolio 6, 3. portfolio 8. Each of these portfolios has an $R^2 > 30\%$.

- Based on the quadratic lag-2 predictor, the following portfolios have the least amount of variance explained by stationary components: 1. portfolio 1, 2. portfolio 11, portfolio 6. Each of these portfolios has an $R^2 < 4\%$.

Table 3.2: Summary of predictive information filter R^2's with standardization

	L 1	L 2		Q 1		Q 2		SR 1		SR 2	
P	$\%R^2_{l1}$	$\%R^2_{l2}$	$\%\Delta$	$\%R^2_{q1}$	$\%\Delta$	$\%R^2_{q2}$	$\%\Delta$	$\%R^2_{sr1}$	$\%\Delta$	$\%R^2_{sr2}$	$\%\Delta$
1	0.63	0.60	-4.55	0.65	3.41	0.99	57.54	0.64	2.10	0.75	18.82
2	13.34	13.96	4.64	11.84	-11.26	11.54	-13.50	11.56	-13.35	11.96	-10.36
3	5.37	5.18	-3.49	5.43	1.25	7.65	42.62	5.48	2.18	7.18	33.75
4	27.48	26.60	-3.21	30.93	12.57	23.03	-16.20	31.22	13.61	26.64	-3.06
5	36.65	35.84	-2.20	35.94	-1.93	41.91	14.35	35.82	-2.25	38.18	4.18
6	3.92	3.80	-3.30	3.61	-8.08	3.89	-0.76	3.48	-11.45	4.65	18.41
7	27.73	27.75	0.08	29.21	5.36	32.95	18.83	28.95	4.39	31.72	14.39
8	13.10	13.51	3.15	12.43	-5.08	8.82	-32.67	12.66	-3.34	8.44	-35.53
9	12.41	13.72	10.58	13.49	8.71	22.54	81.66	13.74	10.75	22.17	78.67
10	61.66	61.06	-0.98	64.22	4.15	63.36	2.75	64.32	4.31	65.43	6.10
11	4.29	4.45	3.65	3.83	-10.87	3.09	-28.05	3.87	-9.94	4.35	1.20
12	18.42	18.25	-0.92	16.52	-10.32	18.98	3.02	16.34	-11.32	16.47	-10.60
13	3.40	3.41	0.21	3.49	2.45	4.09	20.14	3.41	0.32	3.65	7.18
R^2_a	17.57	17.55		17.81		18.68		17.80		18.58	

P is portfolio (see Table 2.1 for a description), L 1 is linear lag-1 prediction, L 2 is linear lag-2 prediction, Q 1 is quadratic lag-1 prediction, Q 2 is quadratic lag-2 prediction, SR 1 is Square-root lag-1 prediction, SR 2 is Square-root lag-2 prediction, $\%\Delta$ is the relative percentage change in R^2 using each of the prediction methods as compared to the R^2 obtained with linear lag-1 prediction. For example the $\%\Delta$ for Quadratic lag-1 was calculated using $\frac{R^2_{q1}-R^2_{l1}}{R^2_{l1}}$. R^2_a is the average R^2 (unweighted).

3.2 The significance of the R^2's and their relevance to the EMH

Tables such as 3.1 and 3.2 alone do not address the question of whether the R^2's obtained are significantly different from those that could occur under the hypothesis that prices follow a RW. We have answered this question by using simulations and bootstrapping to construct hypothetical RW price data, which we then performed a canonical correlation analysis on. These results which are summarized in Tables 2.2, 2.3, 2.4, 2.5 indicate that for the original industries the R^2's that we have obtained can be expected under the hypothesis that prices follow a RW. However, when the data is standardized (which results in 13 portfolios) we find evidence that portfolio 10 can not be explained by a RW. Since this portfolio is a linear combination of each of the industries we view this as evidence that market prices as a whole do not follow a RW.

3.3 Why the quadratic information filter results in the highest R^2's

We believe that one of the main reasons the quadratic predictive information filters works better than the linear predictive information filters is because the data is non-normal. Linear prediction uses the first two order moments of the sampling distribution of the data to perform its prediction whereas the use of quadratic prediction allows the use of up to the fourth order sample moment. This means that quadratic prediction can more accurately reflect the empirical distribution of the data. As shown in Figures A.5, A.6 and A.7 in the Appendix, the industry data is in fact non-normal with very heavy tails.

3.4 Comparison with Tsay's (1990) results

The linear lag-1 predictive information filter is essentially the same as developed by Tsay (1990). In fact, the results obtained by both of these methods are identical. Using this predictor implies an implicit belief that today's prices contain all of the available information on tomorrow's prices. Since he also uses linear prediction, he would like the data to be normally distributed. We have been able to show that both of these assumptions are in error. On the basis of this we used non-linear predictive filters and obtained R^2's that are higher than with Tsay's method for most portfolios (10 out of 13 with the original data and 8 out of 13 with the standardized data). This meets the main objective of this thesis, namely to to be able to account for a larger percentage of stock price variations using stationary components than previously possible. On average our quadratic lag-2 predictor beat Tsay's linear lag-1 predictor by 18.6% with the original data and 6.3% with the standardized data.

3.5 Future Developments

The choice of our predictive information filter depends on the type of filter and also on the number of time lags used. A stopping criteria for the number of lags used would be helpful. One possible solution is to calculate the following statistic of weighted R^2's for each lag used,

$$\sum_{i=1}^{k} W_i R_i^2 = R_l^2 \tag{3.26}$$

where W_i is a weight that could reflect the portfolios market value relative to the total, k is the number of portfolios used in the analysis and l is the lag used by the predictive filter.

When and if R_l^2 does not change much or reaches a maximum, the lag where this occurs is the best lag to choose.

We decided to stop at lag-2 because of two reasons. The first was that at lags greater than two computational difficulties begin to arise for the quadratic and square-root predictive filters. The design matrix becomes singular and prevents the regression from being done, this is a direct result of a very large number of regressors (to be exact there are $3(13+13^2+13(13+1)) = 1092$). The second reason was that we felt two months was enough time for the market to absorb any information released by companies in their quarterly report.

