ESSAYS ON ASSET RETURN PREDICTABILITY

by

SUNG-HWAN SHIN

S.M. Massachusetts Institute of Technology (1988)
B.A. Seoul National University (1985)

Submitted to the
Alfred P. Sloan School of Management
in Partial Fulfillment of the Requirements for the
Degree of

DOCTOR OF PHILOSOPHY IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1993

Signature of Author ______________________________
Sloan School of Management
July 2, 1993

Certified by ______________________________
Andrew W. Lo
Professor of Finance
Thesis Supervisor

Accepted by ______________________________
James B. Orlin
Professor of Operations Research
Chairman, Doctoral Program Committee

Abstract

This thesis entitled "Essays on Asset Return Predictability" is composed of the three essays on various aspects of asset return predictability.

In the first essay, the impact of discrete asset prices on serial and cross-sectional spurious autocorrelations is investigated. Discrete asset prices with a minimum tick of '$1/8' are found to cause asset returns serially autocorrelated by almost 30% when the asset price is $2.5 and a standard deviation of the asset return is 0.01. The serial autocorrelations are decreasing as the asset price and the variance of returns increase. As for cross-sectional autocorrelations, the impact of discrete asset prices is not significant regardless of asset prices. Therefore, the lead-lag relationships among individual stock returns which is implied by positively autocorrelated portfolio returns seem to be caused by some economic reasons. As a first step toward identifying economic sources for the lead-lag relationships, individual stock returns are investigated for each portfolio. Empirical findings suggest that it takes more than a day for individual stocks to reach an equilibrium after new information arrives. The distribution of contemporaneous correlations among individual stocks shifts to the right as a portfolio size increases. Finally, the impact of discrete prices seems to be trivial whereas that of nonsynchronous trading is relatively large for daily returns. However, the impact of nonsynchronous trading becomes also trivial for weekly returns. After all, the lead-lag relationships among individual stocks seem to be most closely related to the size of stocks rather than to the price or the number of shares of stocks.

In the second essay, economic implications of the current state stock market predictability are provided from a stochastic dominance point of view. They are: 1) for short-horizon returns, a representative contrarian trading strategy is superior to a simple buy-and-hold investment strategy, but even a presence of 1 percent one-way per-dollar transaction costs overturns the relationship, 2) for long-horizon returns, return reversals are found, but there are *some* risk-averse investors who prefer the market portfolio to the portfolio composed of stocks which performed poor in the previous period, and 3) for the weekend effect, non-Monday stock returns are preferred to Monday returns, and many pairs among non-Monday returns also have stochastic dominance relationships which are usually from empirically *indistinguishable* return distributions for large stocks and from empirically *distinguishable* return distributions for small stocks. Unfortunately, the stochastic dominance test that is used in this paper is found not very robust to tail events. A weighted stochastic dominance test is suggested to improve the robustness of the test to tail events.

In the third essay, return predictability for foreign exchange rates is investigated by implementing artificial neural network models and moving average trading rules. The variance ratio test fails to reject the random walk hypothesis for exchange rates, which implies that linear auto-regressive models are not effective in predicting returns. Both neural network models and moving average trading rules are found to be successful in predicting returns in the sense that i) 'buy' returns have higher means than 'sell' returns with standard deviations and skewnesses comparable, ii) the predicted returns by neural network models have positive correlations with the actual returns, and iii) the 'buy' returns usually first-degree stochastically dominate the 'sell' returns except for a few cases where those returns are empirically indistinguishable. Finally, two estimating methods for neural network models — NLS and back propagation — are compared by the Monte-Carlo experiments. In finite samples, the NLS seems to be more efficient in terms of R^2 and the time to convergence than the back propagation method. Overall, in spite of some problems associated with estimating the models (problem of local minima), the implementation of neural network models for predicting exchange rate returns is generally found to be successful, and it sheds light on more successful applicability of artificial neural networks to the field of finance by investigating the appropriate models more extensively, and also by mitigating the problem of local minima.

Dissertation Committee

Ernst R. Berndt, Professor of Applied Economics.

John C. Cox, Professor of Finance.

Andrew W. Lo (Chairman), Professor of Finance.

Acknowledgements

To my parents, my wife and my son

I joined the Ph.D. program at the Sloan School after spending two years in the Master's program because I thought that a life in academics would be nice. After all, I came to realize that I was too optimistic about being a doctoral student. It was not like running on a nice flat road where I could see the final destination. There were moments of pain as well as moments of joy. There were moments of boredom as well as moments of excitement.

After fulfilling my commitment which was made five years ago, I'm now proud that I'm ready for the next step in my life . Without the help of many people I have known, I could never have reached this moment of joy.

Andrew Lo has been a role model for me since I first met him. He has been the most talented and the most hard working person I have ever seen. I became interested in empirical finance on capital markets while I was working for him as a research assistant at the NBER, and I came to understand the meaning of an excellent lecturer when I was working for him as a teaching assistant. He stimulated my interests in issues that are related to asset return predictability. The encouragement he gave me throughout the Ph.D. program was invaluable. I am deeply indebted to him.

I would never forget the kindness that John Cox has shown me. He encouraged me to join the Ph.D. program at the Sloan School, and gave me strong morale support throughout the program. It was also a very pleasant experience to work for him as a teaching assistant for two years. His thoughtfulness is deeply appreciated.

Special thanks go to Ernie Berndt for allowing me to finish my defense on time. He was willing to be one of my committee members and spared me his precious time for reading my thesis.

A brief discussion with Jiang Wang dramatically improved the quality of my papers. Thanks to him, I could correct all my mistakes in interpreting empirical results, and in developing the theoretical arguments concerning empirical results.

Chi-fu Huang was kind enough to attend my informal lunch seminar as well as the formal batterymarch finance seminar where I presented my work. He gave me a chance to think about several issues in my work more deeply.

Stewart Myers, John Heaton, Jean-Luc Vila, and other seminar participants were kind enough to have interest in my work and to give me valuable suggestions. Mary Marshall was very helpful in dealing with administrative matters. Jaewoo Lee, Kazuhiko Ohashi and Vivian Ellis have been good friends of mine.

I am thankful to all my friends, especially to my best friend Yong-Shik Choo at the SAIS, for sharing my joy and sorrow during my days at MIT.

My parents gave me a continuous encouragement throughout my graduate life. Without their financial support, my graduate work could have never been completed. Strong encouragement from my parents-in-law is also deeply appreciated. They always gave me a precious advice whenever I was in difficult moments.

Most of all, support from my dear wife and my lovely son has been invaluable. It has been her love that has sustained this work. Words are not enough to thank them, so I will not try.

Table of Contents

Introduction and Overview

1 Introduction

It has been less a decade since asset return predictability received the attention of financial economists. Buried under the strong belief that our financial markets are efficient, almost no attention was paid to asset return predictability until the early eighties. Technical analysis was regarded as useless by academicians although it was widely adopted by practitioners in making their investment decisions. Most academic interests were concentrated on the issues concerning fundamental asset valuation.

During the eighties, there was a flood of documentation which showed that asset returns were predictable in many ways. For example, several anomalies such as the Value-Line Effect, the Weekend Effect, and the January Effect were reported by various authors (i.e. see French (1980)). Most of the anomalies concerned the persistent predictability of asset returns that could not be easily explained by our current economic paradigm.

There were other kinds of documentation which illustrated more general patterns of asset return predictability. Long-horizon stock returns were found to have a tendency to be reversed (i.e. see Debondt and Thaler (1985)). Short-horizon stock returns were also found to be predictable in some interesting manners (i.e. see Lo and MacKinlay (1988)). Intuitive or statistical arguments were made to show that the predictability was meaningful in the sense that it may imply irrational behavior of investors.

Three issues (or questions) may arise with respect to asset return predictability. The first issue is about the sources of predictability, and the second one is about the economic implications of asset return predictability. The last issue is about the methods which utilize asset return predictability. The next three chapters of this thesis will deal with these three issues.

In the first chapter, the first issue will be investigated. Considering the fact that stocks are traded at prices in multiples of 1/8th of a dollar,[1] the impact of price

[1]Stocks in the NYSE and the AMEX with prices of greater than one dollar are traded at prices

discreteness on asset return predictability will be examined. Also the relative sensitivity of stock return predictability to the average market value, to the average price, and to the average number of shares of portfolios will be compared by considering the returns of size-sorted, price-sorted and number of shares-sorted portfolios. Finally the sources of portfolio return predictability will be investigated by examining the relationships among individual stocks in those portfolios. The second issue will be discussed in the second chapter. Considering that investors are more interested in results than in the causes of asset return predictability (and therefore financial economists should be interested in the results), capturing the economic implications of asset return predictability is believed to be no less important than finding the causes of the predictability. Instead of testing for the efficient market hypothesis, an alternative concept of stochastic dominance will be implemented to capture the economic implications of asset return predictability.

It is said that there are stochastic dominance relationships among random variables when there is investors' unanimous agreement on the ranks of the random variables. More specifically, when a random variable X is preferred to a random variable Y by all investors whose preferences are continuous and increasing, it is said that X first-degree stochastically dominates Y. A similar definition applies to the second-degree stochastic dominance for risk-averse investors. The stochastic dominance relationships between dynamic trading strategies and passive buy-and-hold strategies will be examined in the presence of transaction costs as well as in the absence of transaction costs. Weekend stock returns will also be compared with other weekday returns to re-examine the Weekend effect from the stochastic dominance point of view. Finally, a simple modification will be made in this chapter for the method of the stochastic dominance test by Klecan, McFadden and McFadden (1991) to improve the robustness of the test to tail events.

In chapter III, the last issue will be discussed. With respect to the characteristics

of the multiples of 1/8th of a dollar.

of a random variable, there may be three cases where i) the random variable is truly not predictable at all, ii) the random variable is predictable, and the predictability can be easily detected by standard statistical tests, and iii) the random variable is predictable but the predictability is difficult to detect with standard statistical tests. Whereas chapter I and chapter II are related to the second case, chapter III will be related more to the third case. Predictable movements of a random variable, for example, in a highly non-linear manner may be difficult to detect as predictable. The extreme case would be a deterministic chaotic movement.

In this chapter, after examining the random characteristics of foreign exchange rate returns with a variance ratio test, artificial neural network models and moving average trading rules will be implemented to capture the return predictability. Recent theoretical development about the artificial neural network suggests that it can approximate a wide class of functions, and implies the possibility of its successful application to various fields of finance (i.e. see Gallant and White (1988) and Hornik, Stinchcombe and White (1989)). Highly non-linear as well as linear patterns of return predictability may be captured by models of the artificial neural network. The method of moving average trading rules is one of the most frequently used tools for technical analysis. It has already been shown by other authors to be somewhat effective in predicting returns (i.e. see LeBaron (1992)) — and the fact that it is frequently used by practitioners strongly supports the effectiveness of this method.[2] Economic judgment of the effectiveness of both artificial neural network models and moving average trading rules will be made with stochastic dominance criteria. Also, two leading estimating methods for neural network models — NLS and back propagation — are compared by the Monte-Carlo experiments.

[2]In this sense, there is a selection bias associated with moving average trading rules.

2 *Overview* — Price Discreteness and Decomposition of Asset Return Predictability

In the seminal paper by Lo and MacKinlay (1988), it was documented that short-horizon stock returns were predictable. Several interesting patterns were observed in the short-horizon stock return predictability. For example, degrees of the predictability decrease as the size of portfolios increases. Lo and MacKinlay (1990a) showed that nonsynchronous trading may cause small spurious autocorrelations which are negative for individual asset returns, and positive for portfolio asset returns. This chapter will address the impact of price discreteness on short-horizon asset return predictability along with the analysis of the behavior of individual stock returns in each portfolio.

Summarizing interesting patterns for short-horizon stock return predictability, they are as follows: 1) Individual stock returns have slightly negative autocorrelations whereas portfolio returns have large positive autocorrelations, 2) Absolute magnitudes of the autocorrelations decrease as the size of portfolios increases, 3) Two portfolio returns are contemporaneously less correlated as the difference between the sizes of portfolios increases, and 4) As the return-horizon gets longer from daily, to weekly, and to monthly, first-order autocorrelations for large sized portfolio returns tend to decrease much faster than those for small sized portfolio returns. In other words, the lead-lag relationships for small stocks tend to persist up to a certain return horizon where as those for large stocks quickly vanish as the return horizon increases.

The first investigation in this chapter is whether price discreteness has a significant impact on serial and cross-sectional autocorrelations. For stocks listed in the New York Stock Exchange and the American Stock Exchange, they are traded at prices which are multiples of 1/8 of a dollar, unless their prices are less than one dollar. The 1/8 of a dollar is, therefore, called a minimum 'tick'. This institutional constraint may cause some estimation problems or a spurious autocorrelation for asset returns.[3]

The impact of price discreteness on serial and cross-sectional autocorrelations for an individual asset return is investigated by extending the results by Gottlieb and Kalay (1985). As in Harris (1990), it is found that price discreteness causes negative spurious first-order serial autocorrelations up to 29% for an individual asset return. Numerically approximated results for different price levels, different means, and different variances will be presented. On the other hand, discreteness of asset prices is found to have almost no impact on cross-sectional autocorrelations among individual asset returns. Lead-lag relationships among individual stock returns are implied by the recent empirical findings that individual stock returns are negatively serially autocorrelated and portfolio returns are positively autocorrelated. The results about the impact of price discreteness on asset return autocorrelations suggest that the lead-lag relationships among individual stock returns do not seem to be spurious. After all, it is strongly suggested that there must be some economic reasons for such large lead-lag relationships among individual stock returns.

As a first step for investigating the sources for short-horizon stock return predictability, relationships among individual stock returns are analyzes as well as some statistical properties for individual stock returns in each portfolio. For example, probabilities for an individual stock return to be positive are estimated conditioned on its return in the previous period and the market return in the previous and the current periods. Three types of portfolio returns are examined; size-sorted, price-sorted, and number of shares-sorted. The relative impact of price discreteness and nonsynchronous trading can be compared by investigating properties of the three different types of portfolios. Properties of the number of shares-sorted portfolio returns, for example, are believed to be affected more by nonsynchronous trading.

Empirical investigation suggests that nonsynchronous trading has a relatively large impact on autocorrelations for daily portfolio returns, whereas the impact is relatively

[3]Lo and MacKinlay (1990a, 1990b) showed that nonsynchronous trading can cause spurious autocorrelations which are negative for individual asset returns and positive for portfolio asset returns.

small for weekly portfolio returns. The impact of price discreteness does not seem to be large both for daily and weekly portfolio returns. For weekly portfolio returns, the smallest size-sorted portfolio has a larger first-order autocorrelation than the smallest number of shares-sorted portfolio or the smallest price-sorted portfolio. Thus lead-lag relationships among weekly individual stock returns do not seem to be spurious.

With respect to the decomposition of return predictability for each portfolio, a large portion of positive autocorrelations seem to arise from the lead-lag relationships between stocks in two groups; one with positive returns and the other with negative returns. It seems to take more than one day for information to be fully and correctly absorbed in stocks since the conditional probability of a zero return is higher when the previous return was zero than when it was non-zero. In other words, when there are non-zero returns one day, there is a tendency for these returns to be also non-zero the following day. Although the market seems to be the most important factor that governs the movement of contemporaneous individual returns, a movement of the previous individual returns relative to the previous market return also seems to be important, especially for small stocks. Finally, the distribution of contemporaneous correlations among individual stock returns is shifting to the right as a portfolio size increases. Idiosyncratic risks seem be substantial even for large stocks since in most cases, contemporaneous correlations are less than 0.5.

This chapter attempts to investigate the behavior of individual returns in each portfolio as a first step toward an important and ultimate goal in the research of the short-horizon lead-lag relationships among individual stock returns. The next step, of course, would be an attempt to identify economic reasons which drive the lead-lag relationships. For various reasons, information may be absorbed in each stock at different speeds. Nevertheless, it is still puzzling why the market allows it to be absorbed with a lag of more than a week in some cases, for certain stocks. The stock market may be less informationally efficient than was expected, especially for small stocks. However, these are just conjectures which have be verified in the future research.

3 *Overview* — Economic Implications of Asset Return Predictability: Stochastic Dominance Comparisons

Recent empirical findings show that stock returns are predictable in short- and long-horizons (see e.g. Lo and MacKinlay (1988) and Debondt and Thaler (1985,1987)). Although the unpredictability of asset returns is neither a sufficient nor a necessary condition for an economic equilibrium (See e.g. Leroy (1973)), the primary concern for the empirical evidence of the stock market predictability has been whether it leads to any 'excessive' profit opportunities, or equivalently, to any market inefficiency (see e.g. Lehmann (1990) and Debondt and Thaler (1985,1987)). Unfortunately, there has been no consensus among financial economists about the implication of stock return predictability to the efficient market hypothesis.

Identifying 'excessive' return opportunities requires the definition of 'normal' returns which relies on economic models. As Fama (1970) pointed out more than two decades ago, the identification of opportunities for 'excessive' returns, or market inefficiency, therefore, has its intrinsic weakness because the results can be interpreted in two different ways (also see e.g. Shiller (1981) and Marsh and Merton (1986)). When market frictions are considered, the identification results have even weaker implications. For example, it is shown that equilibrium asset prices in the presence of market frictions may deviate from the frictionless market's no-arbitrage prices (see Tuckman and Vila (1992)).

In this chapter, I am not going to design economic models that would determine whether the predictability of stock returns leads to any 'excessive' profit opportunities. Instead I am questioning whether under some circumstances which are less restrictive than those in economic models, the predictability leads to investors' unanimous agreement on the ranks of several uncertain investment or trading opportunities. In other words, I am trying to interpret the current state stock market predictability in the economic context without employing any economic paradigm.

When there is investors' unanimous agreement on the ranks of random variables, it is said that there are stochastic dominance relationships among the random variables. More specifically, when a random variable X is preferred to a random variable Y by any investor whose preference is continuous and increasing, it is said that X first-degree stochastically dominates Y. On the other hand, it is said that X second-degree stochastically dominates Y when a random variable X is preferred to a random variable Y by any risk-averse investor whose preference is continuous and increasing. There is also a third-degree stochastic dominance relationship, but it will not be considered in this chapter.[4] The essential idea of the stochastic dominance is well described by Russel and Seo (1989),

> Stochastic dominance rules dictate procedures for discovering unanimous orderings of uncertain prospects appropriate for utility functions within specified sets. ... The concept of stochastic dominance has introduced a convenient structure for analyzing optimal decisions when information on preferences is limited in various ways.

Three kinds of predictable returns are considered. The first is concerned with short-horizon return predictability (See e.g. Lo and MacKinlay (1988)). A representative 'buy the previous losers and sell the previous winners' strategy, known as the contrarian trading strategy, will be compared with a buy-and-hold investment strategy which is the most efficient strategy to use in the absence of return predictability. From recent findings about weekly stock returns – negative autocorrelation for individual returns and positive autocorrelation for portfolio returns – it is conjectured that when there is no transaction cost, contrarian trading strategies out perform the buy-and-hold strategy. However, a certain level of transaction cost will make contrar-

[4] X is said to third-degree stochastically dominate Y when a random variable X is preferred to a random variable Y by any risk-averse investor who has a continuous and increasing preference, and his preference shows decreasing absolute risk-aversion.

ian trading strategies not so attractive, and eventually a large transaction cost will make contrarian trading strategies inferior to the buy-and-hold strategy.

Empirical evidence show that at 5 percent significance level, one-way per-dollar transaction cost of 0.6 percent is enough to prevent the contrarian trading strategies from stochastically dominating the buy-and-hold strategy for all portfolios. At the same significance level, the buy-and-hold strategy stochastically dominates contrarian trading strategies for all portfolios at one-way per-dollar transaction costs of greater than 1 percent. The second is concerned with long-horizon return reversals and a different kind of contrarian trading strategy, known as Debondt and Thaler's 'extreme performance portfolio strategy', is considered (1985, 1987). In terms of monthly mean returns, we confirm the following results by shown by Debondt and Thaler (1985, 1987): 1) extreme losers for the previous 5 years out perform the market (CRSP value-weighted market index), 2) extreme winners under-perform the market, and 3) the difference is much larger between the losers and the market than between the winners and the market. Stochastic dominance test, however, allows us to interpret their performance differently. The market second-degree stochastically dominates the winners, but the losers do *not* stochastically dominate the market. One possible explanation about this result would be that extreme losers sometimes have very negative returns — they sometimes go bankrupt — and therefore, some risk-averse investors prefer the market to the losers. The impact of January returns to long-horizon return reversals is found significant as shown by Zarowin (1990) in the sense that some of the stochastic dominance relationships change when the January returns are removed from the data set.

The third is concerned with a completely different type of return predictability, known as 'the weekend effect'. French (1980) and Gibbons and Hess (1983) have documented the fact that Monday stock returns are significantly lower than other weekdays stock returns with regard to their sample means. Apart from any profitable trading strategy that can exploit the weekend effect, which is not likely to exist because of relatively high transaction costs, there should be still some upward pressure for Monday stock prices. Suppose that you want to buy some stocks and you know that Monday stock prices usually drop. Then you will wait until Monday and will buy stocks on Monday. On the other hand, if you want to sell stocks, you don't want to sell them on Monday. It is asked if investors actually prefer other weekdays returns to Monday returns.

Empirical evidence shows that, in most cases, Monday returns are stochastically dominated by other weekdays returns and that the relationships are strong — they are from empirically distinguishable return distributions. Non- Monday returns also have stochastic dominance relationships, but the relationships are generally weak for large stocks — they are from empirically indistinguishable return distributions — whereas the relationships are strong for small stocks.

Finally, modification of the test statistics was suggested by weighing the test statistics properly to improve the robustness of the test by Klecan, McFadden and McFadden (1991) to tail events. Monte-Carlo experiments show that the weighted test has better finite sample property with regard to the first-degree null hypotheses. With regard to the second-degree null hypotheses, however, the weighted test sometime perform worse than the conventional test. There may be better ways to improve the test so that it has better finite sample property with regard to second-degree null hypotheses as well as first- degree null hypotheses, but they are yet to be developed in future research.

4 *Overview* — Econometric Implementation of Trading Strategies: Neural Networks and Moving Average Rules

There have been two kinds of financial analysts: the fundamental analyst who conducts fundamental analysis, and the technical analyst who conducts technical analysis. Fundamental analysis is concerned with economic valuation of financial assets whereas technical analysis is related to chasing trends. In other words, technical analysis is regarded to be useful by investors who believe that previous movements of asset prices are valuable information for predicting future movements of the asset prices. For academic researchers, 'technical analysis' had been meaningless at least until the financial markets were believed to be efficient.[5] If, for example, asset prices follow random walks, it is of no use to try to predict future asset prices with current information of asset prices.

Recently, there has been growing empirical evidence that some financial asset prices do not follow random walks. For example, Lo and MacKinlay (1988) showed that stock prices did not follow random walks in weekly investment horizons. Also Debondt and Thaler (1985, 1987) showed that long horizon stock returns had a tendency of mean-reversion. Although the deviation from the random walk does not necessarily imply the inefficiency of the market or equivalently the effectiveness of technical analysis (i.e. see Leroy(1973)), it certainly provides a good possibility for technical analysis to be useful.

Whereas the deviation from the random walk hypothesis suggests that technical analysis can be useful, the inability to reject the random walk hypothesis does not necessarily imply that charting is just a waste of time. Highly nonlinear dynamics may not be detected by conventional tests that are usually related to sample means and sample variances. For example, the variance ratio tests examine only the first and

[5] According to Fama's definition, this efficiency is a weak-form efficiency.

the second moments of samples. Therefore, there may well be the case where drawing a chart can actually help predict future movements of random variables, even though conventional statistical tests cannot. Among the highly nonlinear system is a system which seemingly follows random walks but which is in fact generated by the highly nonlinear *deterministic* process. This process is called deterministic 'chaos'. Hsieh (1991) investigated the stock market to see whether or not it is governed by chaotic dynamics. He found no evidence of chaotic behavior in stock returns. A priori, it seems to be difficult to advocate the possibility of chaotic behavior in financial asset returns since asset returns seem more likely to be stochastic rather than deterministic.

In this chapter, attempts will be made to assess the effectiveness of technical trading rules on foreign exchange rates from the stochastic dominance point of view. Along with the most popular technical trading rules which are moving average trading rules, a relatively new econometric method which is called 'artificial neural network' will be implemented to develop new kinds of technical trading rules.

Artificial neural network is a nonparametric regression method that originated in cognitive science areas inspired by the structure of the brain. Due to its general applicability, it has been used in vast areas of science and engineering from cognitive science to artificial intelligence, and even to computer hardware design. It was only recently applied to economics. For example, White (1988) implemented neural network to predict IBM daily stock returns, but unfortunately, found that using neural network did not out perform the random walk model. Utans and Moody (1991) used neural network to predict corporate bond ratings and showed that it performs much better than a linear bond rating predictor.

Since the main purpose of this chapter is to test the usefulness of technical trading rules with the stochastic dominance criteria, attempts will not be made to refine the theory of neural network, or to exploit all kinds of complicated neural network models.[6] Instead, one of the simplest neural network model — 'single hidden layer

[6]Exploiting all kinds of possible models to see if there is any extraordinary profit opportunity will inevitably face a severe selection bias. Although, the bias cannot be completely avoided if only

feedforward models' — will be used to predict the following movements of exchange rates. Several issues concerning about the proper network architectures will also be briefly discussed.

Before investigating technical trading rules, a simple specification test will be conducted to see if the null hypothesis of random walks for exchange rates can be rejected by the variance ratio test method by Lo and MacKinlay (1988).[7]

Empirical evidence shows that the null hypothesis that foreign exchange rates follow random walks is usually not rejected at a 5 percent significance level for all exchange rates. Therefore, the effectiveness of linear auto regressive models in predicting returns are implicitly denied. Both neural network models and moving average trading rules are found to be effective in predicting returns in the sense that i) 'buy' returns have higher means than 'sell' returns with comparable standard deviations and skewnesses, ii) predicted returns from the neural network models are positively correlated with actual returns, and iii) 'buy' returns usually first-degree stochastically dominate 'sell' returns.

Two estimating methods — NLS and back propagation — were compared by the Monte-Carlo simulations. The NLS method usually had quicker convergence results than the back propagation method. In terms of R^2, the NLS method also had a better performance. The impact of random components in the process on the effectiveness of neural network models seemed to be significant in the sense that R^2s decreases by a large degree. Although the NLS method and the back propagation method with a certain learning rate have the same limit, the NLS method seemed to be more efficient than the back propagation method in finite samples.

Overall, the results from the neural network models can be regarded as promising. In spite of the problem of local minima in estimating the models, the simple models were successful in predicting returns. More extensive investigations into various kinds

the past data are involved in tests.

[7]The test method by Lo and MacKinlay is adopted since it is also robust to certain kinds of heteroskedasticity.

of neural network models would increase the predictability of returns for foreign exchange rates. Finally, since they are able to capture the highly non-linear patterns of asset returns in a very systematic way, it is expected that artificial neural networks will be indispensable for most technical analyses in the near future.

References

[1] Gottlieb, G. and A. Kalay (1985), " Implications of the Discreteness of Observed Stock Prices," Journal of Finance 40, 135-153.

[2] Harris, Lawrence (1990), "Estimation of Stock Price Variances and Serial Covariances from Discrete Observations," Journal of Financial and Quantitative Analysis 25, 291-306.

[3] Debondt, W. and R. Thaler (1985), "Does the Stock Market Overreact?," Journal of Finance 40, 793-805.

[4] Debondt, W. and R. Thaler (1987), "Further Evidence on Investor Overreaction and Market Seasonality," Journal of Finance 42, 557-582.

[5] Fama, Eugene (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," Journal of Finance 25, 383-417.

[6] French, Kenneth (1980), "Stock Returns and the Weekend Effect," Journal of Financial Economics 8, 1-2.