A second consideration is that we have only searched the class of linear, quadratic and square-root filters. It is very possible that a better filter exists which we did not investigate, such as $E(\underline{X}_t|\underline{X}_{t-1}, \underline{X}_{t-2})$. Further studies may be to implement some parametric, semiparametric or non-parametric numerical schemes for approximating

$E(\underline{X}_t|\underline{X}_{t-1}, \underline{X}_{t-2})$.

3.6 Conclusions

In this thesis we have considered three different predictive information filters. In the usual canonical correlation analysis setting only the linear lag-1 filter has been extensively used. Based on Tables 3.1 and 3.2 it is apparent that using different filters can substantially increase the R^2's that we have considered in this thesis. In particular, the quadratic lag-2 filter resulted in a higher average R^2 for both data sets that we considered. This indicates that filtering may lead to better results in the general canonical correlation setting. We have not derived any formal results to support this conclusion, but based on our empirical results we believe that filtering may play an important role in canonical correlation analysis.

At the individual industry level we can not reject the EMH view that prices follow a RW. Some industries such as metals & minerals and paper & forest products indicate that stationary components in their prices can account for around 8% of their total variation in prices. However, when compared to the results from bootstrapping and simulations it indicates that such results are expected from a RW model for prices.

When we consider portfolios of stocks, we find that portfolio ten has an R^2 which lies beyond any of the 95% R^2 quantiles calculated via bootstrapping or simulation. Since this portfolio is a linear combination of stocks from each of the industries, we take this as evidence that market prices as a whole may not be adequately be described as a RW. Our evidence is not overwhelmingly strong however, considering that the other 12 portfolios did not show any departure from the results obtained under the hypothesis that prices follow a RW.

If the market is inefficient, than many new and creative models for prices are possible. Some authors (for example Arrow (1982)) have suggested incorporating psychological models of "irrational decision making" to try to explain the behavior of speculative asset prices. Such models have in the past received little attention because they are based on an inefficient market. However, with the recent doubt cast on the EMH such models should receive more attention in the future.

3.7 Final Comments

The debate surrounding the validity of the EMH is far from over. It is very interesting to note that many people vehemently believed the EMH was true before Summers (1986) proposed his "fads" model for prices . Jensen (1968) referred to the EMH as "best established empirical fact in economics" [10] and Keane (1980) in his monograph "The Efficient Market Hypothesis" [12] said:

> It is perhaps unfortunate that the efficient markets phenomenon should continue conventionally to be described as a "hypothesis", as if it were little more than academic speculation, when the fact is that it is a proposition which has received significant support of two decades of elaborate and rigorous testing

Currently, many researchers strongly believe that the EMH hypothesis is not valid. The *Wall Street Journal* (October 23, 1987) called the EMH "the most remarkable error in the history of economic theory". The moral of the story is that one can never be be too sure about anything. The French philosopher Voltaire realized this two hundred years ago (referring of course to a different subject than statistics) when he wrote:

> It is only charlatans who are certain.

Bibliography

[1] Blanchard, W., (1993), "Forecasting value-weighted real returns of TSE portfolios using dividend yields", UBC Press.

[2] Box, G.E.P. and Tiao, G.C., (1977), "A canonical analysis of multiple time series", *Biometrika*, **64**, 355-365.

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

[4] Daniel, K., and Tourous, W., (1991), "Common Stock Returns and the Business Cycle", *Working Paper, UBC.*

[5] Dickey, D.A. and Fuller, W.A., (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root", *JASA*, **74**, 427-431.

[6] Eckbo, E. and Liu, J., (1993), "Temporary Components of Stock Prices: New Univariate Results", *Journal of Financial and Quantitative Analysis*, **28**, 161-176.

[7] Fama, E.F., (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, **25**, 383-416.

[8] Fama, E.F. and French, K.R., (1988), "Permanent and Temporary Components of Stock Prices", *Journal of Politcal Economy*, **96**, 246-273.

[9] Fuller, W., (1976) *Introduction to Statistical Time Series*, 1st Ed., John Wiley & Sons, New York.

[10] Jensen, M., (1968), "The performance of Mutual Funds in the Period 1945-64", *Journal of Finance*, **23**, 389-416.

[11] Johnson, R.A. and Wichern, D.W., (1992), *Applied Multivariate Statistical analysis*, 3rd Ed., Prentice–Hall, New Jersey.

[12] Keane, S., (1980), *The Efficient Market Hypothesis*, Gee and Co, Oxon.

[13] Keynes, J.M., (1936), *The General Theory of Employment, Interest, and Money*, Harcourt, New York.

[14] Kim, M., Nelson, C., and Startz, R., (1991), "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence", *Review of Economic Studies*, **58**, 515-528.

[15] Lehmann, E.L., (1975), *Nonparametrics: Statistical methods based on ranks*, Holden–Day, San Fransico.

[16] Ljung, G.M. and Box, G.E.P., (1978), "On a Measure of Lack of Fit in Time Series Models", *Biometrika*, No. 2, 297-303.

[17] Phillips, P.C.B., (1987), "Time Series Regression With a Unit Root", *Econometrica*, **55**, No. 2, 277-301.

[18] Richardson, M., (1993), "Temporary Components of Stock Prices: A skeptic's View", *Journal of Business and Economic Statistics*, **11**, No. 2, 199-207.

[19] Summers, L.H., (1986), "Does the Stock Market Rationally Reflect Fundamental Values?", *Journal of Finance*, **41**, 591-601.

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

[21] Tsay, R.S., (1990), "Correlation Transformation and Components of Stock Prices", *University of Chicago technical report.*

Appendix A

TSE Index Formula and Rules [1]

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, $\ldots P_N$: the current board–lot market prices of each stock in the index.

Q_A,Q_B, $\ldots Q_N$: 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.

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 × 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:

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

$$INDEX = 1091.55$$

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$\frac{3}{8}$ 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.

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.

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.

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.

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.

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.

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.

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.