[7] Gibbons, M. and P. Hess (1981), "Day of the Week Effects and Asset Returns," Journal of Business 54, 579-596.

[8] Hadar, J. and W. Russell (1969), "Rules for Ordering Uncertain Prospects," American Economic Review 59, 25-34.

[9] Klecan, L., R. McFadden and D. McFadden (1991), "A Robust Test for Stochastic Dominance," Working Paper, MIT and Cornerstone Research.

[10] Lehmann, Bruce N. (1990), "Fads, Martingales, and Market Efficiency," Quarterly Journal of Economics 105, 1-28.

[11] Leroy, S.F. (1973), "Risk Aversion and the Martingale Property of Stock Returns," International Economic Review 14, 436-446.

[12] Lo, A. and C. MacKinlay (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," Review of Financial Studies 1, 41-66.

[13] Lo, A. and C. MacKinlay (1990), "When Are Contrarian Profits Due To Stock Market Overreaction?," Review of Financial Studies 3, 175-205.

[14] Marsh, T. and R. Merton (1986), "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," American Economic Review 76, 483-498.

[15] Russell, W. and T. Seo (1989), "Representative Sets for Stochastic Dominance Rules," in T. Fomby and K. Seo (eds) *Studies in the Economics of Uncertainty*, Springer, New York, 59-76.

[16] Shiller, R. (1981), "Do Stock Prices Move Too Much To Be Justified By Subsequent Changes in Dividends?," American Economic Review 71, 421-436

[17] Tuckman, B. and Jean-Luc Vila (1992), "Arbitrage With Holding Costs: A Utility-Based Approach," Journal of Finance 47, 1283-1302.

[18] Utan, J. and J. Moody (1991), "Selecting Neural Network Architectures via the Prediction Risk: Application to Corporate Bond Rating Prediction," Proceedings of the *First International Conference on Artificial Intelligence Applications on Wall Street*, IEEE Computer Society Press, Los Alamitos, CA.

[19] White, H. (1988), "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns," Working Paper, University of California, San Diego.

æ

Chapter I

Price Discreteness and Decomposition of Asset Return Predictability

1 Introduction

In the seminal paper by Lo and MacKinlay (1988), it was documented that short-horizon stock returns were predictable. Two important questions were concerned about stock return predictability. The first question is about the economic implications of stock return predictability. In theory, asset return predictability does not necessarily imply market inefficiency (see e.g. Leroy (1973)). Therefore, whether the predictability implies any kind of irrational behavior of investors is an empirical issue. Lehmann (1990) tried to link the short-horizon stock return predictability to market inefficiency by showing that some trading strategies could make a large profit even after transaction costs. Unfortunately, in his work, the assumptions about transaction costs were problematic, and the risk of the trading strategies was not well defined. To avoid the problems that are related to testing for the efficient market hypothesis and to defining risks in dynamic settings, Shin (1992) interpreted stock return predictability from a stochastic dominance point of view by comparing two different trading strategies.

The second question is about reasons for short-horizon stock return predictability. Several interesting patterns were observed in short-horizon stock return predictability. For example, degrees of predictability decrease as the size of a portfolio increases. Lo and MacKinlay (1990a) showed that nonsynchronous trading may cause small spurious autocorrelations which are negative for individual asset returns and positive for portfolio asset returns. This paper of my thesis will address the impact of price discreteness on short-horizon asset return predictability, along with the analysis of the behavior of individual stock returns in each portfolio.

Summarizing interesting patterns for short-horizon stock return predictability, they are as follows: 1) Individual stock returns have slightly negative autocorrelations, whereas portfolio returns have large positive autocorrelations, 2) Absolute magnitudes of the autocorrelations decrease as the size of a portfolio increases, 3) Two portfolio returns are contemporaneously less correlated as the difference between the sizes of portfolios increases, and 4) As the return-horizon gets longer from daily, to weekly, and to monthly, first-order autocorrelations for large sized portfolio returns tend to decrease much faster than those for small sized portfolio returns. In other words, the lead-lag relationships for small stocks tend to persist up to a certain return horizon, whereas those for large stocks quickly vanish as the return horizon increases.

The first investigation in this paper is whether price discreteness has a significant impact on serial and cross-sectional autocorrelations. For stocks listed in the New York Stock Exchange and the American Stock Exchange, they are traded at prices which are multiples of 1/8 of a dollar unless their prices are less than one dollar. The 1/8 of a dollar is, therefore, called a minimum 'tick'. This institutional constraint may cause some estimation problems or a spurious autocorrelation for asset returns.[1]

Although the problem of price discreteness has usually been ignored in financial economics literatures, there are several recent papers by other authors which try to assess the effect of price discreteness. Among them are Ball (1988), Gottlieb and Kalay (1985) (G-K), Cho and Frees (1988) (C-F), Harris (1990), and Hausman, Lo and MacKinlay (1991) (H-L-M). Ball analyzed the bias for estimating the variance when the true price process follows the Brownian motion, and G-K studied it when the process follows geometric Brownian motion. C-F used stopping time to get an unbiased estimator of variance under the continuous monitoring assumption. Harris added the bid-ask spread to G-K and used the maximum likelihood method to estimate the model. Also, Harris showed that the variance was overestimated, and that there was a spurious negative serial autocorrelation for an individual asset return. Finally, H-L-M took into consideration the price discreteness in investigating transaction data.

In this paper, the impact of price discreteness on serial and cross-sectional autocorrelations for an individual asset return is investigated by extending the results by Gottlieb and Kalay (1985). As in Harris (1990), it is found that price discreteness

[1]Lo and MacKinlay (1990a, 1990b) showed that nonsynchronous trading can cause spurious autocorrelations which are negative for individual asset returns and positive for portfolio asset returns.

causes negative spurious first-order serial autocorrelations up to 29% for an individual asset return. Numerically approximated results for different price levels, different means, and different variances will be presented. On the other hand, the discreteness of asset prices is found to have almost no impact on cross-sectional autocorrelations among individual asset returns. Lead-lag relationships among individual stock returns are implied by the recent empirical findings that individual stock returns are negatively serially autocorrelated and portfolio returns are positively autocorrelated. The results about the impact of price discreteness on asset return autocorrelations suggest that the lead-lag relationships among individual stock returns do not seem to be spurious. After all, it is strongly suggested that there must be some economic reasons for such large lead-lag relationships among individual stock returns.

As a first step for investigating the sources for short-horizon stock return predictability, relationships among individual stock returns, as well as some statistical properties for individual stock returns are analyzed for each portfolio. For example, probabilities, for an individual stock return to be positive are estimated conditioned on its return in the previous period and on the market return in the previous and the current periods. All stock returns are sorted by three different variables; period-end market value (or size), price, and number of shares outstanding. Relative impact of price discreteness and nonsynchronous trading can be compared by investigating properties of three different kinds of portfolios. Properties of the number of shares-sorted portfolios, for example, are believed to be affected more by nonsynchronous trading.

Empirical investigation suggests that nonsynchronous trading has a relatively large impact on autocorrelations for daily portfolio returns, whereas the impact is relatively small for weekly portfolio returns. The impact of price discreteness does not seem to be significant for both daily and weekly portfolio returns. For weekly portfolio returns, the smallest size-sorted portfolio has a larger first-order autocorrelation than the smallest number of shares-sorted portfolio, or the smallest price-sorted portfolio.

Thus lead-lag relationships among weekly individual stock returns do not seem to be spurious.

This paper is organized as follows. In section 2, a simple story will illustrate how price discreteness can cause a negative autocorrelation for an individual asset return and why its impact on cross-sectional autocorrelations is negligible. In section 3, the impact of price discreteness on spurious serial and cross-sectional autocorrelations for individual asset returns will be assessed under a hypothetical asset price process. In section 4, the behavior of individual stocks in several different kinds of portfolios will be investigated. Summary and conclusions will follow in section 5.

2 A Simple Example

Suppose that there are two stocks, A and B, whose prices can take only integer values. If the true value of the stock is in the interval of $[n, n+1)$, the observed price of the stock is assumed to be n. At time t, suppose that the true values of the two stocks, V_A and V_B are the same ($V_A = V_B = n\frac{1}{2}$). It is assumed that there is one common factor which affects the values of the two stocks at the same time. Therefore, the value of the stock is assumed to be determined by one common factor and one idiosyncratic factor which does not affect the other stock. The common factor and the idiosyncratic factors are assumed to be independent of one another, and they are also assumed to be i.i.d. over time having zero means.

Based on the above assumptions, the true value of each stock is not serially correlated, and there is also no cross-sectional autocorrelations between the two stocks. Therefore, if stock prices could take any real values, no autocorrelation should have been observed for asset returns. Now let's see what is happening when prices can take only integer values.

At time $t+1$, let's suppose that the common factor has a positive realized value and the idiosyncratic factor for stock B also has a positive value so that the true value of stock B is in the interval of $[n+1, n+2)$. Therefore, the observed price for stock B would be $n+1$. When the interval is relatively large compared with the variance of the common factor and idiosyncratic factors, the true value of the stock B is more likely to be closer to $n+1$ rather than to $n+2$. Therefore, for stock B, the price is more likely to drop in the next period. This will cause a negative serial correlation for individual stock returns.

At the same time ($=t+1$), for stock A, suppose that the idiosyncratic factor has a zero value so that the true value of the stock A increases but still stays in the interval of $[n, n+1)$. The observed price for stock A will be n, but the price has a higher chance to go up in the next period than to go down, since the true value of the stock is closer to $n+1$ than to n. On the other hand, if the price for stock A goes up at time $t+1$ due to a positive realization of the idiosyncratic factor, its price in the next period is more likely to drop for the same reason for stock B. These two conditional covariances will be washed out so that the unconditional impact becomes negligible. Though the story in this section may be too simple and artificial, it serves to illustrate a simple case where price discreteness can cause a spurious autocorrelation for asset returns. More formal analysis of the impact of the price discreteness on a spurious asset return predictability will be addressed in the next section.

3 Impact of Price Discreteness

3.1 Assumptions and A Model

The *true* value processes are assumed to follow a geometric Brownian motion such that the *virtual* return processes have a normal distribution for any finite time interval. There is a minimum allowable price movement which is call a 'tick'. It will be denoted as 'd'. Listing the assumptions and notations, they are as follows.

- P_{it} : '*True*' value of stock 'i' at time t

- $R_{it} = logP_{it} - logP_{it-1}$: Return for stock 'i' from time $t-1$ to t

- $\widehat{P}_{it} = [\ nd \text{ if } nd - d/2 \leq P_{it} \leq nd + 2/d]$: Observed price for stock 'i' at time t

- $\widehat{R}_{it} = \log \widehat{P}_{it} - \log \widehat{P}_{it-1}$: Observed return for stock 'i' from time $t-1$ to t

- $R_{it} = \mu_i + \beta_i \Lambda_t + \varepsilon_{it}$
  where $\Lambda_t \sim N(0, \sigma_\Lambda^2)$ and $\varepsilon_{it} \sim N(0, \sigma_i^2)$.

- Define $\tilde{\sigma_i}^2 = \sigma_\Lambda^2 + \sigma_i^2$.

- $\widehat{P}_{i0} = P_{i0} = P_i^o \ for \ any \ i$

The following theorems show estimators which are required to estimate sample variances and autocorrelations. Let's start from the theorem by Gottlieb and Kalay (1985).

Theorem 1 [Gottlieb and Kalay] $E(\widehat{R}_{it+1}^L \mid \widehat{P}_{it} = nd)$
$= \sum_{j=1}^{\infty} (\log(j) - \log(n))^L \cdot Prob_j$
where
$Prob_j = Prob(\widehat{P}_{it+1} = jd \mid \widehat{P}_{it} = nd) = \int_{nd-d/2}^{nd+d/2} h(n, t, P_{io}, x)$
$\times \left\{ \Phi\left(\frac{1}{\tilde{\sigma}_i} \log\left(\frac{(j+1/2)d}{x}\right) - \frac{\mu_i}{\tilde{\sigma}_i}\right) - \Phi\left(\frac{1}{\tilde{\sigma}_i} \log\left(\frac{(j-1/2)d}{x}\right) - \frac{\mu_i}{\tilde{\sigma}_i}\right)\right\} dx$
Here, $\Phi(\cdot)$ *is the standard normal distribution and*
$h(n, t, P_i, o, x) = \frac{d}{dx} Prob(P_{it} \in [(n-1/2)d, x) \mid P_{i0} = P_i^o, \ \widehat{P}_{it} = nd)$ *which is well expressed in Lemma 1 of Gottlieb and Kalay (1985).*

Proof: See Gottlieb and Kalay (1985) Theorem 1.

$h(n, t, P_i^o, x)$ is the probability density function that the true value at time t is x when the observed price at time t is nd and the true value at time 0 is P_i^0. In the appendix of G-K, it is shown that $h(n, t, P_i^o, x)$ is asymptotically uniformly distributed as t goes to infinity. For computational convenience, the uniform distribution will be used to approximate $h(\cdot)$ as $1/d$ in our numerical evaluation of the estimators. Table 1.1 and 1.2 give numerically approximated values for the estimators in theorem 1.

The following theorem shows an estimator which is necessary to assess first-order autocorrelations for an individual asset return.

Theorem 2 $E(\widehat{R}_{it+2} \cdot \widehat{R}_{it+1} \mid \widehat{P}_{it} = nd)$
$\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} (\log(j) - \log(k)) \cdot (\log(k) - \log(n)) \cdot Prob_{kj}$,
where

$$Prob_{kj} = Prob(\widehat{P}_{it+1} = kd, \widehat{P}_{it+2} = jd \mid \widehat{P}_{it} = nd)$$
$$= \int_{(n-1/2)d}^{(n+1/2)d} \int_{(k-1/2)d}^{(k+1/2)d} h_{0t} \cdot \Phi\left(\frac{1}{\bar{\sigma}_i} \log\left(\frac{(k-1/2)d}{x}\right) - \frac{\mu_i}{\bar{\sigma}_i} \le Z \le \frac{1}{\bar{\sigma}_i} \log\left(\frac{(k+1/2)d}{x}\right) - \frac{\mu_i}{\bar{\sigma}_i}\right) dx$$
$$\times \int_{(n-1/2)d}^{(n+1/2)d} \int_{(k-1/2)d}^{(k+1/2)d} h_{0t} h_{tt+1} \cdot \Phi\left(\frac{1}{\bar{\sigma}_i} \log\left(\frac{(j-1/2)d}{y}\right) - \frac{\mu_i}{\bar{\sigma}_i} \le Z \le \frac{1}{\bar{\sigma}_i} \log\left(\frac{(j+1/2)d}{y}\right) - \frac{\mu_i}{\bar{\sigma}_i}\right) dxdy$$

where

$h_{0t} = h(n, t, P_i^o, x) = Prob(P_{it} = x \mid \widehat{P}_{it} = nd, P_{i0} = P_i^o)$ *in Theorem 1 and*
$h_{tt+1} = h(k, t+1, P_{it} = x, y) = Prob(P_{it+1} = y \mid \widehat{P}_{it+1} = kd, P_{it} = x)$

$$= \frac{\frac{d}{dy}\Phi\left(\frac{1}{\bar{\sigma}_i}\log(y/x) - \frac{\mu_i}{\bar{\sigma}_i}\right)}{\Phi\left(\frac{1}{\bar{\sigma}_i}\log\left(\frac{(k+1/2)d}{x}\right) - \frac{\mu_i}{\bar{\sigma}_i}\right) - \Phi\left(\frac{1}{\bar{\sigma}_i}\log\left(\frac{(k-1/d)d}{x}\right) - \frac{\mu_i}{\bar{\sigma}_i}\right)}.$$

Proof: Using the Bayes' Rule,

$$Prob_{kj} = Prob(\widehat{P}_{it+1} = kd, \widehat{P}_{it+2} = jd \mid \widehat{P}_{it} = nd)$$
$$= Prob(\widehat{P}_{it+1} = kd \mid \widehat{P}_{it} = nd) Prob(\widehat{P}_{it+2} = jd \mid \widehat{P}_{it} = nd, \widehat{P}_{it+1} = kd).$$

Straight forward calculation for each probability will prove the theorem.

In the above equations, h_{tt+1} is the probability that the true value at time $t+1$ is y when the observed price at time $t+1$ is kd and the true value at time t is x. Table 1.3 shows numerically approximated autocorrelations from theorem 1 and theorem 2.

The following theorem shows an estimator which is necessary for cross-sectional autocorrelations.

Theorem 3 $E(\widehat{R}_{1t+1}\widehat{R}_{2t+\tau}^{\tau} \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d)$
$= \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} (\log(k) - \log(n_1)) \cdot (\log(j) - \log(n_2)) \cdot Prob_{kj}$,

where

$Prob_{kj} = Prob(\widehat{P}_{1t+1} = kd, \widehat{P}_{2t+\tau} = jd \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = z_2 d \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d)$
$= \int_{n_1 d-d/2}^{n_1 d+d/2} \int_{n_2 d-d/2}^{n_2 d+d/2} h_1 h_2 \cdot \int_{-\infty}^{\infty} \Phi(B_1)\Phi(B_2)\phi(z)dzdxdy$

where

$h_1 = h(n_1, t, P_1^o, x)$, $h_2 = h(n_2, t, P_2^o, y)$ *and*

$$B_1 = \left(\frac{1}{\sigma_1}\log\left(\frac{(k-1/2)d}{x}\right) - \frac{\mu_1+z}{\sigma_1}, \frac{1}{\sigma_1}\log\left(\frac{(k+1/2)d}{x}\right) - \frac{\mu_1+z}{\sigma_1}\right),$$
$$B_2 = \left(\frac{1}{\sigma_{2\tau}}\log\left(\frac{(j-1/2)d}{y}\right) - \frac{\tau\mu_2+z}{\sigma_{2\tau}}, \frac{1}{\sigma_{2\tau}}\log\left(\frac{(j+1/2)d}{y}\right) - \frac{\tau\mu_2+z}{\sigma_{2\tau}}\right)$$

Here, $\widehat{R}_{2t+\tau}^{\tau}$ *is the* τ *period return for stock 2 from the time* t *to* $t+\tau$. $\sigma_{2\tau}^2 = \tau\sigma_2^2 - \sigma_\Lambda^2$, *and* $\Phi(\cdot)$ *is the standard normal distribution function and* $\phi(\cdot)$ *is the normal density function with mean 0 and variance* σ_Λ^2.

Proof: As $h = h(n, t, P_i^o, x)$ is shown that it is asymptotically uniform on $[(n-1/2)d, (n+1/2)d]$ in the appendix of Gottlieb and Kalay (1985), a very similar proof will show that $H = H(n_1, n_2, t, P_1^o, P_2^o, x, y)$ is asymptotically uniform on $[(n_1 - 1/2)d, (n_1+1/2)d] \times [(n_2-1/2)d, (n_2+1/2)d]$ which implies that
$H = H(n_1, n_2, t, P_1^o, P_2^o, x, y) = h(n_1, t, P_1^o, x)h(n_2, t, P_2^o, y) = h_1 h_2$, where
$H(n_1, n_2, t, P_1^o, P_2^o, x, y) = \frac{d}{dx}\frac{d}{dy}Prob(P_{1t} \in [(n_1 - 1/2)d, x), P_{2t} \in [(n_2-1/2)d, y) \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d)$.
Then straight forward calculation will prove the theorem.

From the theorem 3, the next corollary shows an estimator which is necessary for cross-sectional autocorrelations.

Corollary 1 $E\left(\widehat{R}_{1t+1}\widehat{R}_{2t+2} \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d\right)$
$= E\left(\widehat{R}_{1t+1}\widehat{R}_{2t+2}^2 \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d\right) - E\left(\widehat{R}_{1t+1}\widehat{R}_{2t+1}^1 \mid \widehat{P}_{1t} = n_1 d, \widehat{P}_{2t} = n_2 d\right)$.

Proof: Obvious.

Table 1.4 shows numerically approximated cross-sectional autocorrelations.

3.2 Numerical Approximations

The estimators in the previous subsection are numerically approximated by using the quadrature method of D01FBF with D01BBF and B01BAZ subroutine in the NAG fortran library. For all estimators, the upper and the lower bounds for returns are set as 100% and -100%. An indefinite integral for a normally distributed random variable is modified to a definite integral on the range of [mean+4std, mean-4std] which stands for 99.9% of the whole probability. Parameter values for the approximations are as follows.

$$\mu(\text{mean return}) = 0.0006, 0.003 : \sigma_\Lambda(\text{common factor std.}) = 0.009, 0.02$$

$$\tilde{\sigma}_i(\text{std. of a return}) = 0.09, 0.07, 0.06, 0.05, 0.04, 0.03, 0.02, 0.01$$

$$\text{Prices: } 2.5, 5.0, 10.0, 20.0$$

Table 1.1 and table 1.2 show sample means, sample variances, and [sample variance/true variance] ratios from theorem 1. They are consistent with the results by Gottlieb and Kalay (1985). As the price is smaller and the total variance (σ^2) is larger, the sample mean is more downward biased. On the other hand, the sample standard deviation is more upward biased as the price and the total variance are smaller. The sample variance is as large as four times the true variance when the price is 2.5 and the true standard deviation is 0.1.

Table 1.3 shows spurious serial first-order autocorrelations. Values of μ do not seem to have any significant impact on the autocorrelations. As the price is smaller, there are more negative spurious autocorrelations. Also, as the true standard deviation is smaller, autocorrelation coefficients are more negative. In the extreme case where the price is 2.5 and the standard deviation is 0.1, the spurious first-order autocorrelation coefficient is almost -30%.

Finally, table 1.4 shows spurious cross-sectional autocorrelations. It is clear that price discreteness does not have any significant impact on cross-sectional autocorrelations although all the numbers are non-negative. Those results imply that the high positive autocorrelations of portfolio stock returns are more likely due to some economic sources rather than institutional factors.

In the next section, empirical analysis of relationships among individual stock returns will be performed as a first step for investigating the sources for positive autocorrelations for portfolio returns.

4 Empirical Analysis of Individual and Portfolio Returns

In this section, three types of portfolio returns (and individual returns for each portfolio) are going to be examined; size-sorted, price-sorted and number of shares-sorted. The reason for investigating three different kinds of portfolio returns, rather than just size-sorted portfolio returns is to gauge the relative impact of price discreteness and nonsynchronous trading on the properties of portfolio returns.

4.1 Data and Portfolio Formation

CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31 is used. Daily and weekly portfolio returns are investigated. For daily portfolio returns, depending on each stock's market value (or size), price, and the number of shares outstanding as of the last day of the previous month, 20 size-sorted, 20 price-sorted, and 20 number-of-shares-sorted portfolios are formed and rebalanced every month with stocks that are continuously listed for the month, and have no missing data during the month. For weekly portfolio returns, the same kinds of portfolios are formed and rebalanced every year with stocks which are continuously listed for the year and which have data with no more than 10 missing data during the year. The reason for rebalancing weekly portfolio returns on a yearly, rather than monthly basis is to lose the least number of data points in the investigation of individual stock returns in each portfolio. For example, stocks in one portfolio in a given period may well belong to another portfolio in the following period. This will cause a problem in calculating cross-sectional autocorrelations among individual stocks in each portfolio. Therefore, it is better to remove the first day and the last day returns in each month in calculating first-order autocorrelations among daily individual stock returns. If portfolios are rebalanced every month for weekly returns, two weekly returns (out or four or five weekly returns) would have to be removed each month.

4.2 Portfolio Returns

Table 1.5a and 1.5b show sample moments for the three different portfolios. The sample moments for number-of-shares-sorted portfolios are the least variant upon the ranks of portfolios (i.e. from the smallest to the largest). Those for size-sorted portfolios and price-sorted portfolios are similar for the same ranked portfolios, but the smallest and the largest price-sorted portfolios have slightly higher sample variances. From the results of the impact of price discreteness in the previous section, it is conjectured that 1) the slightly higher sample variances for small price-sorted portfolios are partly due to an estimation bias and 2) on average, firm size determines the idiosyncratic risk for an individual stock return (i.e. inversely proportional relationships).

The next table 1.6 shows the first-order autocorrelations for three different daily and weekly portfolio returns. Daily returns for the smallest number-of-shares-sorted portfolio have a significantly higher first-order autocorrelation than the other two smallest portfolio returns. If the assumption that non-trading probability is higher as the number of shares outstanding is smaller is correct, a large portion of the high autocorrelation seems to be spurious, arising from nonsynchronous trading. The impact of nonsynchronous trading becomes insignificant for weekly portfolio returns — small number-of-shares-sorted portfolios have lower autocorrelations than small size-sorted portfolios. The impact of price discreteness on autocorrelations is insignificant for both daily and weekly portfolio returns. The lead-lag relationships among individual stocks seem to be most closely related to a firm size.[2]

To assess the difference among three kinds of portfolios, characteristics of individual stocks in the portfolios are shown in table 1.7. The probability of a zero return is, on average, the least variant for number-of-shares-sorted portfolios and the most variant for price-sorted portfolios, as the rank of a portfolio moves from the smallest to the largest. Obviously, price discreteness seems to have the closest relationship with the probability of a zero return. As for average prices and average market values, no distinction is made between daily and weekly portfolios since they are almost identical. For all three kinds of portfolios, average prices and average market values increase as the rank of a portfolio increases. Nevertheless, there is a large difference among the three kinds of portfolios in absolute values. For example, the average price and the average market value for the smallest size-sorted portfolio is \$4.2 and \$3.5 million, whereas that for the smallest number-of-shares-sorted portfolio is \$19.9 and \$15.7 million. Therefore, it can safely be assumed that a large portion of return properties (i.e. autocorrelations) for each type of portfolio can be attributed to the portfolios' sorting variable.

4.3 Decomposition of Autocorrelations

Let's define some notations as follows;

- R_t^p: Portfolio return at time t
- R_{it}: Individual return for stock 'i' at time t
- G_{pt}: Group of individual stocks whose returns at time t is positive
- G_{nt}: Group of individual stocks whose returns at time t is negative

[2]A high positive autocorrelation for portfolio returns generally implies large lead-lag relationships among individual stock returns in the portfolio since individual returns usually exhibit negative serial autocorrelations.

- G_{zt}: Group of individual stocks whose returns at time t is zero
- N_{pt}, N_{nt}, N_{zt}: Number of stocks in G_{pt}, G_{nt}, G_{zt} respectively
- N: Total number of stocks in the portfolio ($= N_{pt} + N_{nt} + N_{zt}$).