Results using linear lag-1 prediction

Table A.1: ARIMA models using linear lag-1 prediction

Model for Transformed series	Q(12)	DF
$(1-.104B+.115B^2)(1-B)y_{1,t}=\epsilon_{1,t}$	17.0	0.2
$(1-.913B)(1-B)y_{2,t}=(1+.84B)\epsilon_{2,t}$	12.3	-1.0
(1-.993B)$y_{3,t}$=$\epsilon_{3,t}$	3.9	-2.9
(1-.12B)(1-B)$y_{4,t}$=$\epsilon_{4,t}$	14.9	-8.5
(1-.985B)$y_{5,t}$=$\epsilon_{5,t}$	29.7	-7.0
(1-.977B)$y_{6,t}$=$\epsilon_{6,t}$	18.7	-10.9
(1-.967B)$y_{7,t}$=$\epsilon_{7,t}$	9.7	-15.2
(1-.952B)$y_{8,t}$=$\epsilon_{8,t}$	16.7	-22.7
(1-.944B)$y_{9,t}$=$\epsilon_{9,t}$	14.4	-26.0
(1-.928B)$y_{10,t}$=$\epsilon_{10,t}$	10.8	-33.6
(1-.913B)$y_{11,t}$=$\epsilon_{11,t}$	9.8	-41.0
(1-.871B)$y_{12,t}$=$\epsilon_{12,t}$	10.5	-60.7
(1-.854B)$y_{13,t}$=$\epsilon_{13,t}$	12.1	-68.8

All parameters are greater than their two standard errors. Series 5 has a significant Q(12) statistic. A runs test was performed on the residuals of series 5 and indicated no serial dependence.

Table A.2: Portfolio R^2's using linear lag-1 prediction

Portfolio	$\%R^2$ O	$\%R^2$ S	Portfolio	$\%R^2$ O	$\%R^2$ S
1	0.50	0.63	8	4.25	13.10
2	0.50	13.34	9	4.50	12.41
3	3.95	5.37	10	5.70	61.66
4	5.85	27.48	11	1.96	4.29
5	3.31	36.65	12	3.84	18.42
6	0.51	3.92	13	1.46	3.40
7	9.49	27.73			

where O is the original data and S is the standardized data. Portfolio i for the original data corresponds to industry i.

Results using linear lag-2 prediction

Table A.3: ARIMA models using linear lag-2 prediction

Model for Transformed Series	Q(12)	DF
$(1+.475B)(1-B)y_{1,t}=(1-.60B)\epsilon_{1,t}$	19.8	0.2
$(1-.911B)(1-B)y_{2,t}=(1+.83B)\epsilon_{2,t}$	11.6	-1.1
$(1\text{-}.993\text{B})y_{3,t}{=}\epsilon_{3,t}$	3.5	-3.1
$(1\text{-}.12\text{B})(1\text{-B})y_{4,t}{=}\epsilon_{4,t}$	13.2	-7.8
$(1\text{-}.984\text{B})y_{5,t}{=}\epsilon_{5,t}$	28.2	-7.1
$(1\text{-}.976\text{B})y_{6,t}{=}\epsilon_{6,t}$	18.4	-10.9
$(1\text{-}.967\text{B})y_{7,t}{=}\epsilon_{7,t}$	9.3	-15.4
$(1\text{-}.953\text{B})y_{8,t}{=}\epsilon_{8,t}$	16.2	-21.7
$(1\text{-}.945\text{B})y_{9,t}{=}\epsilon_{9,t}$	13.6	-25.7
$(1\text{-}.931\text{B})y_{10,t}{=}\epsilon_{10,t}$	10.0	-32.4
$(1\text{-}.913\text{B})y_{11,t}{=}\epsilon_{11,t}$	10.1	-40.8
$(1\text{-}.863\text{B})y_{12,t}{=}\epsilon_{12,t}$	12.0	-64.5
$(1\text{-}.852\text{B})y_{13,t}{=}\epsilon_{13,t}$	11.55	-69.4

All parameters are greater than their two standard errors. Series 1 and 5 have a significant Q(12) statistic. A runs test was performed on the residuals of series 1 and 5 and indicated no serial dependence.

Table A.4: Portfolio R^2's using linear lag-2 prediction

Portfolio	**%R^2 O**	**%R^2 S**	**Portfolio**	**%R^2 O**	**%R^2 S**
1	0.48	0.60	8	4.35	13.50
2	0.52	13.96	9	4.74	13.72
3	3.83	5.18	10	5.61	61.06
4	5.84	26.60	11	2.04	4.45
5	3.30	35.84	12	3.79	18.25
6	0.49	3.80	13	1.49	3.41
7	9.52	27.75			

where O is the original data and S is the standardized data.

Results using quadratic lag-1 prediction

Table A.5: ARIMA models using quadratic lag-1 prediction

Model for Transformed Series	Q(12)	DF
$(1-.10B+.11B^2)(1-B)y_{1,t}=\epsilon_{1,t}$	17.1	0.2
$(1-B^2)y_{2,t}=(1+.96B)\epsilon_{2,t}$	16.1	-1.1
(1-.994B)$y_{3,t}$=$\epsilon_{3,t}$	3.4	-2.8
(1-.040B)(1-B)$y_{4,t}=(1-.08B)\epsilon_{4,t}$	14.7	-8.7
(1-.984B)$y_{5,t}$=$\epsilon_{5,t}$	29.5	-7.4
(1-.977B)$y_{6,t}$=$\epsilon_{6,t}$	15.4	-11.0
(1-.967B)$y_{7,t}$=$\epsilon_{7,t}$	10.4	-15.4
(1-.952B)$y_{8,t}$=$\epsilon_{8,t}$	12.9	-22.6
(1-.945B)$y_{9,t}$=$\epsilon_{9,t}$	12.8	-25.6
(1-.922B)$y_{10,t}$=$\epsilon_{10,t}$	7.1	-36.7
(1-.919B)$y_{11,t}$=$\epsilon_{11,t}$	5.4	-38.4
(1-.869B)$y_{12,t}$=$\epsilon_{12,t}$	9.6	-61.9
(1-.859B)$y_{13,t}$=$\epsilon_{13,t}$	12.6	-66.5

All parameters are greater than their two standard errors. The Q(12) statistic for series 5 is significant at the 5% level. A runs test was performed on the residuals of series 5 and indicated no serial dependence.