Now consider the following decomposition of a portfolio return and its serial first-order autocorrelation:

$$R_t^p = \frac{1}{N}\left(\sum_{i\in G_{pt}} R_{it} + \sum_{i\in G_{nt}} R_{it} + \sum_{i\in G_{zt}} R_{it}\right),$$

$$Corr(R_t^p, R_{t+1}^p) = \frac{Cov(R_t^p, R_{t+1}^p)}{Var(R_t^p)},$$

$$Cov(R_t^p, R_{t+1}^p) = E(R_t^p R_{t+1}^p) - [E(R_t^p)]^2,$$

$$E(R_t^p R_{t+1}^p) = \frac{1}{N^2}\{\sum_{i\in G_{pt}}\sum_{j\in G_{pt}} E(R_{it}R_{jt+1}) + \sum_{i\in G_{nt}}\sum_{j\in G_{nt}} E(R_{it}R_{jt+1}) + \sum_{i\in G_{pt}}\sum_{j\in G_{nt}} E(R_{it}R_{jt+1})$$

$$+ \sum_{i\in G_{nt}}\sum_{j\in G_{pt}} E(R_{it}R_{jt+1}) + \sum_{i\in G_{pt}}\sum_{j\in G_{zt}} E(R_{it}R_{jt+1}) + \sum_{i\in G_{nt}}\sum_{j\in G_{zt}} E(R_{it}R_{jt+1})\}$$

$$\approx \frac{1}{N^2}\{Avg_{(t)}[N_{pt}^2 Avg_{(i,j\in G_{pt})}(R_{it}R_{jt+1})] + Avg_{(t)}[N_{nt}^2 Avg_{(i,j\in G_{nt})}(R_{it}R_{jt+1})]$$

$$+Avg_{(t)}[N_{pt}N_{nt}Avg_{(i\in G_{pt}, j\in G_{nt})}(R_{it}R_{jt+1})] + Avg_{(t)}[N_{pt}N_{nt}Avg_{(i\in G_{nt}, j\in G_{pt})}(R_{it}R_{jt+1})]$$

$$+Avg_{(t)}[N_{pt}N_{zt}Avg_{(i\in G_{pt}, j\in G_{zt})}(R_{it}R_{jt+1})] + Avg_{(t)}[N_{pt}N_{zt}Avg_{(i\in G_{nt}, j\in G_{zt})}(R_{it}R_{jt+1})]\}$$

$$\equiv_{define} \{(\text{P-P}) + (\text{Ng-Ng}) + (\text{P-Ng}) + (\text{Ng-P}) + (\text{P-Z}) + (\text{Ng-Z})\} \times Var(R_t^p). \quad (1)$$

Here, $Avg_{(t)}$ denotes an average over time and $Avg_{(i,j...)}$ denotes an average over different stocks. The last line of the above equations defines the decomposition of a portfolio autocorrelation among three different groups of stocks; stocks with positive returns, stocks with negative returns, and stocks with zero returns. P-Z, for example, stands for the approximate contribution of lead-lag relationships between stocks in G_{pt} and G_{zt} to the total portfolio autocorrelation. Table 1.8a and 1.8b show the results of the autocorrelation decomposition for daily and weekly returns. For the smallest portfolio daily returns, most of the positive autocorrelations are due to large lead-lag relationships between stocks in G_{pt} and G_{nt} (P-Ng, Ng-P) which decrease as the rank of a portfolio increases. On the other hand, for the largest portfolio daily returns, positive lead-lag relationships among stocks in G_{pt} (P-P) and stocks in G_{nt} (Ng-Ng) consist for the most part of positive autocorrelations. P-P and Ng-Ng generally increase first and then slightly decrease a the rank of a portfolio increases.

The decomposition for weekly portfolio returns is similar. P-Ng and Ng-P decrease as the rank of a portfolio increases, and they are the most responsible for positive autocorrelations of small portfolios. There are almost consistent positive lead-lag relationships among stocks in G_{pt} (P-P) that are equally responsible with P-Ng and Ng-P for positive autocorrelations of medium portfolios (i.e. portfolio 10 — portfolio 15). Tables 1.9a through 1.10c show conditional probabilities of individual stock returns in each portfolio for three different kinds of daily and weekly portfolio returns, conditioned on the previous period's individual returns (positive, negative, or zero), market return in the same period (positive or negative), and the current period's market return (positive or negative). These conditions will be denoted as $X = (R_{t-1}, M_{t-1}, M_t)$. Now let's consider daily and weekly size-sorted portfolios (table 1.9a and 1.10a).

The first thing to be noted is that for all daily portfolio returns, the conditional probability of a zero return is higher when the previous period return is zero than when it is positive or negative. This suggests that it takes more than one day for information to be completely and correctly absorbed in individual stocks, unless there is truly a high probability of consecutive insignificant information flows. As theory suggests (i.e. Capital Asset Pricing Model), the contemporaneous movement of a market (M_t) seems to be the most important condition for individual returns in all portfolios. Other than M_t, a relative movement of R_{t-1} and M_{t-1} also seems to be important, especially for small sized portfolios. For example, for all size-sorted portfolios, the conditional probability of a negative return ($\Pr(R_t < 0 \mid X)$) is larger when ($R_{t-1} = n, M_{t-1} = n$) than ($R_{t-1} = n, M_{t-1} = p$) given the same movement of M_t. Similarly, the conditional probability of a positive return ($\Pr(R_t > 0 \mid X)$)

is larger when $(R_{t-1} = p, M_{t-1} = p)$ than $(R_{t-1} = p, M_{t-1} = n)$ given the same movement of M_t. As a result, there seems to be a tendency for the market to have a lagged impact on individual returns when they have not moved together with the market in the previous period. Conditional probabilities for price- and number-of-shares sorted portfolios also show similar patterns. Finally, table 1.11 and figure 1.1 and 1.2 show a distribution of contemporaneous correlations for daily and weekly size-sorted portfolio returns. Distributions for price- and number- of-shares-sorted portfolio returns are not separately shown since they are almost the same as those for size-sorted portfolios. Contemporaneous correlations for weekly returns are more concentrated on a certain interval than daily returns — distributions are shifting to the right (see figure 1.1 and 1.2). As a portfolio size increases, returns are more positively contemporaneously correlated. However, contemporaneous correlations are usually less than 0.5, which implies large idiosyncratic risks even for large stocks.

5 Summary and Conclusions

The impact of discrete asset prices on serial and cross-sectional autocorrelations were investigated. Under a hypothetical price process, it was found that price discreteness could cause a relatively large spurious serial autocorrelation when asset prices are low (i.e. less than \$5 or \$10) and when standard deviations of asset returns are low (i.e. less than 0.04 or 0.05). Cross-sectional autocorrelations seemed not to be affected by price discreteness. This implies that empirically documented lead-lag relationships among individual stock returns are not spurious. It is strongly conjectured that there must be some economic reasons which cause individual stocks to be cross-sectionally autocorrelated.

Three kinds of portfolios — size-, price-, and number of shares-sorted portfolios — were empirically investigated. Price discreteness doesn't seem to have an impact on portfolio autocorrelations, which is consistent with the theoretical investigation about price discreteness. Nonsynchronous trading has a relatively large impact on daily portfolio returns, but the impact becomes insignificant for weekly portfolio returns. A large portion of positive autocorrelations seem to arise from the lead-lag relationships between stocks in two groups; one with positive returns and the other with negative returns. It seems to take more than one day for information to be fully and correctly absorbed in stocks since the conditional probability of a zero return is higher when the previous return was zero than when it was non-zero. In other words, when there are non-zero returns one day, there is a tendency for these returns to remain non-zero the following day. Although the market seems to be the most important factor that governs the movement of contemporaneous individual returns, a movement of the previous individual returns relative to the previous market return also seems to be important, especially for small stocks. Finally, the distribution of contemporaneous correlations among individual stock returns is shifting to the right as a portfolio's size increases. Idiosyncratic risks seem to be substantial even for large stocks since in most cases, contemporaneous correlations are less than 0.5.

This paper attempted to investigate the behavior of individual returns in each portfolio as a first step toward an important and ultimate goal in the research of the short-horizon lead-lag relationships among individual stock returns. The next step, of course, would be an attempt to identify economic reasons which drive the lead-lag relationships. For various reasons, information may be absorbed in each stock at different speeds. Nevertheless, it is still puzzling why the market allows it to be absorbed with a lag of more than a week in some cases, for certain stocks. The stock market may be less informationally efficient than was expected, especially for small stocks. However, these are just conjectures which have be verified in the future research.

References

[1] Ball, Clifford (1988), "Estimation Bias Induced by Discrete Security Prices," Journal of Finance 43, 841-865.

[2] Chan, K.C. and N. Chen (1991), "Structural and Return Characteristics of Small and Large Firms," Journal of Finance 46, 1467- 1484.

[3] Cho, D.C. and E.W. Frees (1988), "Estimating the Volatility of Discrete Stock Prices," Journal of Finance 43, 451-466.

[4] Gottlieb, G. and A. Kalay (1985), " Implications of the Discreteness of Observed Stock Prices," Journal of Finance 40, 135-153.

[5] Harris, Lawrence (1990), "Estimation of Stock Price Variances and Serial Covariances from Discrete Observations," Journal of Financial and Quantitative Analysis 25, 291-306.

[6] Hausman, J., A. Lo, and C. MacKinlay (1992), "An Ordered Probit Analysis of Transaction Stock Prices," Journal of Financial Economics 31, 319-379.

[7] Lehmann, Bruce N. (1990), "Fads, Martingales, and Market Efficiency," Quarterly Journal of Economics 105, 1-28.

[8] Leroy, S.F. (1973), "Risk Aversion and the Martingale Property of Stock Returns," International Economic Review 14, 436-446.

[9] Lo, A. and C. MacKinlay (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," Review of Financial Studies 1, 41-66.

[10] Lo, A. and C. MacKinlay (1990a), "An Economic Analysis of Nonsynchronous Trading?," Journal of Economic Theory 45, 181-211.

[11] Lo, A. and C. MacKinlay (1990b), "When Are Contrarian Profits Due To Stock Market Overreaction?," Review of Financial Studies 3, 175-205.

[12] Shin, Sung-Hwan (1992), "Economic Implications of Asset Return Predictability: Stochastic Dominance Comparisons," Working Paper, MIT.

Table 1.1
Mean and Standard Deviation of Returns for Discrete Prices

To approximate the theorem 1, the quadrature method of D01FBF with D01BBF–D01BAZ in the NAG Fortran Library is used. The upper and the lower bounds for returns are set as 100% and -100% respectively. μ stands for the true mean and σ stands for the true standard deviation. The 'tick' is assumed to be 1/8 and P stands for the price of an asset. Standard deviations for $\mu = 0.003$ are not shown in this table since they are almost the same with those for $\mu = 0.0006$.

σ	$P = 2.5$	$P = 5.0$	$P = 10.0$	$P = 20.0$
		Mean		
$\mu = 0.0006$				
0.09	0.000390	0.000548	0.000587	0.000597
0.07	0.000391	0.000548	0.000587	0.000597
0.06	0.000391	0.000548	0.000587	0.000597
0.05	0.000391	0.000548	0.000587	0.000597
0.04	0.000391	0.000548	0.000587	0.000597
0.03	0.000393	0.000548	0.000587	0.000597
0.02	0.000431	0.000548	0.000587	0.000597
0.01	0.000550	0.000558	0.000587	0.000597
$\mu = 0.003$				
0.09	0.002790	0.002948	0.002987	0.002997
0.07	0.002791	0.002948	0.002987	0.002997
0.06	0.002791	0.002948	0.002987	0.002997
0.05	0.002792	0.002948	0.002987	0.002997
0.04	0.002792	0.002948	0.002987	0.002997
0.03	0.002794	0.002948	0.002987	0.002997
0.02	0.002831	0.002948	0.002987	0.002997
0.01	0.002946	0.002956	0.002987	0.002997
		STD		
$\mu = 0.0006$				
0.09	0.092323	0.090586	0.090147	0.090037
0.07	0.072944	0.070747	0.070187	0.070047
0.06	0.063402	0.060868	0.060218	0.060055
0.05	0.054026	0.051036	0.050261	0.050065
0.04	0.044923	0.041285	0.040325	0.040082
0.03	0.036299	0.031691	0.030431	0.030108
0.02	0.028401	0.022455	0.020641	0.020162
0.01	0.019971	0.014192	0.011226	0.010320

Table 1.2
[Sample Variance/True Variance] Ratios

To approximate the theorem 1, the quadrature method of D01FBF with D01BBF–D01BAZ in the NAG Fortran Library is used. The upper and the lower bounds for returns are set as 100% and -100% respectively. μ stands for the true mean and σ stands for the true standard deviation. The 'tick' is assumed to be 1/8 and P stands for the price of an asset.

σ	$P = 2.5$	$P = 5.0$	$P = 10.0$	$P = 20.0$
$\mu = 0.0006$				
0.09	1.052	1.013	1.003	1.001
0.07	1.086	1.021	1.005	1.001
0.06	1.117	1.029	1.007	1.002
0.05	1.168	1.042	1.010	1.003
0.04	1.261	1.065	1.016	1.004
0.03	1.464	1.116	1.029	1.007
0.02	2.016	1.261	1.065	1.016
0.01	3.989	2.014	1.260	1.065
$\mu = 0.003$				
0.09	1.052	1.013	1.003	1.001
0.07	1.086	1.021	1.005	1.001
0.06	1.116	1.029	1.007	1.002
0.05	1.167	1.042	1.010	1.003
0.04	1.261	1.065	1.016	1.004
0.03	1.463	1.116	1.029	1.007
0.02	2.011	1.260	1.065	1.016
0.01	4.036	2.014	1.260	1.065

Table 1.3
Spurious Autocorrelations for Individual Assets

To approximate the theorem 1 and theorem 2, the quadrature method of D01FBF with D01BBF-D01BAZ in the NAG Fortran Library is used. The upper and the lower bounds for returns are set as 100% and -100% respectively. μ stands for the true mean and σ stands for the true standard deviation. The 'tick' is assumed to be 1/8 and P stands for the price of an asset.

Unit=%

σ	$P = 2.5$	$P = 5.0$	$P = 10.0$	$P = 20.0$
$\mu = 0.0006$				
0.09	-2.48	-0.64	-0.16	-0.04
0.07	-3.95	-1.05	-0.27	-0.07
0.06	-5.22	-1.41	-0.36	-0.09
0.05	-7.17	-2.01	-0.52	-0.13
0.04	-10.35	-3.06	-0.80	-0.20
0.03	-15.83	-5.19	-1.41	-0.36
0.02	-24.53	-10.32	-3.05	-0.80
0.01	-29.30	-24.51	-10.32	-3.05
$\mu = 0.003$				
0.09	-2.46	-0.64	-0.16	-0.04
0.07	-3.92	-1.04	-0.26	-0.07
0.06	-5.18	-1.40	-0.36	-0.09
0.05	-7.12	-1.99	-0.51	-0.13
0.04	-10.28	-3.04	-0.80	-0.20
0.03	-15.72	-5.15	-1.40	-0.36
0.02	-24.40	-10.25	-3.03	-0.79
0.01	-29.25	-24.55	-10.25	-3.03

Table 1.4
Spurious Cross-Sectional Autocorrelations for Individual Assets

To approximate the theorem 1, theorem 3, the lemma 3, the quadrature method of D01FBF with D01BBF-D01BAZ in the NAG Fortran Library is used. The upper and the lower bounds for returns are set as 100% and -100% respectively. The indefinite integral for a Gaussian random variable with mean α and variance β is modified as a definite integral whose range is $[\alpha + 4\beta, \alpha - 4\beta]$. Each asset returns are assumed to have two random components; one common factor and one idiosyncratic factor. μ stands for the true mean and σ stands for the true total standard deviation of returns. σ_λ stands for the true standard deviation of the common factor. The 'tick' is assumed to be 1/8 and P stands for the price of an asset.

Unit=%

$\mu = 0.0006, \sigma_\lambda = 0.009$	$R_{t+1}(P = 2.5)$	$R_{t+1}(P = 5)$	$R_{t+1}(P = 10)$	$R_{t+1}(P = 20)$
$\sigma = 0.06$				
$R_t(P = 2.5)$	-5.220	0.010	0.010	0.010
$R_t(P = 5)$	0.007	-1.410	0.008	0.007
$R_t(P = 10)$	0.007	0.008	-0.360	0.007
$R_t(P = 20)$	0.007	0.007	0.007	-0.090
$\sigma = 0.01$				
$R_t(P = 2.5)$	-29.300	0.139	0.124	0.132
$R_t(P = 5)$	0.139	-24.510	0.171	0.182
$R_t(P = 10)$	0.124	0.171	-10.320	0.216
$R_t(P = 20)$	0.132	0.182	0.216	-3.050
$\mu = 0.003, \sigma_\lambda = 0.02$	$R_{t+1}(P = 2.5)$	$R_{t+1}(P = 5)$	$R_{t+1}(P = 10)$	$R_{t+1}(P = 20)$
$\sigma = 0.09$				
$R_t(P = 2.5)$	-2.460	0.007	0.007	0.008
$R_t(P = 5.0)$	0.001	-0.640	0.001	0.002
$R_t(P = 10.0)$	0.000	0.000	-0.160	0.001
$R_t(P = 20.0)$	0.000	0.000	0.001	-0.040
$\sigma = 0.03$				
$R_t(P = 2.5)$	-15.720	0.049	0.050	0.058
$R_t(P = 5)$	0.014	-5.150	0.010	0.017
$R - t(P = 10)$	0.005	-0.001	-1.400	0.005
$R - t(P = 20)$	0.003	-0.004	0.005	-0.360

Table 1.5a
Sample Statistics for Portfolios
(Daily)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. Portfolio mean returns are not stated in the table since they are essentially the same with average means returns for individual stocks.

	Size-Sorted	Price-Sorted	Shares-Sorted
P1 (smallest)			
Mean	0.00197	0.00232	0.00092
Avg Individual Std	0.04575	0.05591	0.02653
Portfolio Std	0.01053	0.01220	0.00732
P5 (5th smallest)			
Mean	0.00074	0.00063	0.00074
Avg Individual Std	0.02800	0.02719	0.02743
Portfolio Std	0.00908	0.00942	0.00878
P10 (10th smallest)			
Mean	0.00059	0.00060	0.00068
Avg Individual Std	0.02226	0.02065	0.02452
Portfolio Std	0.00882	0.00824	0.00899
P15 (15th smallest)			
Mean	0.00056	0.00054	0.00063
Avg Individual Std	0.01870	0.01769	0.02186
Portfolio Std	0.00840	0.00786	0.00888
P20 (largest)			
Mean	0.00043	0.00054	0.00046
Avg Individual Std	0.01472	0.01520	0.01550
Portfolio Std	0.00894	0.00849	0.00846

Table 1.5b
Sample Statistics for Portfolios
(Weekly Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year and which had data missing no more than 10 data points during the year. Portfolio mean returns are not stated in the table since they are essentially the same with average means returns for individual stocks.

	Size-Sorted	Price-Sorted	Shares-Sorted
P1 (smallest)			
Mean	0.00459	0.00443	0.00303
Avg Individual Std	0.09047	0.10066	0.06298
Portfolio Std	0.02833	0.03305	0.02406
P5 (5th smallest)			
Mean	0.00234	0.00220	0.00247
Avg Individual Std	0.06081	0.06107	0.06107
Portfolio Std	0.02382	0.02705	0.02594
P10 (10th smallest)			
Mean	0.00204	0.00219	0.00214
Avg Individual Std	0.05110	0.04874	0.05518
Portfolio Std	0.02312	0.02343	0.02603
P15 (15th smallest)			
Mean	0.00214	0.00202	0.00225
Avg Individual Std	0.04356	0.04285	0.04913
Portfolio Std	0.02184	0.02200	0.02449
P20 (largest)			
Mean	0.00193	0.00235	0.00182
Avg Individual Std	0.03510	0.04368	0.03536
Portfolio Std	0.02115	0.02438	0.02038

Table 1.6
First-Order Autocorrelations for Portfolio Returns[1]

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. For daily portfolio returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. For weekly portfolio returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year and which had data missing no more than 10 data points during the year.

	(Daily)			(Weekly)		
	Size-S	Price-S	Shares-S	Size-S	Price-S	Shares- S
P1 (smallest)	0.29	0.26	0.39	0.35	0.24	0.23
P2	0.33	0.31	0.37	0.29	0.22	0.22
P3	0.33	0.33	0.34	0.29	0.19	0.20
P4	0.32	0.31	0.33	0.26	0.18	0.18
P5	0.32	0.31	0.33	0.24	0.17	0.19
P6	0.32	0.31	0.32	0.23	0.17	0.18
P7	0.31	0.32	0.32	0.23	0.13	0.15
P8	0.30	0.31	0.32	0.23	0.14	0.16
P9	0.30	0.30	0.31	0.22	0.13	0.15
P10	0.31	0.30	0.30	0.21	0.15	0.13
P11	0.29	0.31	0.28	0.22	0.12	0.13
P12	0.30	0.31	0.28	0.20	0.13	0.12
P13	0.31	0.31	0.28	0.19	0.12	0.10
P14	0.29	0.30	0.27	0.17	0.11	0.12
P15	0.28	0.30	0.26	0.16	0.11	0.10
P16	0.27	0.31	0.24	0.13	0.08	0.08
P17	0.26	0.31	0.23	0.13	0.09	0.07
P18	0.24	0.28	0.21	0.09	0.07	0.06
P19	0.22	0.28	0.19	0.06	0.05	0.05
P20 (largest)	0.15	0.24	0.15	0.01	0.03	0.00

[1]Complete matrices for cross-sectional autocorrelations are also available upon request.

Table 1.7
Portfolio Characteristics

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. For daily returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. For weekly portfolio returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year and which had data missing no more than 10 data points during the year. All numbers are based on portfolios for daily returns unless 'weekly' is specified. Portfolio characteristics for weekly returns are almost the same with those for daily returns.

	Size-Sorted	Price-Sorted	Shares-Sorted
P1 (smallest)			
Prob(Zero Return) – Daily	0.368	0.412	0.284
Prob(Zero Return) – Weekly	0.046	0.063	0.028
Avg Prices ($)	4.2	2.0	19.9
Avg Market Values ($MM)	3.8	15.5	15.7
P5 (5th smallest)			
Prob(Zero Return) – Daily	0.268	0.271	0.251
Prob(Zero Return) – Weekly	0.017	0.017	0.015
Avg Prices ($)	12.5	9.9	16.0
Avg Market Values ($MM)	24.9	61.5	36.0
P10 (10th smallest)			
Prob(Zero Return) – Daily	0.207	0.203	0.219
Prob(Zero Return) – Weekly	0.007	0.008	0.013
Avg Prices ($)	20.7	18.5	22.0
Avg Market Values ($MM)	83.4	218.6	108.9
P15 (15th smallest)			
Prob(Zero Return) – Daily	0.162	0.153	0.185
Prob(Zero Return) – Weekly	0.003	0.005	0.011
Avg Prices ($)	31.0	29.9	27.9
Avg Market Values ($MM)	324.6	619.4	372.2
P20 (largest)			
Prob(Zero Return) – Daily	0.105	0.087	0.146
Prob(Zero Return) – Weekly	0.001	0.003	0.005
Avg Prices ($)	62.3	88.5	42.3
Avg Market Values ($MM)	7,445.3	4,333.7	6,962.0

Table 1.8a
Decomposition of First-Order Autocorrelations
(Daily)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. Every day, for each portfolio, all stocks are classified into three groups depending on whether returns are positive (G_{pt}), negative (G_{nt}) or zero(G_{zt}). P–Ng, for example, stands for $Avg(\frac{N_{pt}N_{nt}}{N_t^2}(Avg(R_{it}R_{jt+1}))$, where N_{pt} and N_{nt} are the number of stocks in G_{pt} and G_{nt} and the first average is over time and the second average is over $i \in G_{pt}$ and $j \in G_{nt}$. Similarly, Z stands for stocks with zero returns.

	Port 1 (Smallest)	Port 5	Port 10	Port 15	Port 20 (Largest)
Size-Sort					
P–P	-0.272	0.072	0.139	0.156	0.116
Ng–Ng	-0.526	-0.045	0.073	0.105	0.054
P–Ng, Ng–P	0.956	0.210	0.050	-0.008	-0.033
P–Z, Z–P	0.267	0.067	0.042	0.034	0.013
Ng–Z, Z–Ng	-0.008	0.081	0.057	0.025	0.008
TOTAL	0.418	0.385	0.361	0.312	0.158
Price-Sort					
P–P	-0.512	0.071	0.142	0.178	0.171
Ng–Ng	-0.747	-0.018	0.077	0.129	0.128
P–Ng, Ng–P	1.368	0.170	0.035	-0.032	-0.064
P–Z, Z–P	0.359	0.065	0.052	0.036	0.018
Ng–Z, Z–Ng	-0.074	0.092	0.044	0.028	0.007
TOTAL	0.394	0.381	0.349	0.338	0.260
Shares-Sort					
P–P	0.151	0.078	0.088	0.105	0.103
Ng–Ng	-0.032	-0.018	0.011	0.038	0.027
P–Ng, Ng–P	0.185	0.187	0.137	0.071	-0.003
P–Z, Z–P	0.073	0.090	0.073	0.055	0.021
Ng–Z, Z–Ng	0.106	0.071	0.056	0.042	0.020
TOTAL	0.484	0.408	0.365	0.311	0.168

Table 1.8b
Decomposition of First-Order Autocorrelations
(Weekly)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year with data having less than 10 missing data points. Every week, for each portfolio, all stocks are classified into three groups depending on whether returns are positive (G_{pt}), negative (G_{nt}) or zero(G_{zt}). P–Ng, for example, stands for $Avg(\frac{N_{pt}N_{nt}}{N_t^2}(Avg(R_{it}R_{jt+1}))$ is calculated, where N_{pt} and N_{nt} are the number of stocks in G_{pt} and G_{nt} and the first average is over time and the second average is over $i \in G_{pt}$ and $j \in G_{nt}$. Similarly, Z stands for stocks with zero returns.

	Port 1 (Smallest)	Port 5	Port 10	Port 15	Port 20 (Largest)
Size-Sort					
P–P	0.038	0.079	0.105	0.095	0.035
Ng–Ng	-0.097	0.030	0.048	0.016	-0.068
P–Ng, Ng–P	0.463	0.198	0.101	0.074	0.033
P–Z, Z–P	0.020	0.013	0.002	0.000	0.000
Ng–Z, Z–Ng	0.011	0.000	0.001	0.001	0.000
TOTAL	0.434	0.319	0.256	0.185	0.000
Price-Sort					
P–P	-0.060	0.049	0.072	0.081	0.056
Ng–Ng	-0.158	0.001	0.016	0.011	-0.032
P–Ng, Ng–P	0.488	0.142	0.095	0.059	0.022
P–Z, Z–P	0.032	0.007	0.002	0.000	0.000
Ng–Z, Z–Ng	0.007	0.001	0.000	0.000	0.000
TOTAL	0.309	0.200	0.185	0.152	0.047
Shares-Sort					
P–P	0.096	0.054	0.046	0.054	0.030
Ng–Ng	0.000	0.003	0.001	-0.011	-0.060
P–Ng, Ng–P	0.167	0.155	0.113	0.087	0.038
P–Z, Z–P	0.002	0.003	0.003	0.003	0.000
Ng–Z, Z–Ng	0.007	0.005	0.003	0.000	0.001
TOTAL	0.271	0.220	0.166	0.133	0.009

Table 1.9a
Conditional Probabilities for Individual Returns
(Size-Sorted Portfolios — Daily Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	8	7	6	9	8	7	7	12	6	5	7	11
$\Pr(R_t < 0 \vert X)$	34	24	27	22	34	28	30	26	50	42	46	39
$\Pr(R_t = 0 \vert X)$	34	34	35	34	45	45	45	45	29	31	29	29
$\Pr(R_t > 0 \vert X)$	32	42	38	44	21	27	25	30	21	27	25	32
P5												
Pr(X)	11	9	6	10	6	5	5	8	6	6	8	13
$\Pr(R_t < 0 \vert X)$	46	31	37	28	43	33	38	29	54	41	48	37
$\Pr(R_t = 0 \vert X)$	25	26	27	25	35	36	35	34	21	23	22	22
$\Pr(R_t > 0 \vert X)$	29	43	36	47	22	32	27	37	25	36	30	41
P10												
Pr(X)	12	10	7	10	4	4	4	6	7	6	9	15
$\Pr(R_t < 0 \vert X)$	54	33	44	31	50	33	41	30	56	39	50	35
$\Pr(R_t = 0 \vert X)$	19	20	21	20	27	29	28	27	17	19	18	17
$\Pr(R_t > 0 \vert X)$	27	47	36	50	23	38	30	42	27	42	33	48
P15												
Pr(X)	13	11	7	10	3	3	3	5	7	6	10	16
$\Pr(R_t < 0 \vert X)$	58	34	49	32	53	33	45	30	57	37	51	32
$\Pr(R_t = 0 \vert X)$	15	16	16	16	23	24	24	22	14	15	14	14
$\Pr(R_t > 0 \vert X)$	27	50	35	52	25	43	31	48	28	48	35	54
P20 (largest)												
Pr(X)	15	13	6	9	2	2	2	3	6	5	12	19
$\Pr(R_t < 0 \vert X)$	66	29	62	32	61	30	58	29	62	29	61	29
$\Pr(R_t = 0 \vert X)$	9	10	11	11	15	16	15	15	10	11	9	9
$\Pr(R_t > 0 \vert X)$	25	60	28	57	24	54	27	55	28	60	30	62

Table 1.9b
Conditional Probabilities for Individual Returns
(Price-Sorted Portfolios — Daily Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	7	6	5	8	10	9	9	14	6	5	6	9
$\Pr(R_t < 0 \mid X)$	23	17	18	16	31	25	27	23	52	45	47	42
$\Pr(R_t = 0 \mid X)$	43	42	43	41	49	50	50	49	32	35	34	35
$\Pr(R_t > 0 \mid X)$	33	41	38	43	21	25	23	28	16	20	19	23
P5												
Pr(X)	10	9	6	10	6	6	6	9	6	5	8	12
$\Pr(R_t < 0 \mid X)$	44	29	36	27	44	32	38	29	54	41	49	37
$\Pr(R_t = 0 \mid X)$	28	29	29	27	34	36	35	35	23	26	24	24
$\Pr(R_t > 0 \mid X)$	28	43	35	46	21	32	27	36	23	34	27	38
P10												
Pr(X)	12	10	7	10	5	4	4	6	7	6	9	15
$\Pr(R_t < 0 \mid X)$	52	32	43	31	49	33	42	30	56	39	50	35
$\Pr(R_t = 0 \mid X)$	20	20	21	20	26	29	28	27	18	20	18	18
$\Pr(R_t > 0 \mid X)$	28	47	36	49	24	38	31	43	26	41	31	47
P15												
Pr(X)	13	11	7	10	3	3	3	5	7	6	10	16
$\Pr(R_t < 0 \mid X)$	58	35	49	33	52	35	45	31	56	37	51	33
$\Pr(R_t = 0 \mid X)$	15	16	16	15	22	22	22	21	14	16	14	14
$\Pr(R_t > 0 \mid X)$	28	50	35	52	26	43	33	48	29	47	35	53
P20 (largest)												
Pr(X)	15	12	6	10	2	1	1	3	7	6	12	19
$\Pr(R_t < 0 \mid X)$	67	34	60	35	59	32	52	31	59	31	57	29
$\Pr(R_t = 0 \mid X)$	8	9	9	9	16	17	15	16	9	9	8	8
$\Pr(R_t > 0 \mid X)$	25	57	31	56	26	51	32	54	32	59	35	63

Table 1.9c
Conditional Probabilities for Individual Returns
(Shares-Sorted Portfolios — Daily Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	10	8	6	10	6	6	6	9	7	6	8	12
$\Pr(R_t < 0 \mid X)$	45	33	36	30	39	31	33	28	51	40	45	36
$\Pr(R_t = 0 \mid X)$	24	25	26	25	39	40	40	39	23	24	23	23
$\Pr(R_t > 0 \mid X)$	31	42	37	45	22	29	27	33	26	36	32	41
P5												
Pr(X)	11	9	6	10	6	5	5	8	7	6	8	13
$\Pr(R_t < 0 \mid X)$	48	32	39	29	44	33	37	29	53	39	48	36
$\Pr(R_t = 0 \mid X)$	23	24	24	24	34	34	34	33	21	23	20	20
$\Pr(R_t > 0 \mid X)$	29	44	37	47	23	33	29	38	26	38	32	43
P10												
Pr(X)	12	10	7	10	5	4	4	7	6	6	9	15
$\Pr(R_t < 0 \mid X)$	53	33	43	30	47	33	40	29	56	39	49	35
$\Pr(R_t = 0 \mid X)$	20	21	22	20	30	32	31	30	18	20	18	18
$\Pr(R_t > 0 \mid X)$	27	47	35	50	23	36	29	41	26	42	33	48
P15												
Pr(X)	12	11	6	10	4	3	3	6	6	6	10	16
$\Pr(R_t < 0 \mid X)$	56	32	47	31	50	32	44	29	58	37	52	33
$\Pr(R_t = 0 \mid X)$	17	18	18	18	27	29	27	27	16	17	15	15
$\Pr(R_t > 0 \mid X)$	27	50	34	52	23	40	29	44	27	47	33	52
P20 (largest)												
Pr(X)	14	12	6	9	3	3	3	4	6	5	11	18
$\Pr(R_t < 0 \mid X)$	61	28	57	30	56	30	51	29	61	31	59	29
$\Pr(R_t = 0 \mid X)$	13	13	15	15	22	23	22	22	14	15	13	13
$\Pr(R_t > 0 \mid X)$	25	59	29	55	23	47	27	50	26	55	28	58

Table 1.10a
Conditional Probabilities for Individual Returns
(Size-Sorted Portfolios — Weekly Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year with data having less than 10 missing data points. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	13	12	15	18	0	1	1	1	6	9	7	14
$\Pr(R_t < 0 \mid X)$	68	53	58	50	62	53	54	48	73	62	67	57
$\Pr(R_t = 0 \mid X)$	4	5	5	5	18	17	18	17	3	4	4	3
$\Pr(R_t > 0 \mid X)$	28	42	36	45	20	30	28	36	24	34	30	40
P5												
Pr(X)	14	16	11	17	0	0	0	0	6	7	11	17
$\Pr(R_t < 0 \mid X)$	73	49	62	47	65	51	60	46	74	56	66	51
$\Pr(R_t = 0 \mid X)$	2	2	2	2	13	11	13	10	1	2	1	1
$\Pr(R_t > 0 \mid X)$	25	48	36	51	23	38	28	44	25	42	32	48
P10												
Pr(X)	14	16	10	15	0	0	0	0	6	7	12	19
$\Pr(R_t < 0 \mid X)$	73	45	64	44	70	52	54	46	73	49	66	46
$\Pr(R_t = 0 \mid X)$	1	1	1	1	5	8	12	8	1	1	1	1
$\Pr(R_t > 0 \mid X)$	25	54	35	55	25	40	33	46	26	51	33	53
P15												
Pr(X)	14	16	9	14	0	0	0	0	6	7	13	20
$\Pr(R_t < 0 \mid X)$	72	40	64	40	69	50	64	43	72	44	67	41
$\Pr(R_t = 0 \mid X)$	1	1	1	1	8	7	7	4	0	1	0	1
$\Pr(R_t > 0 \mid X)$	27	59	35	59	23	43	29	52	27	55	33	58
P20 (largest)												
Pr(X)	14	17	8	12	0	0	0	0	5	6	15	22
$\Pr(R_t < 0 \mid X)$	74	31	69	37	75	33	71	45	75	36	73	37
$\Pr(R_t = 0 \mid X)$	0	0	0	0	0	0	0	0	0	0	0	0
$\Pr(R_t > 0 \mid X)$	26	69	30	63	25	67	29	55	25	64	27	63

Table 1.10b
Conditional Probabilities for Individual Returns
(Price-Sorted Portfolios — Weekly Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year with data having less than 10 missing data points. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	14	15	12	18	1	1	1	2	5	6	8	13
$Pr(R_t < 0 \mid X)$	69	52	61	51	59	50	55	43	76	66	73	62
$Pr(R_t = 0 \mid X)$	5	7	6	6	21	22	21	26	3	5	4	4
$Pr(R_t > 0 \mid X)$	26	41	33	43	21	28	24	30	20	30	23	34
P5												
Pr(X)	14	16	11	15	0	0	0	0	5	6	11	16
$Pr(R_t < 0 \mid X)$	74	46	65	46	71	48	58	32	75	53	70	51
$Pr(R_t = 0 \mid X)$	1	2	2	2	8	9	8	31	1	2	1	1
$Pr(R_t > 0 \mid X)$	25	52	34	52	21	42	33	37	24	45	29	48
P10												
Pr(X)	14	15	10	14	0	0	0	0	6	7	12	18
$Pr(R_t < 0 \mid X)$	73	41	64	42	67	50	64	21	71	48	69	45
$Pr(R_t = 0 \mid X)$	0	1	1	1	0	8	7	45	0	1	0	0
$Pr(R_t > 0 \mid X)$	27	58	36	58	33	42	29	33	28	51	31	54
P15												
Pr(X)	14	15	9	13	0	0	0	0	6	7	13	20
$Pr(R_t < 0 \mid X)$	73	37	64	40	60	40	67	14	71	42	67	41
$Pr(R_t = 0 \mid X)$	0	1	0	0	20	0	0	71	0	1	0	0
$Pr(R_t > 0 \mid X)$	27	62	36	60	20	60	33	14	29	57	33	59
P20 (largest)												
Pr(X)	14	16	7	11	0	0	0	0	5	6	15	22
$Pr(R_t < 0 \mid X)$	73	31	68	36	0	50	0	6	71	33	71	35
$Pr(R_t = 0 \mid X)$	0	0	0	0	0	0	0	82	0	0	0	0
$Pr(R_t > 0 \mid X)$	27	69	32	64	0	50	0	12	29	67	29	64

Table 1.10c
Conditional Probabilities for Individual Returns
(Shares-Sorted Portfolios — Weekly Returns)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. Three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year with data having less than 10 missing data points. 'n', 'p', and 'z' mean negative, positive, and zero respectively. Each column stands for probabilities that are related to an event for (R_{t-1}, M_{t-1}, M_t), where R_t is the return for individual stocks at time t, and M_t is the return for value-weighted market index at time t. 'zpn', for example stands for an event of $(R_{t-1} = 0, M_{t-1} > 0, M_t < 0)$.

Event X = (R_{t-1}, M_{t-1}, M_t)

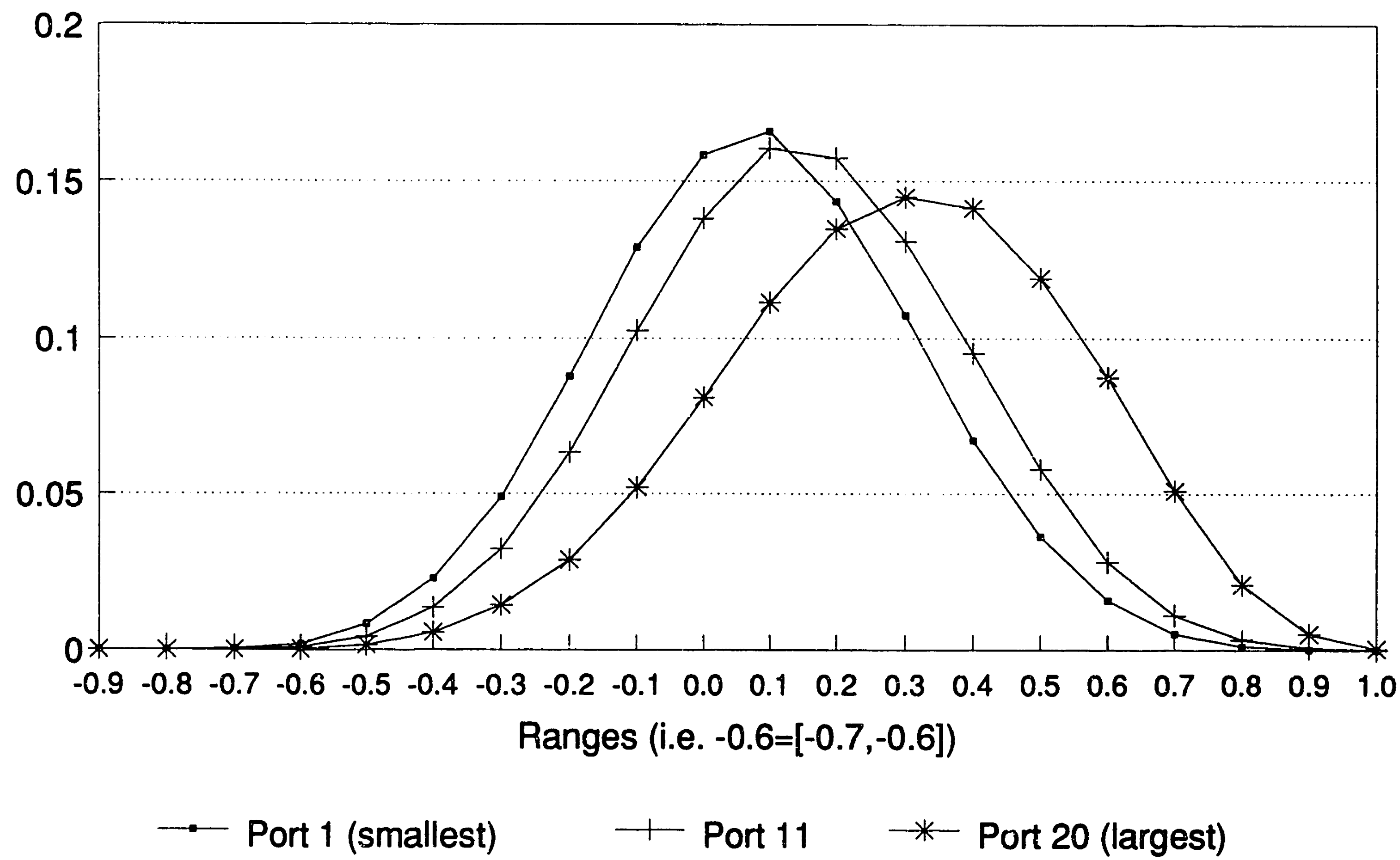
(%)	nnn	nnp	npn	npp	znn	znp	zpn	zpp	pnn	pnp	ppn	ppp
P1 (smallest)												
Pr(X)	13	14	11	16	0	0	0	0	6	7	10	16
$\Pr(R_t < 0 \mid X)$	70	49	60	47	58	49	56	42	70	53	64	50
$\Pr(R_t = 0 \mid X)$	2	3	3	3	18	20	19	28	2	3	2	2
$\Pr(R_t > 0 \mid X)$	28	47	37	50	24	31	25	29	28	44	34	48
P5												
Pr(X)	14	15	10	15	0	0	0	0	6	6	11	17
$\Pr(R_t < 0 \mid X)$	73	46	62	45	70	50	64	38	73	52	68	49
$\Pr(R_t = 0 \mid X)$	1	2	2	1	13	13	9	31	1	2	1	1
$\Pr(R_t > 0 \mid X)$	26	52	36	53	17	37	27	31	26	46	31	50
P10												
Pr(X)	14	15	10	14	0	0	0	0	6	6	12	18
$\Pr(R_t < 0 \mid X)$	73	42	65	43	68	48	62	29	74	47	68	45
$\Pr(R_t = 0 \mid X)$	1	2	1	1	16	16	15	37	1	1	1	1
$\Pr(R_t > 0 \mid X)$	26	57	34	56	16	36	23	33	26	52	32	54
P15												
Pr(X)	14	16	9	13	0	0	0	0	5	6	13	19
$\Pr(R_t < 0 \mid X)$	73	38	66	40	73	45	57	32	73	44	69	42
$\Pr(R_t = 0 \mid X)$	1	1	1	1	9	15	14	43	1	1	0	1
$\Pr(R_t > 0 \mid X)$	27	61	33	59	18	40	29	26	27	55	31	57
P20 (largest)												
Pr(X)	15	16	8	12	0	0	0	0	5	6	14	21
$\Pr(R_t < 0 \mid X)$	73	33	69	38	80	50	56	16	74	37	73	39
$\Pr(R_t = 0 \mid X)$	0	1	0	0	0	12	11	64	0	1	0	0
$\Pr(R_t > 0 \mid X)$	27	66	31	62	20	37	33	20	26	63	26	61

Table 1.11
Contemporaneous Correlations among Individual Returns
(Size-Sorted Portfolios, See Figure 1.1 and 1.2.)

Data are from CRSP daily return file for NYSE and AMEX stocks from the period of 1962.8.1 to 1990.12.31. For daily returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every month based on the previous month ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the month with no missing data point. For weekly portfolio returns, three kinds of twenty equally-weighted portfolios were formed and rebalanced every year based on the previous year's ending market values, prices, and number of shares outstanding from stocks which were continuously listed in NYSE and AMEX during the year and which had data missing no more than 10 data points during the year. For each portfolio, contemporaneous correlations among individual returns are calculated (= $N(N-1)$) and the number of them which belong to a certain range are counted (=N_a). Numbers in the table stands for $\frac{N_a}{N(N-1)}$.

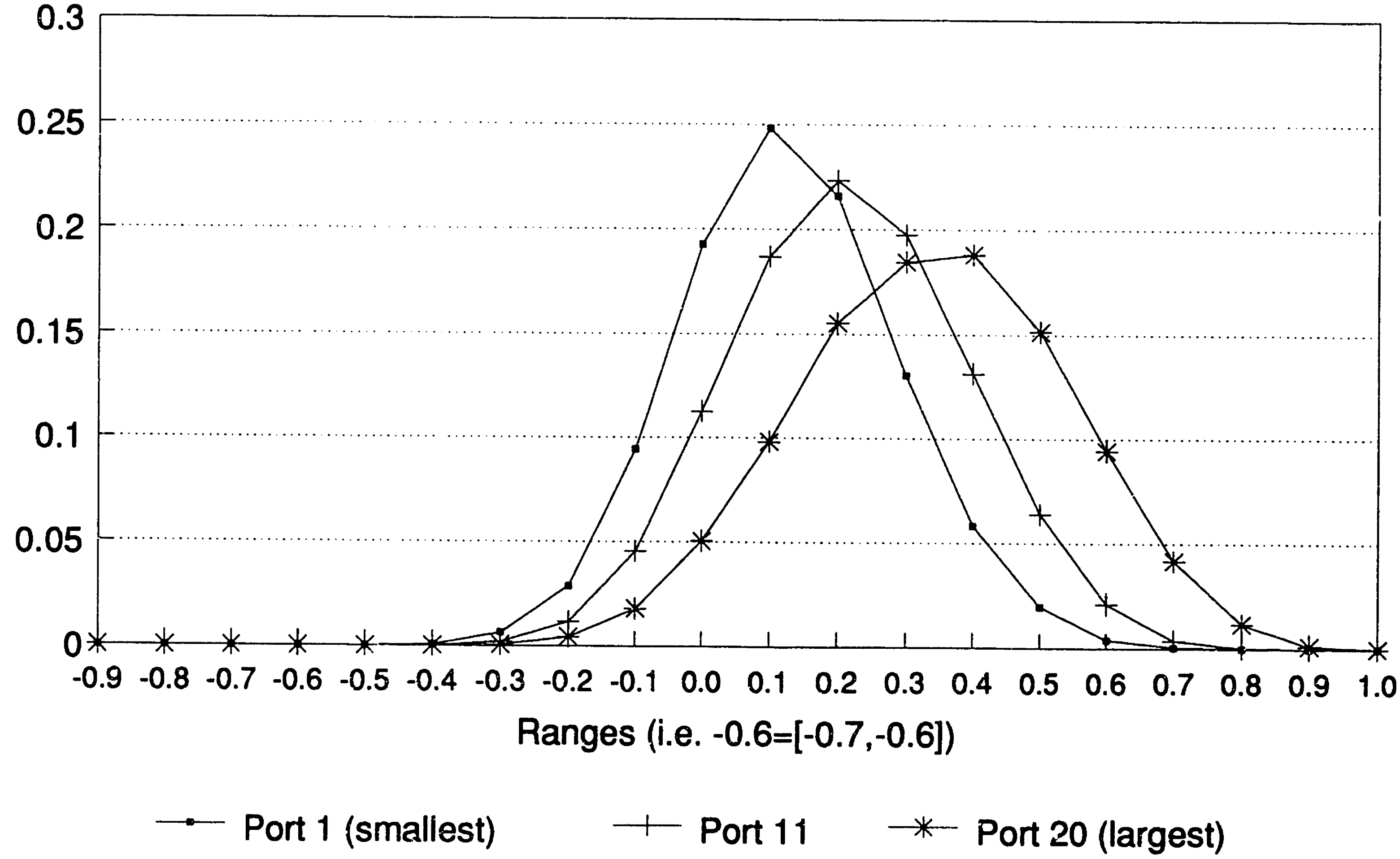
Range	(Daily) Port 1 (smallest)	Port11	Port 20 (largest)	(Weekly) Port1 (smallest)	Port 11	Port20 (largest)
-1.0 to -0.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-0.9 to -0.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-0.8 to -0.7	0.0004	0.0001	0.0001	0.0000	0.0000	0.0000
-0.7 to -0.6	0.0019	0.0010	0.0003	0.0000	0.0000	0.0000
-0.6 to -0.5	0.0083	0.0042	0.0016	0.0000	0.0000	0.0000
-0.5 to -0.4	0.0227	0.0135	0.0055	0.0006	0.0002	0.0001
-0.4 to -0.3	0.0489	0.0322	0.0142	0.0064	0.0020	0.0007
-0.3 to -0.2	0.0878	0.0633	0.0287	0.0288	0.0117	0.0043
-0.2 to -0.1	0.1291	0.1024	0.0520	0.0947	0.0454	0.0180
-0.1 to 0.0	0.1585	0.1382	0.0811	0.1927	0.1125	0.0504
0.0 to 0.1	0.1659	0.1605	0.1114	0.2483	0.1871	0.0980
0.1 to 0.2	0.1437	0.1574	0.1349	0.2159	0.2233	0.1552
0.2 to 0.3	0.1073	0.1309	0.1451	0.1302	0.1976	0.1846
0.3 to 0.4	0.0671	0.0951	0.1415	0.0582	0.1312	0.1882
0.4 to 0.5	0.0362	0.0579	0.1191	0.0197	0.0639	0.1515
0.5 to 0.6	0.0157	0.0280	0.0874	0.0042	0.0211	0.0943
0.6 to 0.7	0.0052	0.0110	0.0509	0.0007	0.0038	0.0417
0.7 to 0.8	0.0012	0.0033	0.0208	0.0001	0.0003	0.0117
0.8 to 0.9	0.0003	0.0007	0.0050	0.0000	0.0000	0.0013
0.9 to 1.0	0.0000	0.0001	0.0005	0.0000	0.0000	0.0000

Figure 1.1
Dist. for Contemporaneous Correlations
(Size-Sorted, Daily Returns)



See table 1.11 for details.

Figure 1.2
Dist. for Contemporaneous Correlations
(Size-Sorted, Weekly Returns)



See table 1.11 for details.

Chapter II

Economic Implications of Asset Return Predictability: Stochastic Dominance Comparisons

1 Introduction

Recent empirical findings show that stock returns are predictable in short- and long-horizons (see e.g. Lo and MacKinlay (1988) and Debondt and Thaler (1985,1987)). Although the unpredictability of asset returns is neither a sufficient nor a necessary condition for an economic equilibrium (See e.g. Leroy (1973)), the primary concern for the empirical evidence of the stock market predictability has been whether it leads to any 'excessive' profit opportunities, or equivalently, to any market inefficiency (see e.g. Lehmann (1990) and Debondt and Thaler (1985,1987)). Unfortunately, there has been no consensus among financial economists about the implication of stock return predictability to the efficient market hypothesis. Identifying 'excessive' return opportunities requires the definition of 'normal' returns which relies on economic models. As Fama (1970) pointed out more than two decades ago, the identification of opportunities for 'excessive' returns, or market inefficiency, therefore, has its intrinsic weakness because the results can be interpreted in two different ways (also see e.g. Shiller (1981) and Marsh and Merton (1986)). When market frictions are considered, the identification results have even weaker implications. For example, it is shown that equilibrium asset prices in the presence of market frictions may deviate from the frictionless market's no-arbitrage prices (see Tuckman and Vila (1992)).

In this paper, I am not going to design economic models that would determine whether the predictability of stock returns leads to any 'excessive' profit opportunities. Instead I am questioning whether under some circumstances which are less restrictive than those in economic models, the predictability leads to investors' unanimous agreement on the ranks of several uncertain investment or trading opportunities. In other words, I am trying to interpret the current state stock market predictability in the economic context without employing any economic paradigm.

When there is investors' unanimous agreement on the ranks of random variables, it is said that there are stochastic dominance relationships among the random variables. More specifically, when a random variable X is preferred to a random variable Y

by any investor whose preference is continuous and increasing, it is said that X first-degree stochastically dominates Y. On the other hand, it is said that X second-degree stochastically dominates Y when a random variable X is preferred to a random variable Y by any risk-averse investor whose preference is continuous and increasing. There is also a third-degree stochastic dominance relationship, but it will not be considered in this paper.[1] The essential idea of the stochastic dominance is well described by Russel and Seo (1989),

> Stochastic dominance rules dictate procedures for discovering unanimous orderings of uncertain prospects appropriate for utility functions within specified sets. ... The concept of stochastic dominance has introduced a convenient structure for analyzing optimal decisions when information on preferences is limited in various ways.

Although it has been more than three decades since stochastic dominance rules were developed as a decision theory, methods of testing for stochastic dominance were developed only recently. Major advances in testing for stochastic dominance were made by McFadden (1989) and Klecan, McFadden and McFadden (1991). Under the null hypothesis that one of two random prospects first-degree (or second-degree) stochastically dominates the other, McFadden(1989) developed a method which applies when observations for random variables are independent serially and cross-sectionally. Klecan, McFadden and McFadden(1991) extended McFadden's results by relaxing the independence assumptions. The extension by Klecan, McFadden and McFadden is important in its application to stock market return data since recent empirical findings show that stock returns are not serially independent and that returns covary cross-sectionally.

Three kinds of predictable returns are considered. The first is concerned with short-horizon return predictability (See e.g. Lo and MacKinlay (1988)). A representative 'buy the previous losers and sell the previous winners' strategy, known as the contrarian trading strategy, will be compared with a buy-and-hold investment strategy which is the most efficient strategy to use in the absence of return predictability.

[1] X is said to third-degree stochastically dominate Y when a random variable X is preferred to a random variable Y by any risk-averse investor who has a continuous and increasing preference, and his preference shows decreasing absolute risk-aversion.

From recent findings about weekly stock returns – negative autocorrelation for individual returns and positive autocorrelation for portfolio returns – it is conjectured that when there is no transaction cost, contrarian trading strategies out perform the buy-and-hold strategy. However, a certain level of transaction cost will make contrarian trading strategies not so attractive, and eventually a large transaction cost will make contrarian trading strategies inferior to the buy-and-hold strategy.

Empirical evidence show that the conjectures are right. Interestingly enough, at 5 percent significance level, one-way per-dollar transaction cost of 0.6 percent is enough to prevent the contrarian trading strategies from stochastically dominating the buy-and-hold strategy for all portfolios. At the same significance level, the buy-and-hold strategy stochastically dominates contrarian trading strategies for all portfolios at one-way per-dollar transaction costs of greater than 1 percent.

The second is concerned with long-horizon return reversals and a different kind of contrarian trading strategy, known as Debondt and Thaler's 'extreme performance portfolio strategy', is considered (1985, 1987). In terms of monthly mean returns, we confirm the following results by shown by Debondt and Thaler (1985, 1987): 1) extreme losers for the previous 5 years out perform the market (CRSP value-weighted market index), 2) extreme winners under-perform the market, and 3) the difference is much larger between the losers and the market than between the winners and the market. Stochastic dominance test, however, allows us to interpret their performance differently. The market second-degree stochastically dominates the winners, but the losers do *not* stochastically dominate the market. One possible explanation about this result would be that extreme losers sometimes have very negative returns — they sometimes go bankrupt — and therefore, some risk-averse investors prefer the market to the losers. The impact of January returns to long-horizon return reversals is found significant as shown by Zarowin (1990) in the sense that some of the stochastic dominance relationships change when the January returns are removed from the data set.

The third is concerned with a completely different type of return predictability, known as 'the weekend effect'. French (1980) and Gibbons and Hess (1983) have documented the fact that Monday stock returns are significantly lower than other weekdays stock returns with regard to their sample means. Apart from any profitable trading strategy that can exploit the weekend effect, which is not likely to exist because of relatively high transaction costs, there should be still some upward pressure for Monday stock prices. Suppose that you want to buy some stocks and you know that Monday stock prices usually drop. Then you will wait until Monday and will buy stocks on Monday. On the other hand, if you want to sell stocks, you don't want to sell them on Monday. It is asked if investors actually prefer other weekdays returns to Monday returns.