Table A.6: Portfolio industry R^2's using quadratic lag-1 prediction

Portfolio	**%R^2 O**	**%R^2 S**	**Portfolio**	**%R^2 O**	**%R^2 S**
1	0.51	0.65	8	4.20	12.43
2	0.58	11.84	9	4.53	13.49
3	3.94	5.43	10	6.03	64.22
4	5.86	30.93	11	1.68	3.83
5	3.16	35.94	12	3.59	16.52
6	0.43	3.61	13	1.47	3.49
7	9.43	29.21			

where O is the original data and S is the standardized data.

Results using quadratic lag-2 prediction

Table A.7: ARIMA models using quadratic lag-2 prediction

Model for Transformed Series	Q(12)	DF
(1-.999B)$y_{1,t}$=$\epsilon_{1,t}$	21.2	0.2
(1-.876B)(1-B)$y_{2,t}$=$(1+.77B)\epsilon_{2,t}$	6.62	-1.2
(1+.002B)(1-B)$y_{3,t}$=$\epsilon_{3,t}$	2.6	-2.9
(1-.156B)(1-B)$y_{4,t}$=$\epsilon_{4,t}$	13.1	-10.0
(1-.986B)$y_{5,t}$=$\epsilon_{5,t}$	15.3	-6.3
(1-.977B)$y_{6,t}$=$\epsilon_{6,t}$	15.3	-10.8
(1-.962B)$y_{7,t}$=$\epsilon_{7,t}$	7.9	-17.9
(1-.958B)$y_{8,t}$=$\epsilon_{8,t}$	18.2	-20.0
(1-.942B)$y_{9,t}$=$\epsilon_{9,t}$	16.2	-27.5
(1-.925B)$y_{10,t}$=$\epsilon_{10,t}$	15.9	-34.9
(1-.858B)$y_{11,t}$=$\epsilon_{11,t}$	9.9	-66.6
(1-.909B)$y_{12,t}$=$\epsilon_{12,t}$	8.6	-42.8
(1-.875B)$y_{13,t}$=$\epsilon_{13,t}$	10.2	-58.8

All parameters are greater than their two standard errors. The Q(12) statistics for series 1 is significant at the 5% level. A runs test was performed on the residuals of series 1 and indicated no serial dependence.

Table A.8: Portfolio industry R^2's using quadratic lag-2 prediction

Portfolio	**%R^2 O**	**%R^2 S**	**Portfolio**	**%R^2 O**	**%R^2 S**
1	0.88	0.99	8	3.56	8.82
2	1.07	11.54	9	7.82	22.54
3	5.34	7.65	10	5.88	63.36
4	6.79	23.03	11	2.00	3.09
5	3.61	41.91	12	4.93	18.98
6	0.51	3.89	13	2.04	4.09
7	9.84	32.95			

where O is the original data and S is the standardized data.

Results using square-root lag-1 prediction

Table A.9: ARIMA models using square-root lag-1 prediction

Model for Transformed Series	Q(12)	DF
$(1-.103B+.110B^2)(1-B)y_{1,t}=\epsilon_{1,t}$	17.3	0.2
$(1\text{-}.909\text{B})(1\text{-B})y_{2,t} = (1+.827B)\epsilon_{2,t}$	10.4	-1.1
$(1\text{-}.994\text{B})y_{3,t}=\epsilon_{3,t}$	3.1	-2.8
$(1\text{-}.12\text{B})(1\text{-B})y_{4,t}=\epsilon_{4,t}$	14.8	-8.7
$(1\text{-}.984\text{B})y_{5,t}=\epsilon_{5,t}$	29.6	-7.4
$(1\text{-}.976\text{B})y_{6,t}=\epsilon_{6,t}$	15.4	-11.0
$(1\text{-}.967\text{B})y_{7,t}=\epsilon_{7,t}$	10.7	-15.5
$(1\text{-}.952\text{B})y_{8,t}=\epsilon_{8,t}$	16.1	-22.6
$(1\text{-}.946\text{B})y_{9,t}=\epsilon_{9,t}$	12.7	-25.6
$(1\text{-}.923\text{B})y_{10,t}=\epsilon_{10,t}$	7.1	-36.6
$(1\text{-}.918\text{B})y_{11,t}=\epsilon_{11,t}$	5.6	-38.7
$(1\text{-}.869\text{B})y_{12,t}=\epsilon_{12,t}$	9.8	-61.7
$(1\text{-}.859\text{B})y_{13,t}=\epsilon_{13,t}$	12.8	-66.5

All parameters are greater than their two standard errors. The Q(12) statistic for series 5 is significant at the 5% level. A runs test was performed on the residuals of series 5 and indicated no serial dependence.

Table A.10: Portfolio industry R^2's using square-root lag-1 prediction

Portfolio	%R^2 **O**	%R^2 **S**	**Portfolio**	%R^2 **O**	%R^2 **S**
1	0.51	0.64	8	4.21	12.66
2	0.59	11.56	9	4.57	13.74
3	3.94	5.48	10	6.06	64.32
4	5.88	31.22	11	1.72	3.87
5	3.15	35.82	12	3.56	16.34
6	0.43	3.48	13	1.49	3.41
7	9.50	28.95			

where O is the original data and S is the standardized data.

Results using square-root lag-2 prediction

Table A.11: ARIMA models using square-root lag-2 prediction

Model for Transformed Series	Q(12)	DF
(1-B)$y_{1,t}$=(1 − .05B)$\epsilon_{1,t}$	18.4	0.2
(1-.883B)(1-B)$y_{2,t}$=(1 + .78B)$\epsilon_{2,t}$	7.1	-1.2
(1-.993B)$y_{3,t}$=$\epsilon_{3,t}$	2.5	-2.9
(1-.158B)(1-B)$y_{4,t}$=$\epsilon_{4,t}$	13.8	-10.0
(1-.986B)$y_{5,t}$=$\epsilon_{5,t}$	15.8	-6.4
(1-.976B)$y_{6,t}$=$\epsilon_{6,t}$	13.1	-10.9
(1-.961B)$y_{7,t}$=$\epsilon_{7,t}$	7.9	-18.6
(1-.958B)$y_{8,t}$=$\epsilon_{8,t}$	19.7	-19.6
(1-.943B)$y_{9,t}$=$\epsilon_{9,t}$	15.8	-27.0
(1-.922B)$y_{10,t}$=$\epsilon_{10,t}$	15.9	-36.6
(1-.864B)$y_{11,t}$=$\epsilon_{11,t}$	8.4	-64.0
(1-.906B)$y_{12,t}$=$\epsilon_{12,t}$	8.8	-44.3
(1-.876B)$y_{13,t}$=$\epsilon_{13,t}$	10.3	-58.3

All parameters are greater than their two standard errors. The Q(12) statistic for series 5 is significant at the 5% level. A runs test was performed on the residuals of series 5 and indicated no serial dependence.