Empirical evidence shows that, in most cases, Monday returns are stochastically dominated by other weekdays returns and that the relationships are strong — they are from empirically distinguishable return distributions. Non-Monday returns also have stochastic dominance relationships, but the relationships are generally weak for large stocks — they are from empirically indistinguishable return distributions — whereas the relationships are strong for small stocks.

This paper is organized as follows. In section 2, theory and test methods for the stochastic dominance will be briefly reviewed. In section 3, short-horizon return predictability is discussed. Discussion of long-horizon returns will follow in section 4 and review of predictably low Monday stock returns will follow in section 5. A simple comparisons will be illustrated between the stochastic dominance criteria and conventional performance evaluation measures in section 6. In section 7, a weighted stochastic dominance test will be suggested and will be compared with the test statistics by Klecan, McFadden and McFadden. Conclusion will follow in section 8. Appendix A will contain simple derivations of weights for the weighted test statistics. Theorems by Klecan, McFadden and McFadden(1991) and some mathematical definitions will be stated without proofs in Appendix B.

2 Theory for Stochastic Dominance and Methodological Development

2.1 Stochastic Dominance as a Decision Theory

Consider two random variables, A and B, whose cumulative distribution functions (CDFs) are $F_A(z)$ and $F_B(z)$ respectively. Without loss of generality, assume that all random prospects have an upper bound of 1 and a lower bound of 0 and that all von Neumann-Morgenstern utility functions are defined on $[0, 1]$.

Definition 1 *It is said that A First-Degree Weakly Stochastically Dominates B if all investors who have continuous and increasing Von Neumann-Morgenstern utility functions in wealth weakly prefer A to B, denoted as* $A \succeq_{FSD} B$. *In a similar way, it is said that A Second-Degree Weakly Stochastically Dominates B if all investors who have continuous, increasing and concave Von Neumann-Morgenstern utility functions in wealth weakly prefer A to B, denoted as* $A \succeq_{SSD} B$.

The theorem by Hadar and Russell (1969) represents the definition of stochastic dominance in terms of cumulative probability distribution functions. It is restated below without proof.

Theorem 1 *(Hadar and Russell)* $A \succeq_{FSD} B$ *if and only if* $F_A(z) \leq F_B(z)$ *for all z.* $A \succeq_{SSD} B$ *if and only if* $\int_0^z F_A(y)dy \leq \int_0^z F_B(y)dy$ *for all z.*

Proof: See Hadar and Russell (1969).

A version of the theorem of Hadar and Russell (1969) will be restated in Appendix B. Figure 1 and Figure 2 depict first- and second-degree stochastic dominance relationships.

Stochastic dominance rules have one shortcoming to be noted. That is, in most cases, the relationship between two random variables predicts nothing about the relationship between the two when they are combined with other random variables. For example, although A stochastically dominates B and C stochastically dominates D, $A + C$ may not stochastically dominates $B + D$. This is one reason why the stochastic dominance test results are not related to the efficient market hypothesis. No stochastic dominance comparison, for example, between mutual funds and the market can lead to conclusions about the efficient market hypothesis. In the extreme case, investors may prefer the market since they hold other assets along with the mutual funds, even though mutual funds stochastically dominate the market. There are two exceptional cases where we can generalize the stochastic dominance relationships. The first case is where A stochastically dominates B conditionally on every possible value of Z and where any linear combination of A and Z in turn stochastically dominates the linear combination of B and Z (Levy and Levy (1984)). This is a very unusual case. Furthermore, this case is not very interesting since estimating the conditional probability functions generally requires a huge data set. Implementation is virtually impossible when Z can have many values. The second case is where A stochastically dominates B, and Z is independent of A and B. In this case, $A+Z$ still stochastically dominates $B + Z$ (Levy and Kroll (1978)).[2] This is also a very unusual case. To see the difference between considering the marginal distributions and considering the joint distributions with the market, the CRSP value-weighted market index will be combined when it is appropriate.

2.2 Methodological Development

Since Hadar and Russel (1969) introduced stochastic dominance rules as a decision theory, only a few papers have attempted to implement the rules empirically. Among them are Joy and Porter (1974) (J-P), Broske and Levy (1989) (B-L) and Seyhun (1992). J-P tested whether mutual funds out performed the market on average and

[2]But A+Z may stochastically dominate B+Z even though A does not stochastically dominate B and Z is independent of A and B.

found that there were more mutual funds which were dominated by the market than funds which dominated the market. But they, however, did not explicitly state their test method. B-L estimated a default risk of corporate bonds using stochastic dominance rules. They tried to define the default risk which allows the returns from corporate bonds to not stochastically dominate the returns from government bonds. Although they emphasized the introduction of a new method more than the provision of empirical results, they suggested no statistical property of the method. Seyhun reviewed the January effect by implementing stochastic dominance rules, but his method, like that of J-P and B-L, had the same problem in that it was not a statistical test.

The first paper which introduced a method for a stochastic dominance test as a statistical tool was by McFadden (1989). It not only suggested test statistics but also offered asymptotic distributions' property of the test statistics. One serious drawback for applying the method to financial economics was the assumption of independence upon which the test statistics were based.[3] A subsequent paper by Klecan, McFadden and McFadden (1991) (KMM) relaxes the assumption of independence so that the method can be applied to financial data. Since the method by KMM is relatively new, it is summarized as follows without heavy mathematical detail. For readers who are more interested in the method, theorems by KMM are quoted without proof in Appendix B.

The method by KMM is based on the following set of assumptions:

A1) A stochastic process is strictly stationary.

A2) A stochastic process is α-mixing with $\alpha(j) = O(j^{-\delta})$ for some $\delta > 1$. A precise definition of an α-mixing process is stated in Appendix B.

A3) The joint distribution of random variables satisfies generalized exchangeability. A precise definition of generalized exchangeability is stated in Appendix B.

The assumption of strict stationarity is required since a probability distributions

[3]The assumptions are that the observations are independent and the prospects are statistically independent.

should be estimated from a set of data. The second assumption is the condition for asymptotic independence. The mixing condition allows observations to be dependent over time and two random variables to be statistically dependent while the law of large numbers and the central limit theorem still work. The Monte-Carlo calculation of the significance level is based upon the third assumption. The generalized exchangeability condition allows the joint distribution of random variables to be unchanged even after random permutations among random variables in the Monte-Carlo calculation.

The first- and second-degree test statistics defined by KMM are

$$d_{2N}^{*} = \min\left(\max_{x}[F_{AN}(x) - F_{BN}(x)], \max_{x}[F_{BN}(x) - F_{AN}(x)]\right); \tag{1}$$

$$s_{2N}^{*} = \min\left(\max_{x}\int_{0}^{x}[F_{AN}(y) - F_{BN}(y)]dy, \max_{x}\int_{0}^{x}[F_{BN}(y) - F_{AN}(y)]dy\right), \tag{2}$$

where $F_{AN}(x)$ and $F_{BN}(x)$ denote empirical CDFs of random variables A and B. KMM shows that $d_{2N}^{*} \longrightarrow^{p} d^{*}$ and $s_{2N}^{*} \longrightarrow^{p} s^{*}$ where d^{*} and s^{*} were defined as

$$d^{*} = \min\left(\max_{x}[F_{A}(x) - F_{B}(x)], \max_{x}[F_{B}(x) - F_{A}(x)]\right), \tag{3}$$

$$s^{*} = \min\left(\max_{x}\int_{0}^{x}[F_{A}(y) - F_{B}(y)]dy, \max_{x}\int_{0}^{x}[F_{B}(y) - F_{A}(y)]dy\right), \tag{4}$$

where $F_{A}(x)$ and $F_{B}(x)$ denote '*true*' CDFs of A and B. Under the null hypothesis that one of two random variables first-degree (second-degree) stochastically dominates the other, d^{*} (s^{*}) should not be greater than 0. The null hypothesis is, therefore, rejected if d_{2N}^{*} (s_{2N}^{*}) is significantly larger than 0. Like other statistical tests, it is determined by the asymptotic distributions of the test statistics whether deviations from the null are significant or not. When appropriately scaled (multiplied by $\sqrt{N}$) and under the null hypothesis that one random variable (first- or second-degree) stochastically dominates the other, KMM shows that the test statistics converge in distribution to a maximum of a Gaussian process with a covariance function of ρ, where the maximum is taken over a set of domain which makes the two probability distributions equivalent. Under the alternative that neither of two random variables dominates the other, KMM shows that the scaled test statistics explode without bounds.

There are infinitely many cases within the null hypothesis. The case of identical CDFs is specifically used for deriving the asymptotic distribution of test statistics and is identified by having the largest size. Since the size is defined as the probability of rejecting the null hypothesis when the null is true, the case of identical CDFs is called a 'least favorable case'.

Unfortunately, the analytic assessment of the asymptotic distributions is difficult because of unknown dependence structures of random variables. By randomly permuting between two random variables, KMM suggests a Monte-Carlo calculation to approximate a significance level. The significance level calculated from the Monte-Carlo calculation is shown to converge in probability to a '*true*' significance level under the null of identical CDFs as the sample size and the number of random permutations approach infinity.

Here, the Kolmogorov-Smirnov (K-S) test should be briefly mentioned. When there are two empirically indistinguishable distributions, the stochastic dominance test usually shows a stochastic dominance relationship. It is not, however, interesting to conclude, for example, that Monday returns are stochastically dominated by other weekdays returns although their return distributions are not empirically distinguishable. Therefore, in some cases, a test for the hypothesis of identical distributions is recommended in addition to the stochastic dominance test. The Kolmogorov-Smirnov (K-S) test is designed for the hypothesis of identical distributions. By considering whether the maximum difference between two distributions is significantly different from zero, the K-S detects deviations from the null hypothesis. Consider the following definitions for kd^* and ks^*.

$$kd^* = \max\left(\max_x[F_A(x) - F_B(x)], \max_x[F_B(x) - F_A(x)]\right) \tag{5}$$

$$ks^* = \max\left(\max_x \int_0^x [F_A(y) - F_B(y)]dy, \max_x \int_0^x [F_B(y) - F_A(y)]dy\right) \tag{6}$$

and their empirical analog of

$$kd^*_{2N} = \max\left(\max_x[F_{AN}(x) - F_{BN}(x)], \max_x[F_{BN}(x) - F_{AN}(x)]\right), \tag{7}$$

$$ks^*_{2N} = \max\left(\max_x \int_0^x [F_{AN}(y) - F_{BN}(y)]dy, \max_x \int_0^x [F_{BN}(y) - F_{AN}(y)]dy\right). \quad (8)$$

These statistics are almost the same with stochastic dominance test statistics. The K-S statistics are max max(.) whereas stochastic dominance test statistics are min max(.). All the mathematical results apply to K-S statistics except for Theorem 5 in Appendix B (theorem 8 in Klecan, McFadden and McFadden (1991)). kd^*_{2N} detects the deviation from the null using the differences of CDFs. Similarly, sd^*_{2N} detects the deviation from the null using the differences of integrated CDFs. If the null hypothesis is correct, those two values should not be much different from zero. If one of the two statistics is significantly different from zero, it is concluded that the distributions are different.

3 Predictable Short-Horizon Stock Returns

Recent empirical findings show that stock prices do not follow random walks. Weekly individual stock returns show, on average, a slightly negative first-order autocorrelations which are not statistically significant, and weekly portfolio returns show a large positive first-order autocorrelations which are statistically significant. The absolute values of the first-order autocorrelations increase as the size, or equivalently the market value, of the stocks or of the portfolio decreases (see e.g. Lo and MacKinlay (1988)). A significant lead-lag relationship among individual stocks is found to be the main reason for the positive autocorrelation of portfolio returns in spite of the presence of negative autocorrelations for the individual stock returns. These predictability of stock returns make contrarian trading strategies, defined as 'strategies that buy the previous losers and sell the previous winners', profitable (see e.g. Lehmann (1990)). Furthermore, a large portion of the profit by contrarian trading strategies is found to be attributed to significant lead-lag relationships among individual stock returns (see e.g. Lo and MacKinlay (1990)). Unfortunately, whether these profits from the contrarian trading strategies imply any market inefficiency is still an open question.

When stock returns are completely unpredictable, no dynamic trading strategy is fruitful. In this case, a simple buy-and-hold strategy, defined as 'a strategy which buy stocks and hold them without further rebalancing', will perform no worse than any dynamic trading strategy even in the absence of transaction costs. A stochastic dominance comparison between a representative contrarian trading strategy and a simple buy-and-hold strategy, therefore, will enable us to interpret the current state short-horizon stock market predictability in the economic context without employing any economic paradigm.[4]

3.1 Description of Strategies and Data

At this point we define the notations as follows.

- R_{it}: Return from stock i at time t
- R_{mt}: Return from the market portfolio at time t
- ω_{it}: Number of dollars invested in stock i at time t (weight for stock i)
- tc: Per-dollar one-way transaction cost
- R_{ct}: Return from contrarian trading strategies at time t
- R_{bt}: Return from a buy-and-hold strategy at time t
- PV_t: Portfolio dollar value at time t.

Consider the following buy-and-hold strategy. At the initial date of the investment period, an equally weighted portfolio is formed and the portfolio is held without

[4] Note that the comparison is *not* between the best (contrarian) trading strategy and a buy-and-hold strategy. Once realized returns are observed, it is almost always possible to find some effective trading strategies unless the realized returns show complete unpredictability. Therefore, when trading strategies are considered based upon realized returns, there is always a problem of a so-called 'back testing bias', meaning that the trading strategies may depend too much on specific characteristics of realized returns rather than on more general characteristics of them. To reduce the bias, a representative contrarian trading strategy is employed which is similar to those employed by previous authors.

further rebalancing until the end of investment period. As stock prices change, the composition of the portfolio is constantly changed after the initial date, and R_{bt} is defined as

$$R_{bt} = PV_t / PV_{t-1}. \tag{9}$$

Now consider the following portfolio rebalancing strategy. Every week (or every two weeks) the portfolio is rebalanced depending on the weekly (or bi-weekly) returns of k weeks ago (or k bi-weeks ago). The weight for each stock is determined by its relative performance to the market, defined as 'an equally weighted size-sorted portfolio to which the stock belongs'. Specifically, the weight is determined as

$$\omega_{it} = \frac{-(R_{it-k} - R_{mt-k})}{N} + \frac{-(R_{it-k} - R_{mt-k})}{\sum_{R_{it-k} < R_{mt-k}} (R_{mt-k} - R_{it-k})} \text{ if } R_{it-k} < R_{mt-k},$$

$$\omega_{it} = \frac{-(R_{it-k} - R_{mt-k})}{N}, \text{ otherwise.} \tag{10}$$

The above weights sum to one, and therefore, a contrarian return is defined as[5]

$$R_{ct} = \sum_{i=0}^{N} \omega_{it} \times R_{it}. \tag{11}$$

Tests are based on weekly and bi-weekly returns.[6] Compared with contrarian strategies based on weekly returns, those based on bi-weekly returns require only half frequencies of portfolio rebalancing. When transaction costs are involved, bi-weekly contrarian trading strategies may be more effective than weekly contrarian strategies. Weekly returns are formed from daily returns from Wednesday to Tuesday. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five

[5]The weights considered here are a combined form of weights used by Lo and MacKinlay (1990) and by Lehmann (1990). Since the contrarian strategy is to be compared with buy-and-hold strategy, it is convenient to be able to define a return. Actually the contrarian trading strategy considered here puts more weight on the previous losers than do the strategies considered by Lo and MacKinlay (1990) and by Lehmann (1990).

[6]The reason for choosing bi-weekly instead of monthly returns is that autocorrelations for bi-weekly returns are much higher than those for monthly returns. Therefore, contrarian strategies based on monthly returns do not seem to be effective as much as those based on bi-weekly returns.

portfolios based on their initial market value. Once they were sorted, no rebalancing was made between portfolios. This method of sorting stocks facilitates the calculation of contrarian returns since the composition does not change over the sample period.[7]

An important aspect of testing for stochastic dominance between contrarian and buy-and-hold strategies is the incorporation of transaction costs. It will be assumed that the per-dollar transaction cost is the same for all stocks within each portfolio. Consider a total investment of one dollar (after transaction costs) — ω_{it-1} dollar for *'i'th* stock — at the beginning of week $t-1$. The transaction costs for the *'i'th* stock at the end of week $t-1$ are

$$TC_{it} = tc \times |\omega_{it} \times Y_t - \omega_{it-1} \times (1 + R_{it-1})|, \tag{12}$$

where $Y_t = 1 + \sum_{i=1}^{N} \omega_{it} \times R_{it-1} - \sum_{i=1}^{N} TC_{it}$. Solving for Y_t is not impossible, but involves complicated numerical computation. Instead of computing the exact values for Y_t, we will approximate the transaction costs as

$$TC_{it} \doteq tc \times |\omega_{it} - \omega_{it-1}|. \tag{13}$$

The approximation will generally make the transaction costs downward biased, but the bias is supposed to be small when weekly or bi-weekly returns are considered. For weekly or bi-weekly returns, R_{it-1} is generally less than 0.01. Also, Y_t cannot be greater than $1 + \sum_{i=1}^{N} \omega_{it-1} R_{it-1}$, where $\sum_{i=1}^{N} \omega_{it-1} R_{it-1}$ is generally less than 0.02 for both weekly and bi-weekly returns. Therefore, the bias would be generally less than 1 or 2 percent of tc. In other words, when we consider transaction costs of 1 percent per dollar, the bias would be less than 0.02 percent per dollar transaction. In reality, transaction costs are usually composed of two components: the clearing house costs of 0.05 percent per dollar and one-half of the bid-ask spread (see Lehmann (1990)). The per-dollar transaction costs for each portfolio change as stock prices move over time. The bias from the approximation for the transaction costs is relatively small compared

[7]There arises, however, a problem of survivorship bias in selecting stocks. Although the bias is difficult to assess, it is not believed that it is large enough to make the contrarian trading strategies void.

with the measurement errors arising from estimating the per-dollar transaction costs from the data. For example, for stocks with prices less than one hundred dollars, one dollar increase in stock prices implies a decrease in transaction cost from the bid-ask spread by more than 1 percent of the transaction cost, or 0.01 cent for 1 cent per-dollar transaction cost. Thus, it is conjectured that any conclusion from the test results would not be much affected by the bias from the transaction cost approximation. The portfolio return after transaction costs is

$$\sum_{i=1}^{N} \omega_{it} R_{it} - \sum_{i=1}^{N} TC_{it}. \tag{14}$$

Initially, eight levels from 0.05% to 1.0% of transaction costs will be considered. Then the break-even level of transaction costs will be determined at the 5 percent significance level.

3.2 Empirical assessment without transaction costs

Weekly and bi-weekly contrarian returns are considered where the weights depend on one to five period lagged returns. Table 2.1a and 2.1b show the first, second, and third sample moments of weekly and bi-weekly contrarian returns based on different lagged returns as well as those moments of buy-and-hold returns. Note that the means and the standard deviations for both contrarian returns and buy-and-hold returns are monotone decreasing as the size of portfolio increases. Skewness fluctuates slightly across different portfolios and also across different lags. Buy-and-hold returns for all portfolios are skewed left, and contrarian returns for large portfolios tend to be left-skewed. It is clear that contrarian returns with lag 1 have much higher mean returns than those with greater lags. Nevertheless, the differences in terms of standard deviations and skewnesses are not significant. Recent empirical findings show that *'k'th* autocorrelations for weekly portfolio returns decrease as k increases. As Lo and MacKinlay (1990) showed, a large part of contrarian profits comes from the cross sectional lead-lag relationships which cause the portfolio returns to be positively autocorrelated. The decreasing autocorrelations cause the contrarian strategies with higher lags to have lower mean returns.

Table 2.2 summarizes results from the test for stochastic dominance. Contrarian returns with lag 1 first- and second-degree stochastically dominate buy-and-hold returns at the significance level of 10%. For contrarian returns with lags of greater than one, they generally do not first- or second-degree stochastically dominate buy-and-hold returns, and vice-versa. Since the interesting question is whether the stock market predictability leads to any stochastic dominance relationship in the presence of transaction costs, only the contrarian strategy with lag 1 will be considered when transaction costs are involved.

Table 2.3 shows the comparisons between contrarian returns and the value weighted CRSP market index. Precisely speaking, these comparisons are not fair in the sense that contrarian trading strategies are implemented to stocks which survived in the whole sample period, whereas the value weighted CRSP market index is calculated from the whole stocks in the market. However, this table reveals how well the market index is diversified. Contrary to the results from the comparisons between contrarian returns and B-H returns, contrarian returns for small sized portfolios tend not to stochastically dominate the market even with lag 1 in spite of their huge mean returns. This is because small stocks usually have too much idiosyncratic risk compared to the market. In this context, it is well understood that contrarian returns with lag 1 for large sized portfolios stochastically dominate the market.

3.3 Empirical assessment with transaction costs

To examine how sensitive the stochastic dominance relationships between contrarian and buy-and-hold returns are to levels of transaction costs,[8] the following different

[8]It is assumed that short-sellings are allowed in rebalancing portfolios. Although short-sellings are more costly in practice, the results in this section do not change much. On average, the portion of short-selling is just 2% of the total investment. The stochastic dominance relationships in this section are found to hardly change even if short-sellings are prohibited.

levels of one-way per-dollar transaction costs are considered:

0.05%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.7%, and 1.0%.

Table 2.4 shows the first three moments of weekly and bi-weekly contrarian returns with transaction costs for each portfolio. As the level of transaction costs increases, the sample means decrease fast, whereas the standard deviations and the skewnesses remain almost unchanged. Therefore, it is obvious that the contrarian strategies become less attractive as the level of transaction costs increases. Table 2.5 shows the test results for weekly and bi-weekly contrarian returns with lag 1 for each portfolio at each level of transaction costs. For all portfolios having transaction costs above the break-even level, the contrarian returns do not stochastically dominate buy-and-hold returns, and eventually, the buy-and-hold returns stochastically dominate the contrarian returns. Interestingly enough, the break-even levels of transaction costs are less than 1 percent for all portfolios and for both weekly an bi-weekly returns. Table 2.6 shows the comparisons between contrarian returns with lag 1 and the value weighted market index in the presence of transaction costs.

Table 2.7 shows the break-even levels of transaction costs for the second-degree test at a significance level of 5%. For example, for weekly returns of portfolio 4, contrarian returns stochastically dominate buy-and-hold returns if the one-way transaction cost is less than 0.09%, and conversely, buy-and-hold returns stochastically dominate the contrarian returns if the one-way transaction cost is greater than 0.244%. The break-even levels of transaction costs for bi-weekly returns are greater than those for weekly returns for all portfolios, and they usually decrease as the portfolio size increases.

What are actual approximate levels of transaction costs for each portfolio? As Lehmann (1990) suggests, it is true that large investors can avoid or minimize transaction fees, but it is doubtful that they can always trade at prices within the bid-ask spread in cases where specialists are not involved. The actual transaction costs, therefore, depend on how frequently traders can trade at prices within the bid-ask spreads. Thus, the actual transaction cost may well differ from trader to trader and in turn may well be difficult to estimate.

As a very crude measure, transaction costs are assumed to be composed of bid-ask spreads only, and the minimum bid-ask spread is assumed to be one-eighth of a dollar, which is true for all NYSE and AMEX stocks whose prices are greater than one dollar. Furthermore, it will be assumed that no trader can trade at prices within the bid-ask spread. Table 2.8 shows the average characteristics of each portfolio and the ratios of one-half of the minimum bid-ask spread, which is one-sixteenth of a dollar, to their average prices. Stocks with large market values tend to have higher prices. For weekly and bi-weekly returns, the contrarian returns for portfolio 1 seem to second-degree stochastically dominate buy-and-hold returns when the actual transaction costs are less than our estimates. The same is true for the bi-weekly contrarian returns for portfolio 3. If the actual transaction costs are less than or close to the ratios in table 2.8 and other problems, such as a non-synchronous trading, are not serious, the test results indicate that the contrarian trading strategies are superior to the simple buy-and-hold strategy.[9]

4 Long-Horizon Return Reversals

Debondt and Thaler (1985,1987) reported that previous extreme losers tend to out perform the market and that previous extreme winners tend to under-perform the market for long-horizon stock returns. They interpreted these empirical findings as the manifestation of investors' irrational behavior. Many authors (e.g. Chan (1986), Ball and Kothari (1989), Chopra et al. (1992)) have reexamined the arguments by Debondt and Thaler by attributing some return reversals to the systematically changing equilibrium expected return. On the other hand, some authors (e.g. Zarowin (1990)) attribute Debondt and Thaler's results to unusually high January returns — so called the 'January Effect'. This section reviews Debondt and Thaler's implications

[9]Lo and MacKinlay (1989, 1990) suggest that the impact of nonsynchronous trading could not be large enough to explain the high autocorrelations of portfolio returns.

from a stochastic dominance point of view.

4.1 Description of Portfolio Formations and Data

Let's define notations as follows.

- R_{Lt} : Returns for a loser portfolio at time t.
- R_{Wt} : Returns for a winner portfolio at time t.
- R_{mt} : Returns for CRSP value-weighted market index at time t.

Tests are based on monthly stock returns data from CRSP monthly return file for NYSE and AMEX stocks. CRSP monthly return file contains data from 1926.1 to 1991.12. Consider the following portfolio formation strategy as in Debondt and Thaler.

1. For each stock, calculate the cumulative return for the last five years and then choose the 50 worst stocks and 50 best stocks.
2. Form an equal weighted portfolio labeled as a loser portfolio composed of the 50 worst stocks and another equal weighted portfolio labeled as a winner portfolio composed of the 50 best stocks.
3. Repeat the same procedure every year.

Two sample periods are considered. The first is the entire period from 1926.1 to 1991.12. Officer (1971) reported that the 1930s was not only a period of economic depression, but also of unusual uncertainty implied by highly volatile stock prices and industrial productions. It is therefore hard to believe that stock returns were stationary in the 1930s and thereafter. Since the stationarity is one of assumptions for the stochastic dominance test, a sample period from 1941.1 to 1991.12 will be considered. For each sample period, four pairs of returns are considered: (R_{Lt}, R_{Wt}), (R_{Lt}, R_{mt}), (R_{Wt}, R_{mt}), and $(R_{Lt} + R_{mt}, R_{Wt} + R_{mt})$. It is hoped that some of the changing equilibrium (conditional) expected return is captured by combining the market portfolio with the loser and the winner portfolios. CRSP value- weighted market index is used as the market portfolio. To determine the impact of January returns on return reversals, portfolio returns without January returns will be considered.

4.2 Empirical Assessment

Table 2.9 shows the first three moments of each portfolio returns for the two sample periods including and excluding January returns. First, let's look at returns including January returns. The loser portfolio far out performs the market portfolio, and the winner portfolio slightly under-performs the market portfolio in terms of mean returns. This is actually what Debondt and Thaler (1985,1987) documented. For the sub-period, all three moments of the loser portfolio returns are much less than those for the whole period. This implies that there was considerable difference between return distributions in the 1930s and thereafter. When January returns are excluded, the mean returns for the loser portfolio decrease remarkably. Returns for the winner portfolio and the market portfolio also decrease when January returns are excluded but not as much as do returns for the loser portfolio. The market portfolio is highly correlated with both the loser and the winner portfolios, implying that it is worth considering the combined returns.