Table A.12: Portfolio industry R^2's using square-root lag-2 prediction

Portfolio	%R^2 O	%R^2 S	**Portfolio**	%R^2 O	%R^2 S
1	0.82	0.75	8	3.51	8.42
2	1.02	11.96	9	7.76	22.17
3	5.34	7.18	10	5.81	65.43
4	6.69	26.64	11	2.12	4.34
5	3.62	38.18	12	4.58	16.47
6	0.49	4.65	13	1.95	3.65
7	9.77	31.72			

where O is the original data and S is the standardized data.

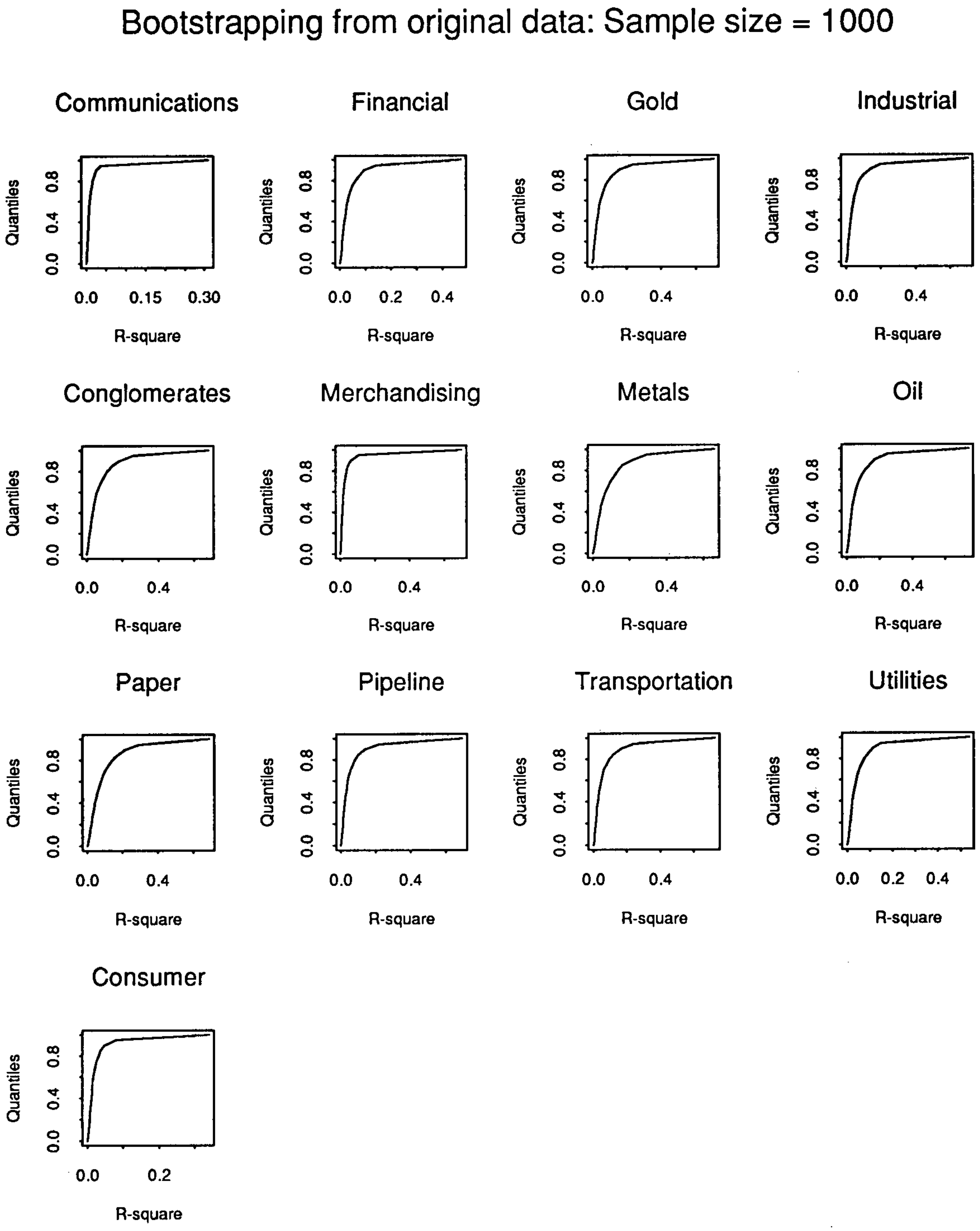


Figure A.1: Quantiles of R^2 using bootstrapping with original data.

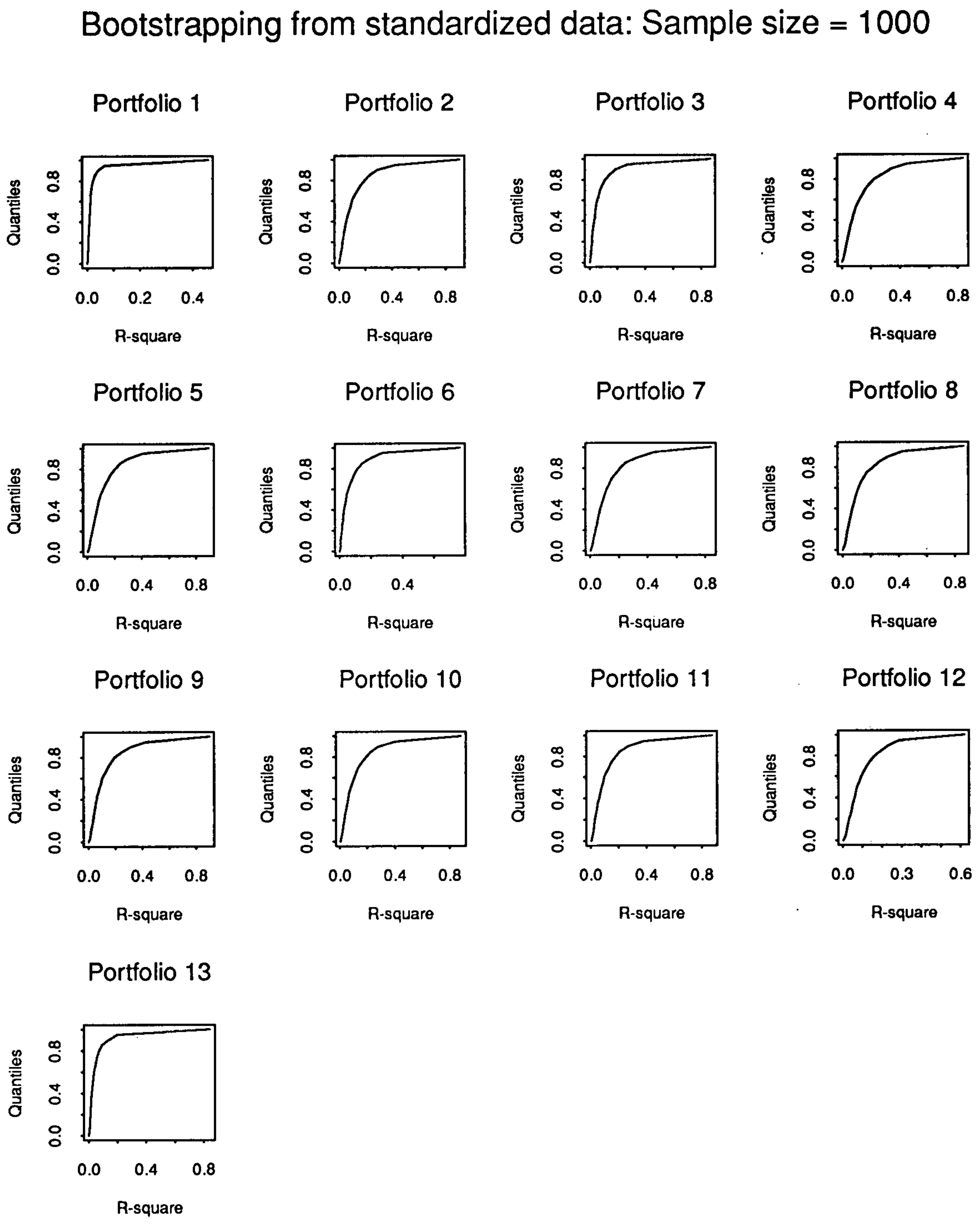


Figure A.2: Quantiles of R^2 using bootstrapping with standardized data.

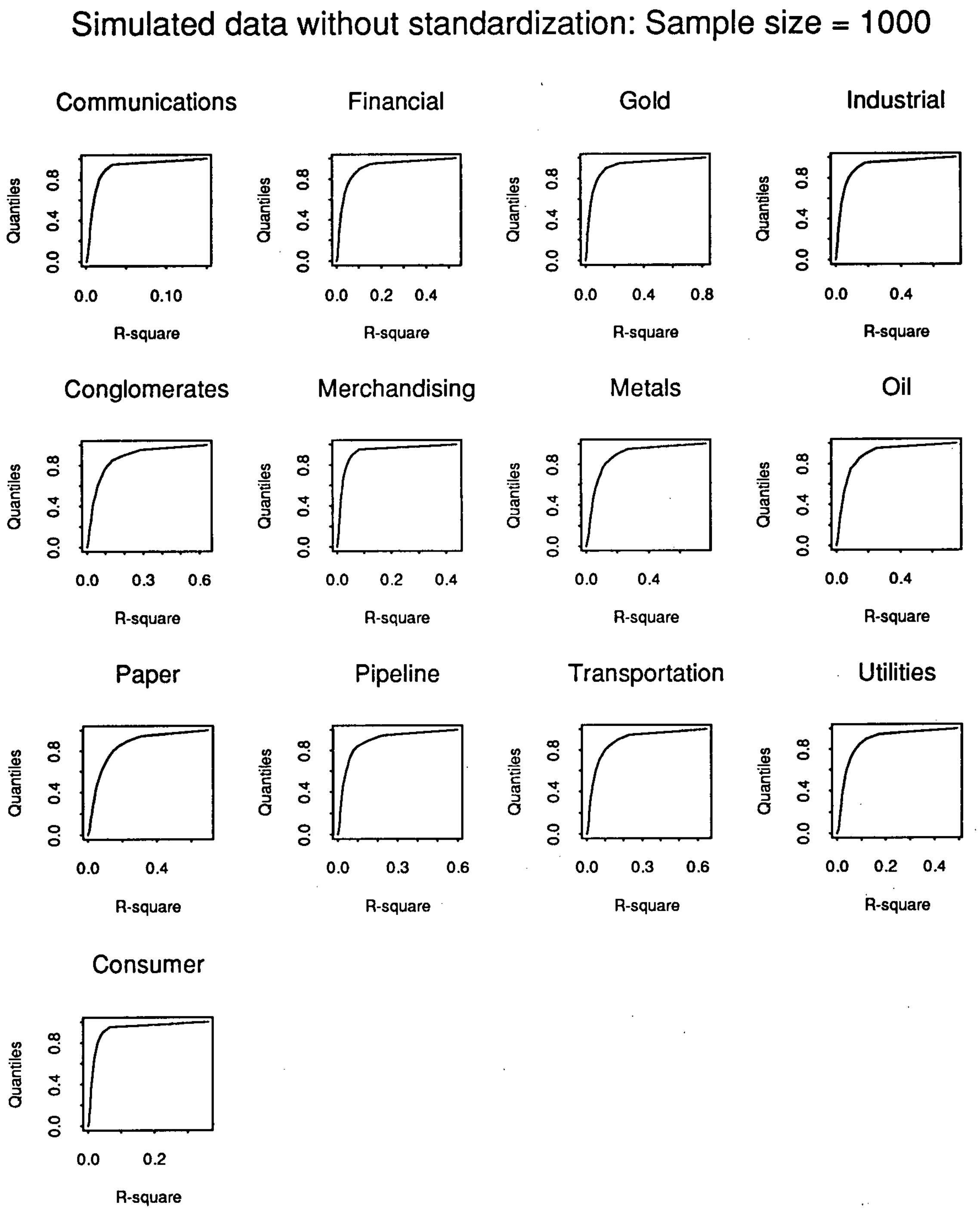


Figure A.3: Quantiles of R^2 using simulations with original data.