The results from the stochastic dominance test are reported in Table 2.10 only for the sub-period since it is not believed that return distributions in the 1930s are the same as those in other periods. For returns including January, in terms of first-(and second-)degree stochastic dominance, only the $R_{Lt} + R_{mt}$ dominates $R_{Wt} + R_{mt}$ at 10% significance level. In terms of second-degree stochastic dominance, R_{mt} dominates R_{Wt}, and R_{Lt} dominates R_{Wt} at 10% significance level[10]. R_{Lt}, however, does not dominate R_{mt}. This result is somehow counter-intuitive since the loser portfolio

[10]Although the actual significance levels are not shown in this paper, the results show that R_{mt} dominated R_{Wt} at any reasonable significance level and that R_{Lt} barely dominated R_{Wt} at 10 percent significance level.

out performs the market portfolio to a greater extent than the market portfolio out performs the winner portfolio in terms of mean returns. One possible explanation for this result is that the loser portfolio sometimes realizes extremely low returns. Consider the probability of bankruptcy for firms which belong to the loser portfolio. They have unusually high probability of bankruptcy due to their extremely poor performance, and some actually go bankrupt within 5 years after the portfolio formation. Therefore, some risk averse investors would may well prefer the market portfolio to the loser portfolio in spite of the loser portfolio's high expected return. This result casts some doubt on Debondt and Thaler's interpretation of the long-horizon return reversals.

When the January returns are excluded, the results are changed. For one, no first-degree stochastic dominance relationship appeared at any reasonable significance level. Still, R_{mt} second-degree stochastically dominates R_{Wt}, and surprisingly, R_{mt} second-degree stochastically dominates R_{Lt}. Therefore, January returns have significant impact on the long horizon return reversals although evidence of return reversals does not completely vanish even after January returns are excluded.

5 The Weekend Effect

French (1980) and Gibbons and Hess (1981) documented that the typical Friday to Monday returns were not larger than the typical other weekday returns, empirically supporting that the market operates on trading time and not on calendar time. In fact, the mean of weekend returns was found to be negative, far less than the mean of other weekday returns. Apart from any profitable trading strategy that can exploit the predictably low Monday returns, it is not likely that any profitable trading strategy exists because of relatively high transaction costs. Yet, there still should be some kind of upward pressure for Monday's stock prices for the following reason. Suppose that you want to buy some stocks and you know that stock prices usually drop on

Monday. Then you will wait until Monday and will buy stocks on Monday. On the other hand, if you want to sell stocks, you don't want to sell them on Monday. These findings, therefore, have been considered as a stock market anomaly. In this section, it is asked if the predictable Monday stock returns lead to any stochastic dominance relationship between Monday stock returns and other weekday stock returns.

5.1 Data description

Data are selected from a CRSP daily return file for NYSE and AMEX stocks from 1962.8.3 to 1991.12.31. Three index returns and three individual company returns are considered. They are value weighted market index, equal weighted market index, S&P500, AT&T, Eastman Kodak, and IBM. The three individual companies are selected randomly among the Dow Jones Industrial companies. The reason for restricting the individual stocks to DJIS is to minimize the non-trading problem.[11] Transaction costs will not be considered since they are too large to make any dynamic trading activity which exploits return differences for each day of the week profitable.

5.2 Empirical Results

Table 2.11 shows the first three moments of returns for each index or each stock. For all market indices, and for AT&T and EK, the Monday returns have negative means with slightly higher standard deviations than do returns for other days. Only IBM has a slightly positive Monday mean return. Monday returns for all indices and stocks are skewed left. Mean returns for each non-Monday weekday are also very different from one another.

Table 2.12 shows results from a stochastic dominance test. The weekend effect is confirmed; For most cases, Monday returns are first- and second-degree dominated by returns for other weekdays at the 10% significance level. A few exceptions are related

[11]The way indices and individual stocks are selected is the same as Gibbons and Hess (1981) except that they considered all thirty Dow Jones Industrial stocks whereas only three are considered here.

to IBM's returns (remember that the lowest mean return for IBM was a Thursday mean return).

Now consider the test results among non-Monday returns. In some cases, the first-degree null is still accepted at the 10% significance level. The second-degree null is accepted at the same significance level in many cases. With regard to returns for market indices, only a few pairs have indistinguishable probability distribution at the same significance level (none for the EW index). As for individual returns, more than half of the pairs of Monday returns and almost all pairs of non-Monday returns have indistinguishable probability distributions. This implies that all weekday returns for small stocks tend to have different distributions, whereas those for large stocks tend to have the same distribution. This is somehow consistent with recent empirical findings that small stocks tend to be more predictable than large stocks. A thorough economic analysis of how the size of firms affect the behavior of stock returns is yet to be done.

6 Performance Evaluations vs. Stochastic Dominance

One common interest for both practitioners and academic researchers has been the issue of performance evaluation. For practitioners it offers a way of identifying better investment opportunities, and for academic researchers it is one of the methods for testing for the market efficiency hypothesis. Since the performance evaluation measures are based on some economic models, the problem of how to evaluate performance of mutual funds can be regarded as the problem of how to adjust the 'risks' associated with them.

A number of performance measures have been developed. Among them are Sharp measure, Jensen measure, Appraisal ratio, and Treynor measure. None is universally superior to another. The relevance of each measure depends on the environment in which the performance is evaluated. Many studies criticize these measures, and suggest modifications. For example, it has been shown by Admati and Ross (1985) and by Dybvig and Ross (1985) that the market timing ability can make upwardly bias the estimate of a systematic risk, resulting in a negative Jensen measure. Cornell (1979) and Grinblatt and Titman (1989) suggest alternative measures which are robust to market timing ability. Modest and Lehmann (1987) investigated the sensitivity of the Jensen measure and the Appraisal ratio to several benchmarks chosen to define the normal performance. They employed the standard CAPM and a variety of APT benchmarks, and found that these measures changed much, depending on the benchmark chosen. More extensive discussion of various performance measures is beyond the scope of this paper. In this section, a simple comparison between the conventional performance measures and the results from the stochastic dominance test is illustrated. Most conventional performance evaluation measures (except for the Sharp measure) are based upon some kind of portfolio theories, whereas the stochastic dominance test dictates the relationship between two random prospects.[12]

For the stochastic dominance criteria, two scenarios are considered here. One is the case where each mutual fund is the total investment, and the other is the case where each mutual fund and the value-weighted market portfolio consist of the total investment. We hope that a part of each fund's market risks is captured by including the market portfolio.[13]

[12]In comparison with the Sharp measure, the stochastic dominance criteria is more robust in the sense that the stochastic dominance criteria considers whole probability distributions, whereas the Sharp measure considers the first two moments of probability distribution.

[13]Including the risk-free asset within the total investment does not change the test results so long as the risk-free asset's returns are independent of returns for other assets and one stochastically dominates the other.

6.1 Performance Evaluation Measures and Data Description

Five close-end mutual funds are selected from CRSP daily return file for NYSE and AMEX stocks.[14] Each has observations of 5193 trading days from June 15, 1971 to December 31, 1991. To mitigate the impact of measurement error associated with daily returns, weekly returns from Wednesday to Tuesday are formed and all comparisons will be based on these weekly returns. For a brief description of each fund, please see Table 2.13. ASA is a high-mean, high-variance fund, JHI is a low-mean, low variance fund, and the remainder are in between.[15]

Now consider the following one-factor market model equation.

$$R_{pt} = R_{ft} + \beta_p \times (R_{mt} - R_{ft}) + \varepsilon_{pt}. \qquad (15)$$

Define $\bar{R}_p$ as a sample mean of R_{pt}, $\bar{R}_f$ as a sample mean of R_{ft}, $\sigma(R_p)$ as a standard deviation of R_{pt} and $\sigma(\varepsilon_p)$ as a standard deviation of ε_{pt}. Conventional measures for performance evaluations are defined as follows.[16]

Sharp Measure $(\bar{R}_p - \bar{R}_f)/\sigma(\varepsilon_p)$ Sharp measure divides average portfolio excess return by the standard deviation of the returns. This is relevant when the investor's utility function is quadratic and the portfolio is the total investment.

Appraisal Ratio $\alpha_p/\sigma(\varepsilon_p)$ Appraisal ratio divides the Jensen measure by the idiosyncratic risk of the portfolio. It measures the abnormal return per unit of risk, which can be diversifiable by holding a market portfolio. The proper environment for this measure is one in which the portfolio and the market portfolio consist of the total investment.

[14] All five funds were considered by Klecan, McFadden and McFadden (1991). Considering the purpose of this section, using the same data set as previous authors would not be an issue.

[15] All descriptions about the funds are quoted from Klecan, McFadden and McFadden (1991).

[16] The following descriptions are quoted from Chapter 24 of 'Investments' by Bodie, Kane, and Marcus (1989).

Jensen Measure $\bar{R_p} - [\bar{R_f} + \beta_p \times (\bar{R_m} - \bar{R_f})]$ Jensen measure, usually denoted as α_p, gauges the excess return under the assumption that the market model defines the normal return. (But the market model need not be one factor model.) When other assets or portfolios consist of the total investment along with the portfolio, the Jensen measure gives some indication of the potential contribution of the portfolio to the total investment.

Treynor Measure $(\bar{R_p} - \bar{R_f})/\beta_p$ Treynor measure divides the excess return for the portfolio by its systematic risk. Whereas the Jensen measure indicates how large the abnormal return is, Treynor measure gives the value for an excess return per unit of systematic risk. So when a number of various funds consist of the total investment, the Treynor measure can be an efficient guide for assigning relative rankings to each fund.

The Sharp measure is comparable with the stochastic dominance test for the first scenario: each fund is the total investment. The other three measures are partly comparable with the stochastic dominance test for the second scenario: each fund and the market portfolio consist of the total investment.

6.2 Empirical Comparisons

A brief description about the five close-end mutual funds is given in Table 2.13, which shows also the values of the first three moments for weekly returns for each mutual fund and the CRSP value-weighted market index. ASA has the highest mean and standard deviation, whereas JHI has the lowest. Table 2.14 shows the results from the stochastic dominance test, and Table 2.13 illustrates the values for the conventional performance evaluation measures for the five mutual funds. Among five mutual funds and the CRSP value-weighted market index, none first-degree stochastically dominates another at 10% significance level. At the same significance level, JHI second-degree stochastically dominates NGS, and TY second-degree stochastically dominates NGS.

Now let's consider the case where returns for each mutual fund are combined with returns from the CRSP value-weighted market index. Combined returns for TY and the value-weighted market index first-degree stochastically dominates combined returns for NGS and the value-weighted market index. Combined returns for JHI and the value-weighted market index second-degree stochastically dominates combined returns for NGS and the value-weighted market index. We would de better to avoid choosing NGS or 'NGS+Market'.

Keeping in mind the above test results, let's consider the conventional performance evaluation measures in Table 2.13. NGS has the lowest Sharp measure, but it is hard to identify that NGS should be avoided. For example, GAM has the highest Sharp measure, but it does not stochastically dominate other funds, even NGS. As for other performance evaluation measures, it is difficult to compare them directly with the stochastic dominance test since portfolio theories are involved with them. But roughly speaking, the stochastic dominance test gives somehow consistent results with the conventional performance evaluation measures: NGS has the lowest values for all measures.

7 A Weighted Stochastic Dominance Test

One estimation and inference problem is detected for the stochastic dominance test – the statistics do not seem to be very robust to tail events. The stochastic dominance test by Klecan, McFadden and McFadden — hereafter 'the conventional stochastic dominance test' — was found not very robust to tail events. In theory, first-degree stochastic dominance should imply second degree-stochastic dominance. However, the conventional test gave the contradictory result that accepted the first-degree null hypothesis but rejected the second-degree null hypothesis.[17] This result was usually

[17]Decisions are suggested to be made from the second-degree test results. Usually, this problem is caused by extremely negative observations. In this case, the CDF difference may not be big enough

observed when the min-max of CDF differences (or integrated CDF differences) was attained in the left tail of empirical distributions. Figure 3 explains the contradictory results. A few extremely negative observations often caused the min-max of the CDF differences to be small and the min-max of the integrated CDF differences to be large.[18] Heuristically speaking, by giving more weights to tail events, the test can be made, in principle, more robust to tail events. But the question is, "what are the 'optimal' weights?"

7.1 Weights for the SD Test

Define[19] the first- and second-degree *weighted* stochastic dominance test statistics as

$$d^*_{w2N} = \min\left(\max_x w(x)[F_{AN}(x) - F_{BN}(x)], \max_x w(x)[F_{BN}(x) - F_{AN}(x)]\right), \quad (16)$$

$$s^*_{w2N} = \min\left(\max_x w(x)\int_0^x [F_{AN}(y) - F_{BN}(y)]dy, \max_x w(x)\int_0^x [F_{BN}(y) - F_{AN}(y)]dy\right). \quad (17)$$

Their probability limits are defined as

$$d^*_w = \min\left(\max_x w(x)[F_A(x) - F_B(x)], \max_x w(x)[F_B(x) - F_A(x)]\right), \quad (18)$$

$$s^*_w = \min\left(\max_x w(x)\int_0^x [F_A(y) - F_B(y)]dy, \max_x w(x)\int_0^x [F_B(y) - F_A(y)]dy\right). \quad (19)$$

First consider $w(x)[F_{AN}(x) - F_{BN}(x)]$. Straightforward calculation shows that under the assumption of independent observations and independent prospects, the variance of $w(x)[F_{AN}(x) - F_{BN}(x)]$ given x is $w(x)^2[F_A(x)(1 - F_A(x)) + F_B(x)(1 - F_B(x))]/N$. (See Appendix B for the derivation.) For the conventional stochastic dominance test, $w(x)$ was 1 for all x, which implies that the variance of $[F_{AN}(x) - F_{BN}(x)]$ given x

to reject the first degree null but the integrated CDF difference can be large enough to reject the second-degree null. Risk-averse investors are likely to care about extremely negative returns.

[18] When results were contradictory, our interpretations were against the null hypothesis of stochastic dominance. In other words, the contradictory results was interpreted so as to reject the first- and the second-degree null hypotheses of stochastic dominance. A similar interpretation was made for the Kolmogorov-Smirnov tests.

[19] I am considerably indebted to Daniel McFadden who provided me with guide-lines how to make the test more robust to tail events.

attains its maximum when $F_A(x)$ and $F_B(x)$ are 1/2 and diminishes as x approaches the tail of distributions. This causes the conventional test not to be very robust to tail events.

For the same amount of deviation from the null hypothesis, the deviation arising from the tail distribution is more accurately measured than the deviation arising from the middle of the distribution, but the conventional stochastic dominance test treats the deviations equally. One way to improve the robustness of the conventional stochastic dominance test statistics to tail events would be to set $w(x) = \sqrt{N/var([F_{AN}(x) - F_{BN}(x)])}$ so that the variance of the test statistics given x is the same for all x. When we allow for dependent observations and dependent prospects, the variance of the CDF difference given x depends on the dependence structures of the stochastic processes that are difficult to identify. In practice, dependence structures can be assumed — i.e. AR1 process with one common factor — and the parameters can be estimated. Considering, however, that the purpose of estimating the variance is to use it as a weight, finding a consistent estimator for the variance is not necessary. Furthermore, it is not even clear which $w(x)$ gives the test a maximum power given the same size since analytical assessment of the asymptotic distribution of the test statistics is not possible.[20] There is a trade-off. If $w(x)$ is taken as the square root of the inverse of the variance, the test is more robust to tail events but less robust to middle events. Therefore,

$$\hat{w(x)} = \sqrt{\frac{N}{[F_{AN}(x)(1 - F_{AN}(x)) + F_{BN}(x)(1 - F_{BN}(x))]}} \tag{20}$$

is suggested as an estimator for $w(x)$ for first-degree stochastic dominance test statistics.

Weights for the second-degree stochastic dominance test are a bit more complicated. Under the assumption of independent observations and independent prospects,

[20] One way to compare the conventional stochastic dominance test with the weighted stochastic dominance test would be Monte-Carlo experiments.

a straightforward calculation shows that the variance of $\int_0^x [F_{AN}(y) - F_{BN}(y)]dy$ is

$$(1/N)\int_0^x \left[(x-y)^2 f_A(y) + (x-y)^2 f_B(y)\right] dy$$
$$-\left(\int_0^x (x-y) f_A(y) dy\right)^2 - \left(\int_0^x (x-y) f_B(y) dy\right)^2,$$

where $f_A(y)$ and $f_B(y)$ are probability density functions (PDF) for random variables of A and B. (See Appendix B for the derivation.) The variance of the estimated integrated CDF difference can be estimated from the empirical distributions of $F_{AN}(.)$ and $F_{BN}(.)$ by using the above expression's discrete analog, which is

$$(1/N)\sum_{i=1}^{N}\left[(x-A_i)^2(\frac{1}{N})1(A_i < x) + (x-B_i)^2(\frac{1}{N})1(B_i < x)\right]$$
$$-\left(\sum_{i=1}^{N}(x-A_i)(\frac{1}{N})1(A_i < x)\right)^2 - \left(\sum_{i=1}^{N}(x-B_i)(\frac{1}{N})1(B_i < x)\right)^2, \qquad (21)$$

where 1(.) denotes an indicator function whose value is 1 if the condition in the parenthesis is satisfied, and 0 otherwise. As in first-degree stochastic dominance test statistics, a weight for second-degree test statistics is suggested as

$$\hat{w(x)} = \sqrt{\frac{1}{Var(\text{estimate of integrated CDF difference})}}. \qquad (22)$$

For distributions similar to normal distributions, the weights for the first-degree test are larger for (left and right) tail events and smaller for middle events. The weights for the second-degree test are the largest for left tail events and diminishes as the events approach the right tail. Figure 4 describes the weights for the first- and the second-degree stochastic dominance tests when empirical distributions are normally distributed. To compare the weighted stochastic dominance tests with the conventional stochastic dominance test, the following subsection will show the comparative results of Monte-Carlo experiments.

7.2 Monte Carlo Experiments for the Weighted SD Test

Table 2.15 compares the finite sample property of the conventional and the weighted stochastic dominance tests. The specific processes for random variables are described in Table 2.15. All cases are classified into three categories: 1) first- and second-degree stochastic dominance, 2) second- not first-degree stochastic dominance, and 3) neither second- nor first-degree stochastic dominance. Summarizing the results, the first-degree weighted stochastic dominance test performs better than the first-degree conventional stochastic dominance test for all three cases. As for the second-degree test, the weighted test sometimes performs better but sometimes performs worse than the conventional tests depending on the cases. When two random variables are highly correlated, both test perform considerably well, but on the other hand, when each variable is highly autocorrelated, both test perform poorly in the small sample size. Now let's look at Monte-Carlo results for each case.

When the true relationship is the first- (and second-) degree stochastic dominance, both tests perform well for the first-degree null hypotheses. For the second-degree null hypothesis, the conventional test performs better than the weighted test since the conventional test has lower rejection rates.

The weighted stochastic dominance test performs much better when the true relationship is the second (not first-) degree stochastic dominance. For example, consider the case where

$$\alpha_1 = 0,\ \alpha_2 = -0.2,\ \beta_1 = 1,\ \beta_2 = 4,\ \lambda = 0.1,\ \rho = 0.5,\ N = 100,\ I = 100.$$

Both tests are good at detecting the deviation from the first- degree null hypothesis: both tests have high power against the first-degree null. As for the second-degree null hypothesis, when the decisions are made at 5% significance level, the conventional test falsely rejects the second-degree null hypothesis 29 times out of 100 experiments whereas the weighted test rejects only 12 times out of 100 experiments. When two random variables are highly correlated, the Monte-Carlo results indicate that a large sample size is required for both tests to have high power against first-degree null hypothesis. Both tests perform poorly in rejecting the false first-degree null hypothesis when random variables are highly autocorrelated ($\rho = 0.9$) and the sample size is small (N=100). When the sample size is increased up to N=1,000, the weighted test is much better in rejecting false first- degree null hypothesis than the conventional test.

When one of two random variables neither second- nor first-degree stochastically dominate the other, the weighted test performs better than the conventional test against the false first-degree null hypotheses. The weighted test, however, performs slightly worse than the conventional test in rejecting false second-degree null hypotheses except when the random variables are highly autocorrelated.

For all cases, the weighted test performs better with regard to the first-degree null hypotheses. On the other hand, the conventional test sometimes performs better with regard to the second-degree null hypotheses. The weighted test, however, performs better for all cases if the random variables are highly autocorrelated. Although the test results for a sample size of 1,000 are not reported in this paper (except for the case where ρ is 0.9), it is shown that both test perform considerably well for all the cases considered in the Monte-Carlo experiments when a large number of observations are available.

The Monte-Carlo experiments in this paper considered only certain kinds of Gaussian random variables. Much more extensive Monte-Carlo experiments are yet to be done. More importantly, a better way of modifying the stochastic dominance test should be searched such that the modified test has improved properties with regard to second-degree null hypotheses as well as first-degree null hypotheses.

8 Summary and Conclusions

In the absence of transaction costs, contrarian trading strategies which exploit the weekly and bi-weekly stock returns' first-order autocorrelations stochastically dominate the passive buy-and-hold investment strategy in the first- and second-degree at any reasonable significance level. Contrarian trading strategies exploiting the second- or higher-order autocorrelations do not first- or second-degree stochastically dominate the buy-and-hold strategy. Yet, when transaction costs are involved, the stochastic dominance relationships change depending on the level of transaction cost. For all portfolios, the second-degree break-even levels of one-way transaction costs are between 0.05% and 1.0%. Roughly, the break-even levels are smaller for bi-weekly returns and decrease as the portfolio size increases. When extreme loser stocks and extreme winner stocks are selected for a five year period, extreme losers seem to second-degree stochastically dominate extreme winners for the following five year period on a monthly return basis. The extreme losers do not stochastically dominate the CRSP value-weighted market index, but the extreme winners are second-degree stochastically dominated by the CRSP value-weighted market index. When January returns are removed from the data set for the following five years after portfolio formation, the results change. The extreme losers do not stochastically dominate the winners, and interestingly enough, the CRSP value-weighted market index second-degree stochastically dominates the losers. Much of the long-horizon return reversals seem to be due to the January effect.

For three market indices and three Dow Jones Industrial stocks, Monday's returns were generally first- and second-degree dominated by other days' returns, confirming the weekend effect. Only a few pairs of returns showed that they do not stochastically dominate each other. As for market indices returns, these relationships were strong in the sense that the comparisons were made for two distinguishable return distributions. For three large individual returns, these relationships were weak; Most non-Monday returns have empirically indistinguishable probability distributions. This implies that small firms have different return distributions for each day of the week. Finally, the results from the stochastic dominance comparisons among several mutual funds were illustrated along with conventional performance evaluation measures for the mutual funds.

It is an open question as to how the stochastic dominance test can be improved so that the test shows results which are more consistent with the theoretical implications of stochastic dominance. Theory tells us that first-degree stochastic dominance implies second-degree stochastic dominance. Sometimes, this was violated in the current test statistics because of extreme tail events. Modification of the test statistics was suggested by weighing the test statistics properly. Monte-Carlo experiments show that the weighted test has better finite sample property with regard to the first-degree null hypotheses. With regard to the second-degree null hypotheses, however, the weighted test sometime perform worse than the conventional test. There may be better ways to improve the test so that it has better finite sample property with regard to second-degree null hypotheses as well as first- degree null hypotheses, but they are yet to be developed in future research.

References

[1] Ball, R. and S. Kothari (1989), "Nonstationary Expected Returns: for Tests of Market Efficiency and Serial Correlation in Returns," Journal of Financial Economics 25, 51-74.

[2] Billingseley, Patrick (1969), *Convergence of Probability Measures*, Wiley & Sons, New York.

[3] Bodie, Z., M. Kane and A. Marcus (1989), *Investments*, R.D. Irwin Inc., Boston, MA.

[4] Broske, M. and H. Levy (1989), "The Stochastic Dominance Estimation of Default Probability," in T. Fomby and K. Seo (eds) *Studies in the Economics of Uncertainty*, Springer, New York, 91-112.

[5] Chan, K.C. (1988), "On the Contrarian Investment Strategy," Journal of Business 61, 147-163.

[6] Chopra, N., J. Lakonishok and J. Ritter (1992), "Measuring Abnormal Performance: Do Stocks Overreact?," Journal of Financial Economics 31, 235-268.

[7] Debondt, W. and R. Thaler (1985), "Does the Stock Market Overreact?," Journal of Finance 40, 793-805.

[8] Debondt, W. and R. Thaler (1987), "Further Evidence on Investor Overreaction and Market Seasonality," Journal of Finance 42, 557-582.

[9] Fama, Eugene (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," Journal of Finance 25, 383-417.

[10] Fama, Eugene (1976), *Foundations of Finance*, Basic Books Inc.

[11] French, Kenneth (1980), "Stock Returns and the Weekend Effect," Journal of Financial Economics 8, 1-2.

[12] Gibbons, M. and P. Hess (1981), "Day of the Week Effects and Asset Returns," Journal of Business 54, 579-596.

[13] Grinblatt, M. and Sheridan Titman (1989), "Portfolio Performance Evaluation: Old Issues and New Insights," Review of Financial Studies2, 393-421.

[14] Hadar, J. and W. Russell (1969), "Rules for Ordering Uncertain Prospects," American Economic Review 59, 25-34.

[15] Hadar, J. and W. Russel (1971), "Stochastic Dominance and Diversification," Journal of Economic Theory 3, 288-305.

[16] Hanoch, G. and H. Levy (1969), "The Efficiency Analysis of Choices Involving Risk," Review of Economic Studies 36, 335-346.

[17] Huang C. and R. Litzenberger (1988), *Foundations for Financial Economics*, North-Holland, New York.

[18] Joy, O. and R. Porter (1974), "Stochastic Dominance and Mutual Fund Performance," Journal of Financial and Quantitative Analysis 9, 25-31.

[19] Klecan, L., R. McFadden and D. McFadden (1991), "A Robust Test for Stochastic Dominance," Working Paper, MIT and Cornerstone Research.

[20] Lehmann, Bruce N. (1990), "Fads, Martingales, and Market Efficiency," Quarterly Journal of Economics 105, 1-28.

[21] Leroy, S.F. (1973), "Risk Aversion and the Martingale Property of Stock Returns," International Economic Review 14, 436-446.

[22] Levy, H. and Y. Kroll (1978), "Investment Decision Rules, Diversification, and the Investor's Initial Wealth," Econometrica 46, 1231-1237.

[23] Levy, H. and A. Levy (1984), "Ordering Uncertain Options Under Inflation," Journal of Finance 39, 1223-1229.

[24] Lo, A. and C. MacKinlay (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," Review of Financial Studies 1, 41-66.

[25] Lo, A. and C. MacKinlay (1990), "When Are Contrarian Profits Due To Stock Market Overreaction?," Review of Financial Studies 3, 175-205.

[26] Marsh, T. and R. Merton (1986), "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," American Economic Review 76, 483-498.

[27] McFadden, Daniel (1989), "Testing for Stochastic Dominance," in T. Fomby and K. Seo (eds) *Studies in the Economics of Uncertainty*, Springer, New York, 113-134.

[28] Merton, Robert, "On the Current State of the Stock Market Rationality Hypothesis," in Dornbusch, Fischer, and Bossons (eds) *Macroeconomics and Finance: Essays in Honor of Franco Modigliani*, MIT Press, Cambridge.

[29] Meyer, J. and R. Rache (1992), "Sufficient Conditions for Expected Utility to Imply Mean-Standard Deviation Rankings," Economic Journal 102, 91-106.

[30] Lehmann, B. and D. Modest(1987), "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons," Journal of Finance42, 233-265.

[31] Officer, Robert (1971), "A Time Series Examination of the Market Factor of the New York Stock Exchange," Ph.D. Dissertation, University of Chicago.

[32] Press, W., B. Flannery. S. Teukolsky, and W. Vetterling (1986), *Numerical Recipes*, Cambridge University Press, Cambridge.

[33] Pollard, D. (1984), *Convergence in Stochastic Processes*, Springer, New York.

[34] Porter, R.B. (1978), "Portfolio Applications: Empirical Studies" in G. Whitemore and M. Findlay (eds) Stochastic Dominance, Lexington Books, Lexington, 117-161.