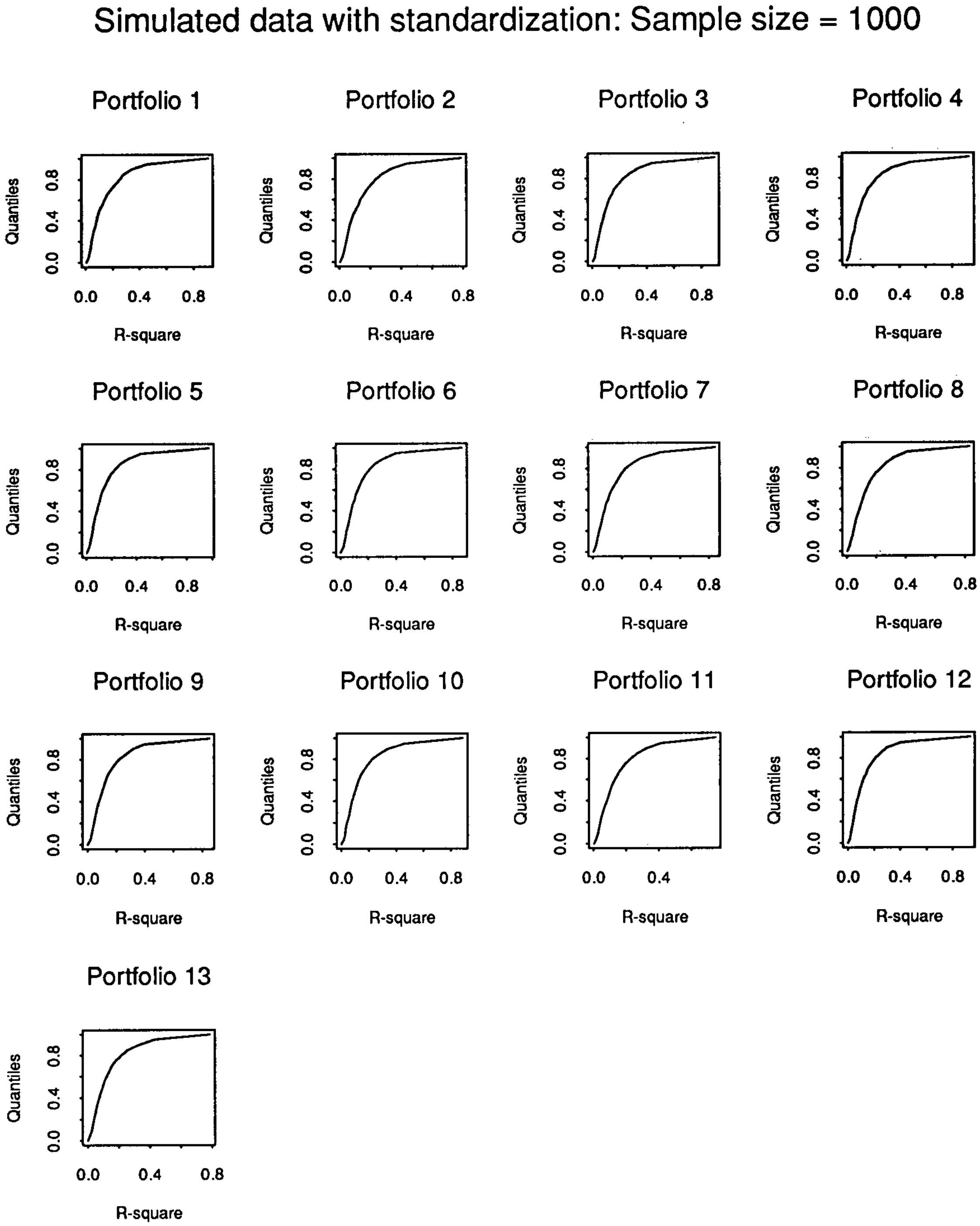


Figure A.4: Quantiles of R^2 using simulations with standardized data.