[35] Rothchild, M. and J. Stiglitz (1970), "Increasing Rist I: A definition," Journal of Economic Theory 2, 225-243.

[36] Russell, W. and T. Seo (1989), "Representative Sets for Stochastic Dominance Rules," in T. Fomby and K. Seo (eds) *Studies in the Economics of Uncertainty*, Springer, New York, 59-76.

[37] Seyhun, H. Nejat (1992), "Can Omitted Risk Factors Explain The January Effect? A Stochastic Dominance Approach," Working Paper, The University of Michigan.

[38] Shiller, R. (1981), "Do Stock Prices Move Too Much To Be Justified By Subsequent Changes in Dividends?," American Economic Review 71, 421-436

[39] Tuckman, B. and Jean-Luc Vila (1992), "Arbitrage With Holding Costs: A Utility-Based Approach," Journal of Finance 47, 1283-1302.

[40] White, Halbert (1984), *Asymptotic Theory for Econometricians*, Academic Press Inc., Orlando, Fl.

[41] Zarowin, Paul (1990), "Size, Seasonality, and Stock Market Overreaction," Journal of Financial and Quantitative Analysis 25, 113-125.

Appendix A

Variance of CDF Difference and Integrated CDF Difference

In the following derivations, 1(.) denotes an indicator function whose value is 1 if the condition in the parenthesis is satisfied , and 0 otherwise. All other notations are the same with those in the main text.

1. Variance of the estimated CDF difference

Under the assumptions that observations are independent over time and two randcm variables are independent of each other, the variance of the estimated CDF difference is as follows.

$$\begin{aligned} Var\left[F_{AN}(x) - F_{BN}(x)\right] &= Var\left[(\frac{1}{N})\sum_{i=1}^{N} 1(A_i \leq x)\right] + Var\left[(\frac{1}{N})\sum_{i=1}^{N} 1(B_i \leq x)\right] \\ &= (\frac{1}{N})(F_A(x)(1 - F_A(x)) + F_B(x)(1 - F_B(x))). \end{aligned}$$

The first equality is from the fact that $F_{kN}(x)$ is defined as $(1/N)(\sum_{i=0} N1(k_i \leq x))$ and from the assumption that two random variables are independent of each other. The second equality comes from the assumption that observations are independent over time.

2. Variance of the estimated integrated CDF difference

Under the assumptions that observations are independent over time and two random variables are independent of each other, the variance of the estimated integrated CDF difference is as follows.

$$\begin{aligned} Var\left[\int_0^x (F_{AN}(y) - F_{BN}(y))\, dy\right] &= Var\left[(\frac{1}{N})\sum_{i=1}^{N}(x - A_i)1(A_i \leq x)\right] \\ &+ Var\left[(\frac{1}{N})\sum_{i=1}^{N}(x - B_i)1(B_i \leq x)\right] \end{aligned}$$

$$
\begin{aligned}
&= (\frac{1}{N})(\int_0^x (x-y)^2 f_A(y)dy \\
&- (\int_0^x (x-y) f_A(y)dy)^2) \\
&+ (\frac{1}{N})\int_0^x ((x-y)^2 f_B(y)dy \\
&- (\int_0^x (x-y) f_B(y)dy)^2).
\end{aligned}
$$

The first equality is from the fact that $\int_0^x F_{kN}(y)dy$ is defined as $(1/N)(\sum_{i=0}^N 1(x - k_i)1(k_i \leq x))$ and from the assumption that two random variables are independent of each other. The second equality comes from the assumption that observations are independent over time.

Appendix B

Theorems by Klecan, McFadden and McFadden and Some Mathematical Definitions

To help readers to understand the method used in this paper, key theorems from Klecan, McFadden and McFadden (1991) will be quoted without proofs along with several mathematical definitions. The theorem of Hadar and Russel (1969) about first- and second-degree stochastic dominance will also be quoted from Klecan, McFadden and McFadden (1991). Throughout this appendix, without loss of generality, random variables are bounded below by 0 and bounded above by 1.

Definition 1 *Let's denote B_n^{n+m} as the Borel field generated by $\{Z_t(\omega), t = n, \ldots, n+m\}$. In other words, B_n^{n+m} is the smallest collection of events that allows us to express the probability of an event, say $[Z_n < a_1, Z_{n+1} < a_2]$, in terms of the probability of an event in B_n^{n+m}, say $[\omega : Z_n(\omega) < a_1, Z_{n+1}(\omega) < a_2]$. Let A and B σ-fields and define*

$$\alpha(a,b) = \sup_{a \in A, b \in B} |P(a \cap b) - P(a)P(b)|.$$

Let's define $\alpha(m) = \sup_n \alpha(B_{-\infty}^n, B_{n+m}^{\infty})$. Then $\{Z_t\}$ is called α-mixing if $\alpha(m) \longrightarrow 0$ as $m \longrightarrow \infty$.

Definition 2 *The finite set $(X_1, \ldots, X_n)$ of random variables is said to be exchangeable if the joint distribution of $(X_{\pi 1}, \ldots, X_{\pi n})$ is the same as that of $(X_1, \ldots, X_n)$ for every permutation π.*

Theorem (LM) 1 *Assume that prospects are bounded above and below. Also assume that von Neumann-Morgenstern utility functions are continuous, increasing functions.*

Consider a set of prospects $A = \{X_1, \ldots, X_k\}$.

The prospects in A are first-degree stochastically maximal (or none of them first-degree stochastically dominates another); i.e.,

$$d^* = \min_{i \neq j} \max_x [F_i(x) - F_j(x)] > 0,$$

if and only if for each i and j, there exists a continuous increasing function u such that $E[u(X_i)] > E[u(X_j)]$.

The prospects in A are second-degree stochastically maximal; i.e.,

$$s^* = \min_{i \neq j} \max_x \int_0^X [F_i(y) - F_j(y)]dy > 0,$$

if and only if for each i and j, there exists a continuous increasing strictly concave function u such that $E[u(X_i)] > E[u(X_j)]$.

Proof: See Theorem 1 in Klecan, McFadden and McFadden (1991).

Now consider the following algorithm. For each pair X_i and X_j with $i < j$, form a vector z of length 2N containing the observations from X_i, followed by the observations from X_j. Form a vector l of length 2N containing the indices of the elements of z in ascending order; i.e., $z_{l_m} \leq z_{l_{m-1}}$. Then the following theorem follows.

Theorem (LM) 2 *Let's define empirical analog of* d^* *and* s^* *in Theorem 1.*

$$d^*_{2N} = \min_{i \neq j} \max_x [F_{iN}(x) - F_{jN}(x)];$$

$$s^*_{2N} = \min_{i \neq j} \max_x \int_0^x [F_{iN}(y) - F_{jN}(y)]dy.$$

Also, let's define the following set of equations. A superscript 'ij' can be added to the following equations to identify the pair of prospects i,j being evaluated.

$$d_0 = d_0^+ = d_0^- = s_0 = s_0^+ = s_0^- = 0,$$

and recursively for $m = 1, \ldots, 2N$,

$$\delta_m = \begin{cases} -1 & \text{if } l_m > N, \\ +1 & \text{if } l_m \le N \end{cases}$$

$$d_m = d_{m-1} + \delta_m,$$

$$d_m^+ = \max(d_{m-1}^+, d_m), \qquad d_m^- = \max(d_{m-1}^-, -d_m)$$

$$s_m = s_{m-1} + d_{m-1} \cdot (z_{l_m} - z_{l_{m-1}}),$$

$$s_m^+ = \max(s_{m-1}^+, s_m), \qquad s_m^- = \max(s_{m-1}^-, -s_m)$$

Then d_{2N}^* *and* s_{2N}^* *can be restated as following equations.*

$$d_{2N}^* = N^{-1} \min_{i<j} \min(d_{2N}^{+ij}, d_{2N}^{-ij}) = N^{-1} \min_{i \neq j}(d_{2N}^{+ij}),$$

$$s_{2N}^* = N^{-1} \min_{i<j} \min(s_{2N}^{+ij}, s_{2N}^{-ij}) = N^{-1} \min_{i \neq j}(s_{2N}^{+ij}).$$

Proof: See Theorem 4 in Klecan, McFadden and McFadden (1991).

The following theorem shows the convergence of d_{2N}^* and s_{2N}^*.

Theorem (LM) 3 *Assume that* $(X_{1n}, \ldots, X_{kn})$, *viewed as a stochastic process indexed by* $n = 1, 2, \ldots$, *with values in* $[0,1]^k$, *is strictly stationary and* α*-mixing with* $\alpha(j) = O(j^{-\delta})$ *for some* $\delta > 1$. *Then* $s_{2N}^* \longrightarrow^p s^*$ *and* $d_{2N}^* \longrightarrow^p d^*$.

Proof: See Theorem 5 in Klecan, McFadden and McFadden (1991).

The following theorem implies that $N^{1/2} s_{2N}^*$ has an asymptotic distribution when $s^* \le 0$, and that this distribution is nondegenerate in the 'least favorable' case of identical marginals.

Theorem (LM) 4 *Assume that* (X_{1n}, X_{2n}) *is a strictly stationary stochastic process, taking values in* $[0,1] \times [0,1]$, *such that the process is* α*-mixing with* $\alpha(j) = O(j^{-\delta})$ *for some* $\delta > 1$. *Define* $\psi_i(x) = \int_0^x F_i(y)dy$. *Consider a random variable*

$$g(w, X_{in}, X_{jn}) = \int_0^w [1(y > X_{in}) - 1(y > X_{jn})]dy = \max(w - X_{in}, 0) - \max(w - X_{jn}, 0).$$

It is trivial that $E[g(w, X_{1n}, X_{2n})] = \psi_1(w) - \psi_2(w)$. *Define*

$$\tilde{g}(w, X_{1n}, X_{2n}) = g(w, X_{1n}, X_{2n}) - E[g(w, X_{1n}, X_{2n})],$$

$$S_{2N}(w) = N^{1/2}\left[\int_0^w [F_{1N}(y) - F_{2N}(y)]\,dy\right] \equiv N^{1/2}s_{2N}(w) \equiv N^{-1/2}\sum_{n=1}^{N} g(w, X_{1n}, X_{2n}),$$

$$\tilde{S}_{2N}(w) = S_{2N}(w) - E[S_{2N}(w)] \equiv N^{-1/2}\sum_{n=1}^{N} \tilde{g}(w, X_{1n}, X_{2n}).$$

Then $E[\tilde{S}_{2N}(w)] = 0$; $\tilde{S}_{2N}(0) = 0$; *there exists* $M > 0$ *such that*

$$E[(\tilde{S}_{2N}(v) - \tilde{S}_{2N}(w))^2] \leq M(v - w)^2;$$

and there exists a covariance function $\rho(w, v)$ *that is uniformly Lipschitz on* $[0,1] \times [0,1]$ *such that* $E[\tilde{S}_{2N}(w)\tilde{S}_{2N}(v)] \to \rho(w, v)$ *uniformly.*
Assume $0 < \rho(1,1) = \lim_{N\to\infty} N_{-1}\sum_{n=1}^{N}\sum_{m=1}^{N} cov[(X_{2n} - X_{1n}), (X_{2m} - X_{1m})]$. *Then the sequence of processes* $\tilde{S}_{2N}(.)$ *for* $N \to \infty$ *is tight: i.e., it has the stochastic boundedness property that for each* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that*

$$\sup_N P(\sup_x |\tilde{S}_{2N}(x)| > \delta) < \varepsilon,$$

and the stochastic equicontinuity property that for each positive η *and* ε, *there exists* $\delta > 0$ *such that for all* N,

$$P\left(\sup_{|w-v|<\delta} \left|\tilde{S}_{2N}(w) - \tilde{S}_{2N}(v)\right| > \eta\right) < \varepsilon.$$

Also, the $\tilde{S}_{2N}$ *converge in distribution to a Gaussian process* $\tilde{S}_\infty$ *that has continuous sample paths with probability one, and the covariance function* ρ.
If $\psi_1(x) \leq \psi_2(x)$, *with equality holding for* $x \in A$, *then* $N^{1/2}s_{2N}^{+12} \equiv \max_x S_{2N}(x)$ *converges in distribution to* $S^*_\infty \equiv \max_{x\in A}\tilde{S}^*_\infty(x)$. *If* $\psi_1(x) > \psi_2(x)$ *for some* x, *then* $P(\max_x S_{2N}(x) < \varepsilon) \to 0$ *for every* $\varepsilon > 0$.

Proof: See Theorem 6 in Klecan, McFadden and McFadden (1991).

The distributions of the test statistics d^*_{2N} and s^*_{2N} under the null hypothesis depend on each case within the null hypothesis. The way to resolve this problem is to establish an asymptotic least favorable case under the null.

Theorem (LM) 5 *Suppose the assumptions and the definition $S_{2N}^{12*} = N^{1/2}s_{2N}^{12*}$ from the previous theorem. Suppose generalized exchangeability, and let $H(F_1(x_1), F_2(x_2))$ denote the joint density of (X_1, X_2) and let $F_{12}(x) = (F_1(x_1)+F_1(x_1))/2$. If (X_{1n}, X_{2n}) are second-degree stochastically maximal, then for N sufficiently large,*

$$P\left(S_{2N}^{12*} > \varepsilon|(X_{1n}, X_{2n}) \sim H(F_1, F_2)\right) > P\left(S_{2N}^{12*} > \varepsilon|(X_{1n}, X_{2n}) \sim H(F_{12}, F_{12})\right).$$

If, on the other hand, X_1 second-degree weakly stochastically dominates X_2, then for N sufficiently large,

$$P\left(S_{2N}^{12*} > \varepsilon|(X_{1n}, X_{2n}) \sim H(F_1, F_2)\right) \leq P\left(S_{2N}^{12*} > \varepsilon|(X_{1n}, X_{2n}) \sim H(F_{12}, F_{12})\right).$$

Proof: See Theorem 7 in Klecan, McFdden and McFadden (1991).

From the above theorem, we can see that the identical case has the largest size among all cases under the null hypothesis.

Since the statistics d_{2N}^* and s_{2N}^* have neither tractable finite-sample distributions, nor asymptotic distributions for which there are convenient computational approximations, Klecan, McFadden and McFadden suggested a method which assesses empirical distribution of the statistics by a Monte-Carlo simulation. For the actual procedures, please see Klecan, McFadden and McFadden. In the following theorem, they prove the asymptotic validity of the Monte-Carlo calculation.

Theorem (LM) 6 *Suppose the assumptions of theorem 4 hold for each pair of random variables X_i and X_j. Suppose the joint distribution of these random variables satisfies generalized exchangeability. Then, the significance level for $N^{1/2}\bar{s}_{2N}^*$ calculated by a Monte-Carlo simulation* [21] *approaches the probability $P\left(S_{2N}^{12*} > N^{1/2}\bar{s}_{2N}^*|(X_{1n}, X_{2n}) \sim H(F_{12}, F_{12})\right)$, as N and the number of Monte-Carlo iterations approaches infinity.*

[21] Briefly describing, the significance level is calculated as follows. Calculate the simulated statistics. Then count the fractions of these simulated statistics that are greater in magnitude that the observed values $\bar{d}_{2N}^*$ and $\bar{s}_{2N}^*$.

Proof) See Theorem 8 in Klecan, McFadden and McFadden (1991).

Klecan, McFadden and McFadden contains all proofs and detailed description of the procedures.

Table 2.1a
Sample Statistics for Cont-Returns Without T.C.
(Weekly Returns)

A weekly contrarian strategy with lag k stands for a contrarian trading strategy which rebalances the portfolio every week and whose weights for each stock depends on weekly returns of k weeks ago. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios.

	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	B-H
P1 (smallest)						
Mean	0.01524	0.00615	0.00480	0.00505	0.00388	0.00287
Std	0.03498	0.03600	0.03217	0.03379	0.03259	0.02736
Skew	0.03435	0.05289	0.01834	0.02650	0.02436	-0.02299
P2						
Mean	0.00970	0.00498	0.00354	0.00323	0.00348	0.00287
Std	0.03122	0.03059	0.02909	0.02945	0.03073	0.02456
Skew	0.02122	0.02230	-0.02039	0.01423	0.02515	-0.02463
P3						
Mean	0.00731	0.00447	0.00327	0.00292	0.00300	0.00272
Std	0.02712	0.02738	0.02556	0.02481	0.02455	0.02331
Skew	0.01479	-0.01451	-0.02029	-0.02036	-0.01187	-0.02325
P4						
Mean	0.00628	0.00393	0.00278	0.00286	0.00274	0.00252
Std	0.02516	0.02427	0.02404	0.02434	0.02323	0.02164
Skew	-0.02136	-0.01128	-0.01476	-0.02400	-0.01577	-0.01984
P5 (largest)						
Mean	0.00530	0.00359	0.00261	0.00259	0.00219	0.00229
Std	0.02440	0.02353	0.02288	0.02277	0.02236	0.02074
Skew	-0.01838	-0.01514	-0.01335	-0.02266	-0.01461	-0.02051

Table 2.1b
Sample Statistics for Cont-Returns Without T.C.
(Bi-Weekly returns)

A bi-weekly contrarian strategy with lag k stands for a contrarian trading strategy which rebalances the portfolio every other week and whose weights for each stock depends on bi-weekly returns of k bi-weeks ago. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios.

	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	B-H
P1 (smallest)						
Mean	0.02046	0.01008	0.00744	0.00697	0.00624	0.00582
Std	0.06231	0.06099	0.06018	0.05917	0.06022	0.04163
Skew	0.03167	0.02295	0.02798	0.03003	0.03724	-0.03733
P2						
Mean	0.01380	0.00687	0.00606	0.00691	0.00578	0.00574
Std	0.05670	0.05502	0.05422	0.05342	0.05232	0.03644
Skew	0.03997	0.01714	-0.02585	-0.00600	0.00827	-0.03648
P3						
Mean	0.01150	0.00632	0.00593	0.00582	0.00647	0.00544
Std	0.04784	0.04484	0.04545	0.04416	0.04346	0.03441
Skew	-0.01204	-0.02264	-0.02618	-0.03286	-0.02701	-0.03310
P4						
Mean	0.00998	0.00645	0.00481	0.00529	0.00559	0.00499
Std	0.04272	0.04193	0.04104	0.04124	0.03955	0.03103
Skew	0.01627	-0.01338	-0.03192	-0.03045	-0.01145	-0.02755
P5 (largest)						
Mean	0.00825	0.00500	0.00464	0.00522	0.00528	0.00454
Std	0.03999	0.03870	0.03947	0.03830	0.03866	0.02907
Skew	-0.01742	-0.02654	-0.03353	-0.03233	-0.02559	-0.02911

Table 2.2
Contrarian vs. B-H Strategies Without T.C.

A weekly (bi-weekly) contrarian strategy with lag k stands for a contrarian trading strategy which rebalances the portfolio every (other) week and whose weights for each stock depends on weekly (bi-weekly) returns of k weeks (k bi-weeks) ago. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

	Port 1 (smallest)	Port 2	Port 3	Port 4	Port 5 (largest)
Lag 1					
Weekly	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$
Bi-weekly	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$
Lag 2					
Weekly	$C \| B$	$C \| B$	$C \| B$	$C \| B$	$C \| B$
Bi-weekly	$C \| B$	$C \| B$	$C \doteq B$	$C \| B$	$C \| B$
Lag 3					
Weekly	$C \| B$	$C \| B$	$C \| B$	$C \doteq B$	$C \| B$
Bi-weekly	$C \| B$	$B \succeq_{SSD} C$	$C \doteq B$	$B \doteq C$	$B \succeq_{SSD} C$
Lag 4					
Weekly	$C \| B$	$C \| B$	$C \doteq B$	$C \| B$	$C \| B$
Bi-weekly	$C \| B$	$C \| B$	$C \doteq B$	$C \doteq B$	$C \| B$
Lag 5					
Weekly	$C \| B$	$C \| B$	$C \doteq B$	$C \doteq B$	$B \succeq_{SSD} C$
Bi-weekly	$C \| B$	$B \succeq_{SSD} C$	$C \doteq B$	$C \| B$	$C \| B$

Table 2.3
Contrarian vs. Value Weighted Market Index Without T.C.

A weekly (bi-weekly) contrarian strategy with lag k stands for a contrarian trading strategy which rebalances the portfolio every (other) week and whose weights for each stock depends on weekly (bi-weekly) returns of k weeks (k bi-weeks) ago. Value Weighted Market Index is CRSP VW-return. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

	Port 1 (smallest)	Port 2	Port 3	Port 4	Port 5 (largest)
Lag 1					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
Lag 2					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \| M$
Lag 3					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \| M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{FSD} C$
Lag 4					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \| M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \| M$
Lag 5					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{SSD} C$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{SSD} C$

Table 2.4a
Sample Statistics for Cont-Returns With T.C.
(Weekly Returns)

A weekly contrarian strategy which rebalances the portfolio every other week and whose rebalancing weight for each stock depends on weekly returns of 1 week ago are considered. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Transaction costs are one-way per-dollar transaction costs. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios.

	0.05%	0.1%	0.2%	0.3%	0.4%	0.5%	0.7%	1.0%	B-H
P1 (smallest)									
Mean	0.015	0.014	0.012	0.011	0.009	0.008	0.005	0.000	0.003
Std	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.027
Skew	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	-0.023
P2									
Mean	0.009	0.008	0.006	0.005	0.004	0.002	-0.001	-0.005	0.003
Std	0.031	0.031	0.031	0.031	0.031	0.031	0.031	0.031	0.025
Skew	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.020	-0.025
P3									
Mean	0.007	0.006	0.004	0.003	0.001	0.000	-0.003	-0.007	0.003
Std	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.023
Skew	0.015	0.015	0.015	0.014	0.014	0.014	0.014	0.013	-0.023
P4									
Mean	0.006	0.005	0.003	0.002	0.000	-0.001	-0.004	-0.009	0.003
Std	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.022
Skew	-0.021	-0.021	-0.021	-0.021	-0.021	-0.021	-0.021	-0.021	-0.020
P5 (largest)									
Mean	0.005	0.003	0.002	0.001	-0.001	-0.002	-0.005	-0.009	0.002
Std	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.021
Skew	-0.018	-0.018	-0.018	-0.018	-0.018	-0.018	-0.018	-0.019	-0.021

Table 2.4b
Sample Statistics for Cont-Returns With T.C.
(Bi-Weekly Returns)

A bi-weekly contrarian strategy which rebalances the portfolio every other week and whose rebalancing weight for each stock depends on bi-weekly returns of 1 bi-week (or two weeks) ago are considered. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Transaction costs are one-way per-dollar transaction costs. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios.

	0.05%	0.1%	0.2%	0.3%	0.4%	0.5%	0.7%	1.0%	B-H
P1 (smallest)									
Mean	0.020	0.019	0.017	0.016	0.014	0.013	0.010	0.005	0.006
Std	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.042
Skew	0.032	0.032	0.031	0.031	0.031	0.031	0.031	0.031	-0.037
P2									
Mean	0.013	0.012	0.011	0.009	0.008	0.006	0.003	-0.001	0.006
Std	0.057	0.057	0.057	0.057	0.056	0.056	0.056	0.056	0.036
Skew	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.039	-0.036
P3									
Mean	0.011	0.010	0.009	0.007	0.006	0.004	0.001	-0.003	0.005
Std	0.048	0.048	0.048	0.048	0.048	0.048	0.047	0.047	0.034
Skew	-0.012	-0.013	-0.013	-0.0144	-0.014	-0.015	-0.015	-0.016	-0.033
P4									
Mean	0.009	0.008	0.007	0.005	0.004	0.003	-0.001	-0.005	0.005
Std	0.043	0.043	0.043	0.043	0.042	0.042	0.042	0.042	0.031
Skew	0.016	0.016	0.016	0.015	0.015	0.015	0.014	0.013	-0.028
P5 (largest)									
Mean	0.008	0.007	0.005	0.004	0.002	0.001	-0.002	-0.007	0.005
Std	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.029
Skew	-0.018	-0.018	-0.018	-0.018	-0.018	-0.019	-0.019	-0.020	-0.029

Table 2.5
Contrarian vs. B-H Strategies With T.C.

A Weekly (bi-weekly) contrarian strategy which rebalances the portfolio every (other) week and whose rebalancing weight for each stock depends on weekly (bi-weekly) returns of 1 week (1 bi-week) ago are considered. B-H strategy stands for the strategy which buys an equally weighted portfolio at the initial date and holds it without any portfolio rebalancing. Transaction costs are one-way per-dollar transaction costs. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

tc(%)	Port 1 (smallest)	Port 2	Port 3	Port 4	Port 5 (largest)
0.05					
Weekly	$C \succeq_{FSD} B$	$C \| B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \| B$
Bi-weekly	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \| B$
0.1					
Weekly	$C \succeq_{FSD} B$	$C \| B$	$C \succeq_{FSD} B$	$C \| B$	$C \| B$
Bi-weekly	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \succeq_{FSD} B$	$C \| B$
0.2					
Weekly	$C \succeq_{FSD} B$	$C \| B$	$C \| B$	$C \| B$	$B \succeq_{SSD} C$
Bi-weekly	$C \succeq_{FSD} B$	$C \| B$	$C \succeq_{FSD} B$	$C \| B$	$C \| B$
0.3					
Weekly	$C \succeq_{FSD} B$	$C \| B$	$C \| B$	$B \succeq_{SSD} C$	$B \succeq_{SSD} C$
Bi-weekly	$C \succeq_{FSD} B$	$C \| B$	$C \| B$	$C \| B$	$B \succeq_{SSD} C$
0.4					
Weekly	$C \succeq_{FSD} B$	$C \| B$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
Bi-weekly	$C \succeq_{FSD} B$	$C \| B$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{SSD} C$
0.5					
Weekly	$C \succeq_{FSD} B$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
Bi-weekly	$C \succeq_{FSD} B$	$C \| B$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
0.7					
Weekly	$C \| B$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
Bi-weekly	$C \| B$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
1.0					
Weekly	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$
Bi-weekly	$B \succeq_{SSD} C$	$B \succeq_{SSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$	$B \succeq_{FSD} C$

Table 2.6
Contrarian vs. Value Weighted Market Index With T.C.

A Weekly (bi-weekly) contrarian strategy which rebalances the portfolio every (other) week and whose rebalancing weight for each stock depends on weekly (bi-weekly) returns of 1 week (1 bi-week) ago are considered. Value Weighted Market Index is the CRSP VW-return. Transaction costs are one-way per-dollar transaction costs. Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

tc(%)	Port 1 (smallest)	Port 2	Port 3	Port 4	Port 5 (largest)
0.05					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
0.1					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \succeq_{FSD} M$	$C \succeq_{FSD} M$
0.2					
Weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{SSD} C$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$C \| M$
0.3					
Weekly	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{FSD} C$
0.4					
Weekly	$C \| M$	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
Bi-weekly	$C \| M$	$C \| M$	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$
0.5					
Weekly	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
Bi-weekly	$C \| M$	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
0.7					
Weekly	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
Bi-weekly	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
1.0					
Weekly	$M \succeq_{SSD} C$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$
Bi-weekly	$C \| M$	$M \succeq_{SSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$	$M \succeq_{FSD} C$

Table 2.7
Break-Even One-Way Transaction Costs

Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios. All decisions are made at the 5 percent significance level.

The second column stands for the level of transaction costs which allow the contrarian trading strategy with lag 1 to second-degree stochastically dominate the buy-and-hold strategy. The third column stands for the level of transaction costs which prevent two strategies from second-degree stochastically dominating each other. The last column stands for the level of transaction costs which allow the buy-and-hold strategy to second-degree stochastically dominate the contrarian trading strategy with lag 1.