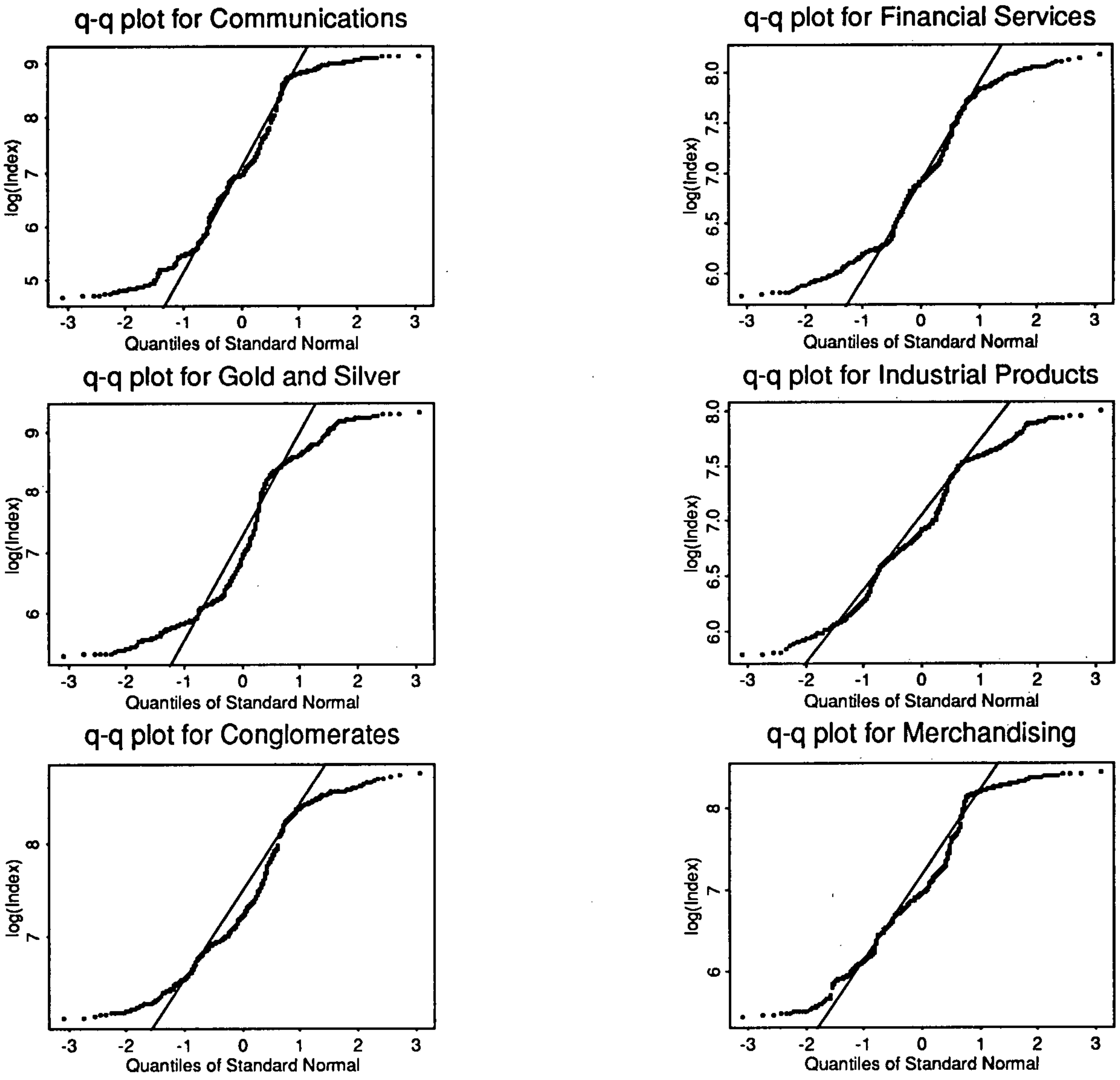
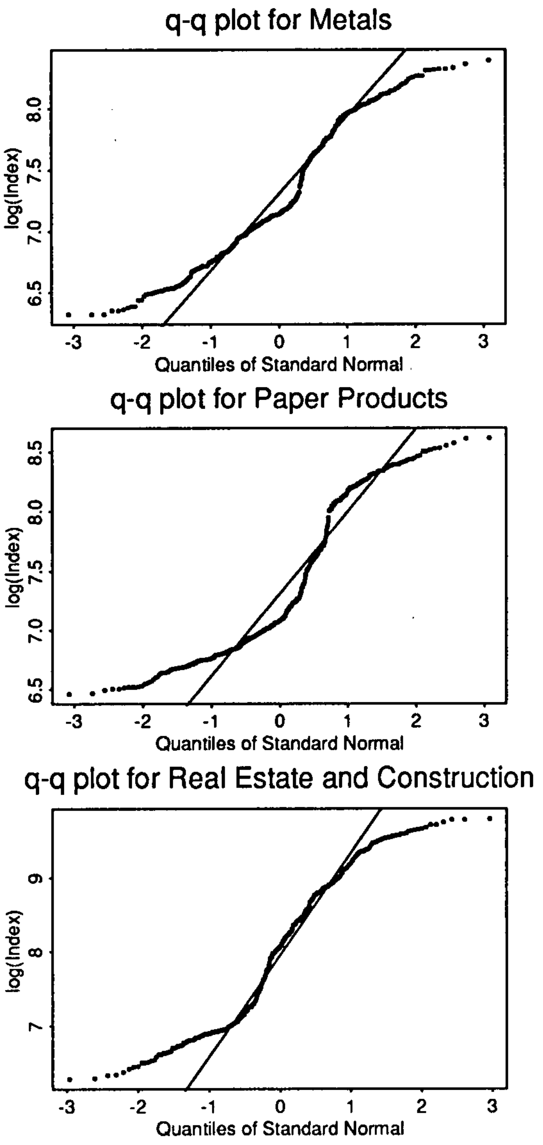


Figure A.5: Q-Q plots of monthly log(index) for TSE industry portfolios



q-q plot for Metals

q-q plot for Oil and Gas

q-q plot for Paper Products

q-q plot for Pipelines

q-q plot for Real Estate and Construction

q-q plot for Transportation

Figure A.6: Q-Q plots of monthly log(index) for TSE industry portfolios

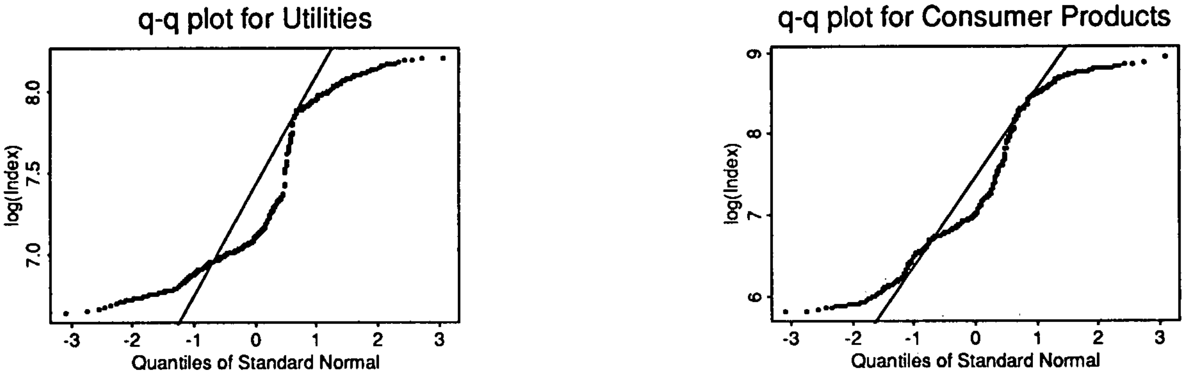


Figure A.7: Q-Q plots of monthly log(index) for TSE industry portfolios