Portfolio	Contrarian S.S.D. B-H	Neither S.S.D. The Other	B-H S.S.D. Contrarian
Weekly	(%)	(%)	(%)
Port 1 (smallest)	[0.0 , 0.534]	[0.534 , 0.086]	[0.086 , -]
Port 2	[0.0 , 0.012]	[0.012 , 0.448]	[0.448 , -]
Port 3	[0.0 , 0.160]	[0.160 , 0.301]	[0.301 , -]
Port 4	[0.0 , 0.090]	[0.090 , 0.244]	[0.244 , -]
Port 5 (largest)	[0.0 , 0.045]	[0.045 , 0.196]	[0.196 , -]
Bi-Weekly	(%)	(%)	(%)
Port 1 (smallest)	[0.0 , 0.544]	[0.544 , 0.940]	[0.940 , -]
Port 2	[0.0 , 0.167]	[0.167 , 0.532]	[0.532 , -]
Port 3	[0.0 , 0.300]	[0.300 , 0.386]	[0.386 , -]
Port 4	[0.0 , 0.161]	[0.161 , 0.313]	[0.313 , -]
Port 5 (largest)	[0.0 , 0.033]	[0.033 , 0.235]	[0.235 , -]

Table 2.8
Average Portfolio Characteristics

Weekly returns are formed from Wednesday to Tuesday daily returns. Data are from CRSP daily return file for NYSE and AMEX stocks. 510 stocks with a complete return history for the period 1963.1.1 to 1991.12.31 are selected which have data missing no more than 30 data points. The stocks were then sorted into five portfolios based on their initial market value. Once they were sorted, no rebalancing was made among portfolios.

Beg, Mid, and End stand for the values at the beginning of the sample period, at the mid of the sample period, and at the end of the sample period, respectively. 1/16 stands for the half of minimum bid-ask spread for NYSE and AMEX stocks whose prices are greater than one dollar. The bid-ask spread over price ratio can be a crude measure for per-dollar transaction costs.

Variable	Port 1 (smallest)	Port 2	Port 3	Port 4	Port 5 (largest)
Shares(000)					
Beg	1,236	2,016	3,815	8,217	31,105
Mid	3,408	7,046	15,700	26,922	73,691
End	13,273	116,914	53,900	105,980	220,356
Price					
Beg	10.46	21.72	33.14	40.66	55.92
Mid	13.29	20.06	26.16	32.54	39.38
End	17.21	31.56	34.93	38.52	46.40
Size(000)					
Beg	7,471	29,689	101,798	291,397	1,786,408
Mid	43,484	147,610	418,673	866,755	3,402,730
End	301,320	6,454,960	2,008,592	4,704,857	11,443,881
$\frac{(1/16)}{Price}(\%)$					
Beg	0.60	0.29	0.19	0.15	0.11
Mid	0.47	0.31	0.24	0.19	0.16
End	0.36	0.20	0.18	0.16	0.13

Table 2.9
Sample Statistics for Loser and Winner Portfolios

Tests are based on CRSP monthly return file for NYSE and AMEX stocks for the sample period from 1926.1 to 1991.12. Every year starting from 1946.1, according to previous five year period's cumulative return performances, 50 best stocks and 50 worst stocks are selected out of all stocks which have complete return history for the five year period. Loser portfolio (L) is an equally weighted portfolio composed of 50 worst stocks and winner portfolio (W) is an equally weighted portfolio composed of 50 best stocks. For the following 5 year period, the performances of loser and winner portfolios are compared on a monthly return basis including and excluding January returns. VW Market is a CRSP value weighted market index.

	Loser	Winner	VW Market	L+VWM	W+VWM
With Jan. Returns					
1926.1 - 1991.12					
Mean	0.02141	0.00944	0.01022	0.01582	0.00983
Std	0.11753	0.07025	0.05104	0.11186	0.08362
Skew	0.16469	0.04531	0.03010	0.15269	0.04858
1941.1 - 1991.12					
Mean	0.01550	0.00834	0.01022	0.01285	0.00928
Std	0.07971	0.06394	0.04187	0.07987	0.07276
Skew	0.07990	-0.02742	-0.02978	0.05867	-0.05377
Without Jan. Returns					
1926.1 - 1991.12					
Mean	0.01172	0.00746	0.00955	0.01063	0.00850
Std	0.10735	0.06930	0.05104	0.10599	0.08309
Skew	0.15232	0.03843	0.02931	0.14474	0.04202
1941.1 - 1991.12					
Mean	0.00748	0.00599	0.00946	0.00847	0.00772
Std	0.07071	0.06147	0.04080	0.07360	0.07044
Skew	0.03841	-0.04548	-0.03288	-0.06129	-0.06377

Table 2.10
A Test for Long-Horizon Return Reversals

Tests are based on CRSP monthly return file for NYSE and AMEX stocks for the sample period from 1941.1 to 1991.12. Every year starting from 1946.1, according to previous five year period's cumulative return performances, 50 best stocks and 50 worst stocks are selected out of all stocks which have complete return history for the five year period. Loser portfolio (L) is an equally weighted portfolio composed of 50 worst stocks and winner portfolio (W) is an equally weighted portfolio composed of 50 best stocks. For the following 5 year period, the performances of loser and winner portfolios are compared on a monthly return basis including January returns and excluding January returns. VW Market is a CRSP value weighted market index. All decisions are made at the 10 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

	Loser Winner	Loser VW Market	Winner VW Market	L+VW W+VW
With Jan. Returns	$L \succeq_{SSD} W$	$L \parallel VW$	$VW \succeq_{SSD} W$	$L + VW \succeq_{FSD} W + VW$
Without Jan. Returns	$L \parallel W$	$VW \succeq_{SSD} L$	$VW \succeq_{SSD} W$	$L + VW \parallel W + VW$

Table 2.11
Sample Statistics for Each Day of the Week Returns

Data are from CRSP daily return file for NYSE and AMEX stocks for the period from 1962.8.1 to 1991.12. VWM stands for the value weighted CRSP market index and EWM stands for the equally weighted CRSP market index.

	Mon	Tue	Wed	Thu	Fri
VW Market					
Mean	-0.000935	0.000472	0.001166	0.000493	0.001027
Std	0.010224	0.007839	0.008237	0.007728	0.00773
Skew	-0.016446	0.004762	0.008567	0.002553	-0.006532
EW Market					
Mean	-0.001162	0.000089	0.001447	0.001106	0.002132
Std	0.009542	0.007126	0.007682	0.007254	0.007124
Skew	-0.013679	-0.006383	0.009200	-0.004146	-0.002655
SP500 Index					
Mean	-0.001055	0.000427	0.001043	0.000272	0.000787
Std	0.010837	0.008450	0.008578	0.008113	0.008214
Skew	-0.018442	0.006606	0.008853	0.003540	-0.007271
AT%T					
Mean	-0.000624	0.000915	0.000749	0.000346	0.000738
Std	0.013323	0.012076	0.011874	0.011380	0.011003
Skew	-0.018577	0.015249	0.011477	0.003744	-0.005098
EK					
Mean	-0.000130	0.000710	0.001318	0.000140	0.000197
Std	0.018300	0.016255	0.017145	0.014621	0.013718
Skew	-0.026159	0.021194	0.015152	0.009912	-0.007818
IBM					
Mean	0.000279	0.000817	0.000775	-0.000067	0.000241
Std	0.015565	0.014395	0.013333	0.012794	0.012298
Skew	-0.020565	0.011295	0.010572	0.009596	0.007153

Table 2.12
A test for the Weekend Effect

Data are from CRSP daily return file for NYSE and AMEX stocks for the period from 1962.8.1 to 1991.12. VWM stands for the value weighted CRSP market index and EWM stands for the equally weighted CRSP market index. All decisions are made at 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$||$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

	VWM	EWM	SP500	AT&T	EK	IBM
M–T	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\doteq$
M–W	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{SSD}$
M–Th	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\doteq$	$\|\|$
M–F	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{SSD}$	$\preceq_{SSD}$
T–W	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\doteq$	$\doteq$	$\doteq$
T–Th	$\doteq$	$\preceq_{FSD}$	$\doteq$	$\doteq$	$\doteq$	$\succeq_{FSD}$
T–F	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\doteq$	$\doteq$	$\|\|$
W–Th	$\succeq_{FSD}$	$\succeq_{FSD}$	$\succeq_{FSD}$	$\doteq$	$\succeq_{FSD}$	$\succeq_{FSD}$
W–F	$\doteq$	$\preceq_{FSD}$	$\doteq$	$\doteq$	$\|\|$	$\doteq$
Th–F	$\preceq_{FSD}$	$\preceq_{FSD}$	$\preceq_{FSD}$	$\doteq$	$\doteq$	$\doteq$

Table 2.13
Performance Measures for 5 Mutual Funds

ASA ASA Ltd. is a closed-end investment company with at least 50% of its funds in South African gold stocks.

GAM General American Investment is a closed-end regulated management company investing primarily in medium and high-quality growth stocks, with the aim of long-term capital appreciation.

JHI John Hancock Investors Trust is a closed-end diversified investment company whose main objective is income distribution to shareholders, and main holdings are in debt securities, up to 50% of which are direct placements.

NGS Niagara Share Corporation is a closed-end management company investing primarily in common stocks, seeking high earnings and dividend potential.

TY Tri-Continental Corp. is a closed-end diversified management company investing in common stocks and equivalents with the aims of long-term appreciation and growth in income.

Sample period is from 1971.6.15 to 1992.12.31. Returns are weekly returns. All descriptions about the mutual funds are quoted from Klecan, McFadden and McFadden (1991).

	ASA	GAM	JHI	NGS	TY	VWRETD
Mean	0.0039	0.0033	0.0023	0.0026	0.0027	0.0024
Std	0.0540	0.0338	0.0233	0.0325	0.0284	0.0227
Skew	0.0370	-0.0261	-0.0160	-0.0169	-0.0183	-0.0236
Beta	0.3540	0.9326	0.3763	0.7617	0.8589	
Sharp Measure	0.0477	0.0578	0.0390	0.0371	0.0486	
Jensen Measure	0.00222	0.00103	0.00054	0.00046	0.00053	
Treynor Measure	0.00727	0.00209	0.00241	0.00158	0.00161	
Appraisal Ratio	0.0421	0.0504	0.0267	0.0195	0.0355	

Table 2.14
A test for 5 mutual funds

Sample period is from 1971.6.15 to 1992.12.31. Returns are weekly returns. For description about each mutual fund, please see table 2.13. M stands for CRSP value weighted market index. Decisions are made at 5 percent significance level.

Funds Alone	$JHI \succeq_{SSD} NGS$, $TY \succeq_{SSD} NGS$, $M \succeq_{SSD} JHI$, $M \succeq_{SSD} NGS$
With VW Market Index Combined	$JHI + M \succeq_{SSD} NGS + M$, $TY \succeq_{SSD} NGS + M$
Indistinguishable	$NGS + M \doteq TY + M$

Table 2.15
Monte-Carlo Size and Power Comparison Between Conventional SD Test and Weighted SD Test

Processes for X_1 and X_2 are assumed as

$$X_{i,n} = (1-\lambda)[\alpha_i + \beta_i(\sqrt{\rho}C_n + \sqrt{1-\rho}Z_{i,n})] \; + \; X_{i,n-1}, i = 1,2$$

where C_n is i.i.d. standard normal over n, $Z_{i,n}$ is i.i.d. standard normal over i and over n, C_n and $Z_{i,n}$ are independent, and α_i, β_i, ρ, λ are parameters satisfying $\rho \in [0,1)$ and $\lambda \in (-1,1)$. Then, $E(X_{i,n}) = \alpha_i$, $Var(X_{i,n}) = \beta_i^2$, and $Corr(X_{i,n}, X_{j,m}) = \rho\lambda^{|n-m|}$, and the processes are α-mixing, and generalized exchangeable.[1]

F10, F05 and F01 stand for the probability of rejecting the first-degree null at 10%, 5% and 1% significance levels, respectively.
S10, S05 and S01 stand for the probability of rejecting the second-degree null at 10%, 5% and 1% significance levels, respectively.
N stands for the number of time-series observations, and I stands for the number of Monte-Carlo experiments.

Continued —

[1]The processes considered here are exactly the same processes in Klecan, McFadden and McFadden (1991) except for parameter values. All the descriptions of the processes are quoted from Klecan, McFadden and McFadden (1991).

(When X_1 first- and second-degree stochastically dominates X_2)

	F10	F05	F01	S10	S05	S01
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 4, \beta_2 = 4$ $\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.03	0.00	0.00	0.05	0.04	0.01
Weighted SD Test	0.05	0.03	0.01	0.17	0.08	0.04
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 4, \beta_2 = 4$ $\lambda = 0.1, \rho = 0.9, N = 100, I = 100$						
Conventional SD Test	0.05	0.03	0.02	0.04	0.02	0.02
Weighted SD Test	0.06	0.06	0.03	0.08	0.06	0.05
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 4, \beta_2 = 4$ $\lambda = 0.9, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.00	0.00	0.00	0.00	0.00	0.00
Weighted SD Test	0.00	0.00	0.00	0.10	0.10	0.08
$\alpha_1 = 0, \alpha_2 = -0.2, \beta_1 = 4, \beta_2 = 4$ $\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.17	0.09	0.01	0.12	0.06	0.03
Weighted SD Test	0.11	0.08	0.03	0.24	0.12	0.06
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 4, \beta_2 = 4$ $\lambda = 0.1, \rho = 0.9, N = 1000, I = 1000$						
Conventional SD Test	0.00	0.00	0.00	0.00	0.00	0.00
Weighted SD Test	0.01	0.00	0.00	0.01	0.01	0.00

Continued —

(When X_1 second- but not first-degree stochastically dominates X_2)

	F10	F05	F01	S10	S05	S01
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.95	0.94	0.86	0.01	0.00	0.00
Weighted SD Test	1.00	1.00	1.00	0.00	0.00	0.00
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.9, N = 100, I = 100$						
Conventional SD Test	0.21	0.18	0.12	0.00	0.00	0.00
Weighted SD Test	0.50	0.46	0.37	0.00	0.00	0.00
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.9, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	1.00	1.00	0.94	0.00	0.00	0.00
Weighted SD Test	1.00	1.00	1.00	0.00	0.00	0.00
$\alpha_1 = 0, \alpha_2 = -0.2, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.99	0.99	0.99	0.30	0.29	0.26
Weighted SD Test	1.00	1.00	1.00	0.12	0.12	0.09
$\alpha_1 = 0, \alpha_2 = -0.4, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	1.00	1.00	0.99	0.11	0.09	0.08
Weighted SD Test	1.00	1.00	1.00	0.09	0.09	0.05
$\alpha_1 = 0, \alpha_2 = -1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.9, N = 1000, I = 1000$						
Conventional SD Test	0.74	0.67	0.55	0.00	0.00	0.00
Weighted SD Test	0.99	0.99	0.98	0.00	0.00	0.00

Continued —

(When X_1 does *neither* second- nor first-degree stochastically dominate X_2)

	F10	F05	F01	S10	S05	S01
$\alpha_1 = 0, \alpha_2 = 1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.96	0.94	0.91	0.98	0.98	0.93
Weighted SD Test	1.00	1.00	0.99	0.98	0.98	0.98
$\alpha_1 = 0, \alpha_2 = 1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.9, N = 100, I = 100$						
Conventional SD Test	0.26	0.24	0.20	0.35	0.26	0.22
Weighted SD Test	0.50	0.49	0.41	0.69	0.63	0.47
$\alpha_1 = 0, \alpha_2 = 1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.9, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	0.97	0.97	0.93	1.00	1.00	0.98
Weighted SD Test	1.00	1.00	1.00	1.00	1.00	1.00
$\alpha_1 = 0, \alpha_2 = 0.2, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	1.00	1.00	1.00	0.68	0.67	0.60
Weighted SD Test	1.00	1.00	1.00	0.51	0.50	0.45
$\alpha_1 = 0, \alpha_2 = 0.4, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	1.00	1.00	1.00	0.82	0.81	0.81
Weighted SD Test	1.00	1.00	1.00	0.73	0.73	0.67
$\alpha_1 = 0, \alpha_2 = 0.6, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.5, N = 100, I = 100$						
Conventional SD Test	1.00	0.98	1.97	0.89	0.86	0.86
Weighted SD Test	1.00	1.00	1.00	0.81	0.81	0.76
$\alpha_1 = 0, \alpha_2 = 1, \beta_1 = 1, \beta_2 = 4$						
$\lambda = 0.1, \rho = 0.9, N = 1000, I = 1000$						
Conventional SD Test	0.73	0.65	0.53	0.88	0.82	0.70
Weighted SD Test	0.99	0.99	0.98	0.99	0.99	0.99

Figure 2.1
First-Degree Stochastic Dominance
(A Stochastically Dominates B)

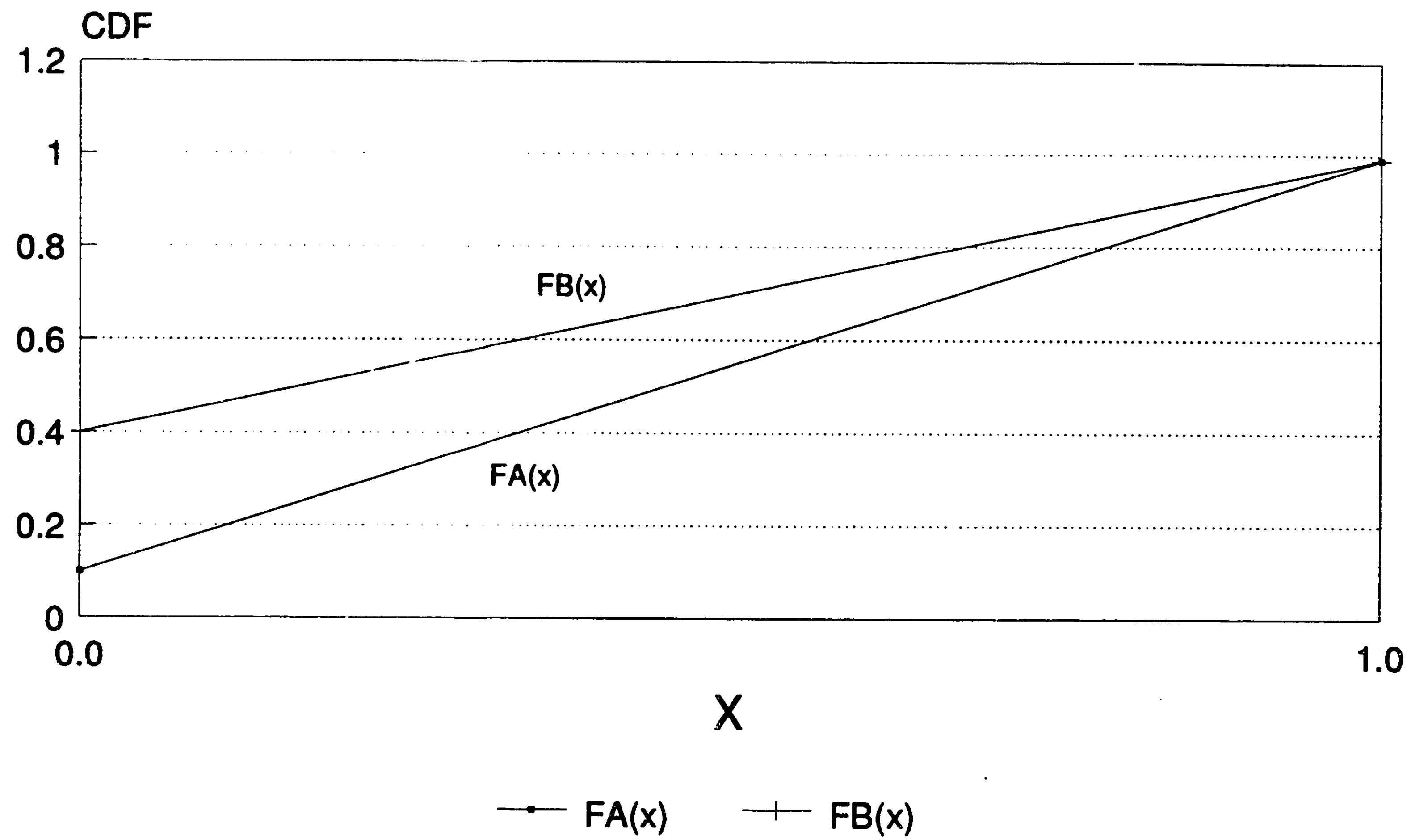


Figure 2.2
Second Degree Stochastic Dominance
(A Stochastically Dominates B)

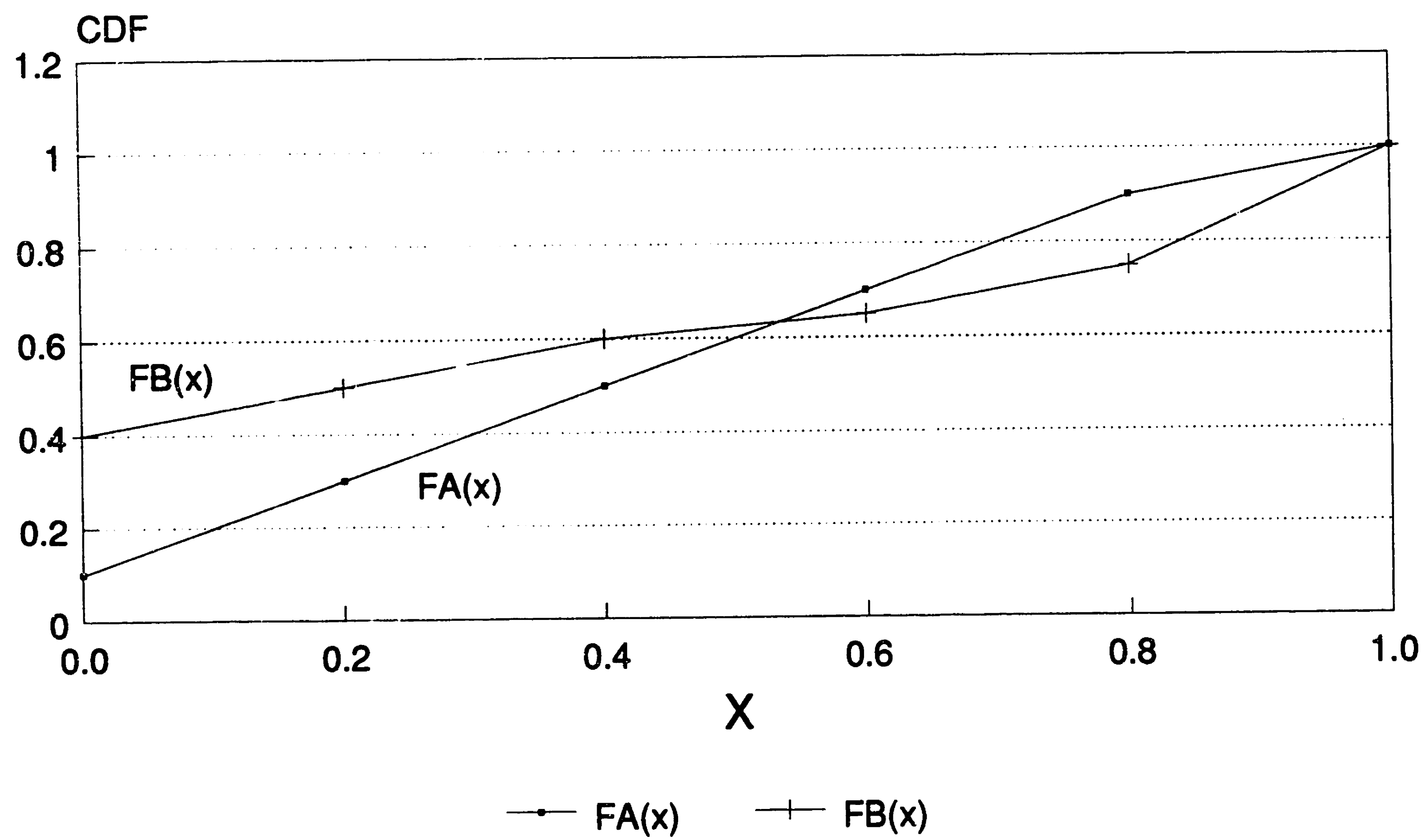


Figure 2.3
Ex : S-dominance but not F-dominance
(Statistically)

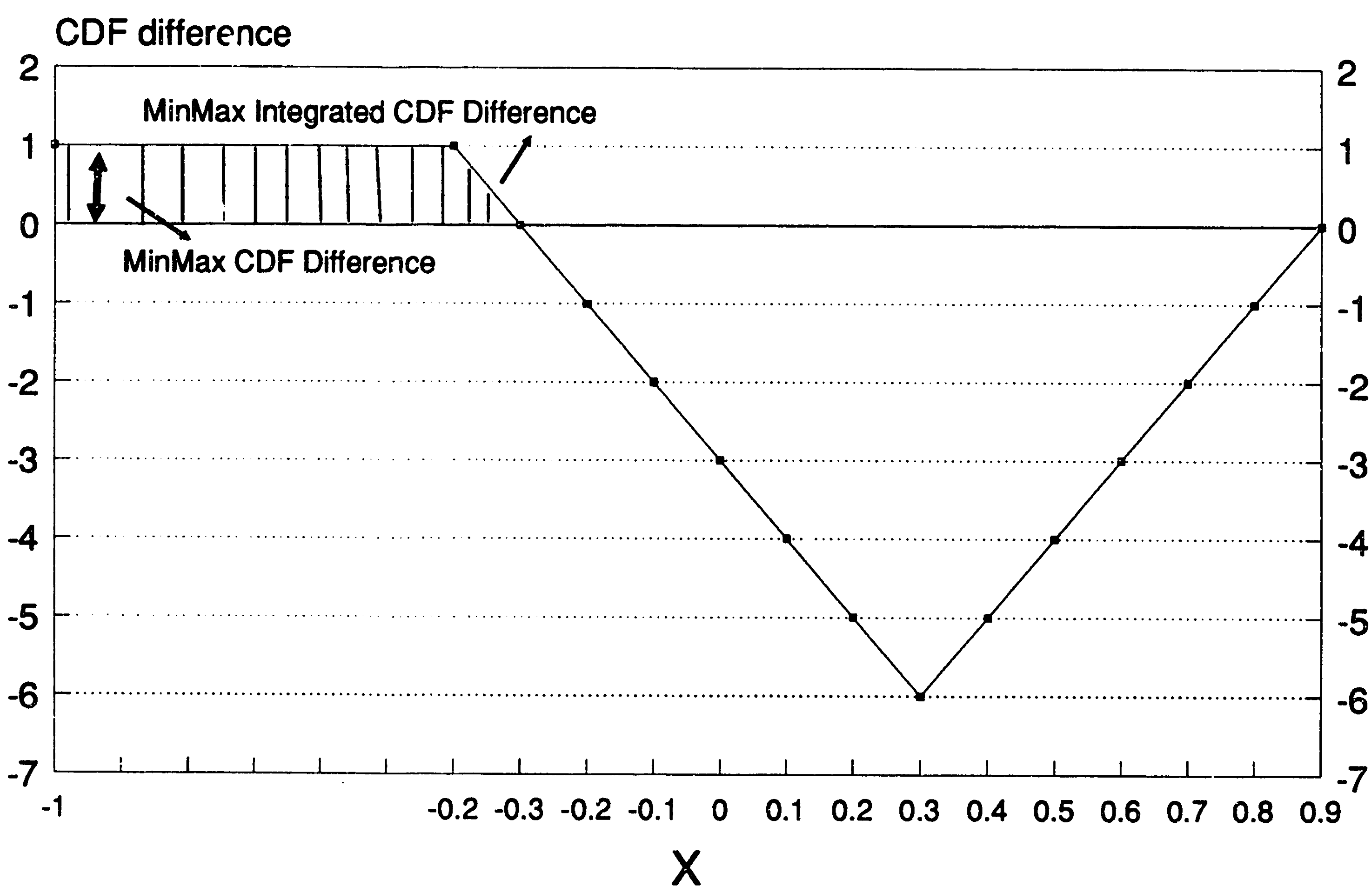
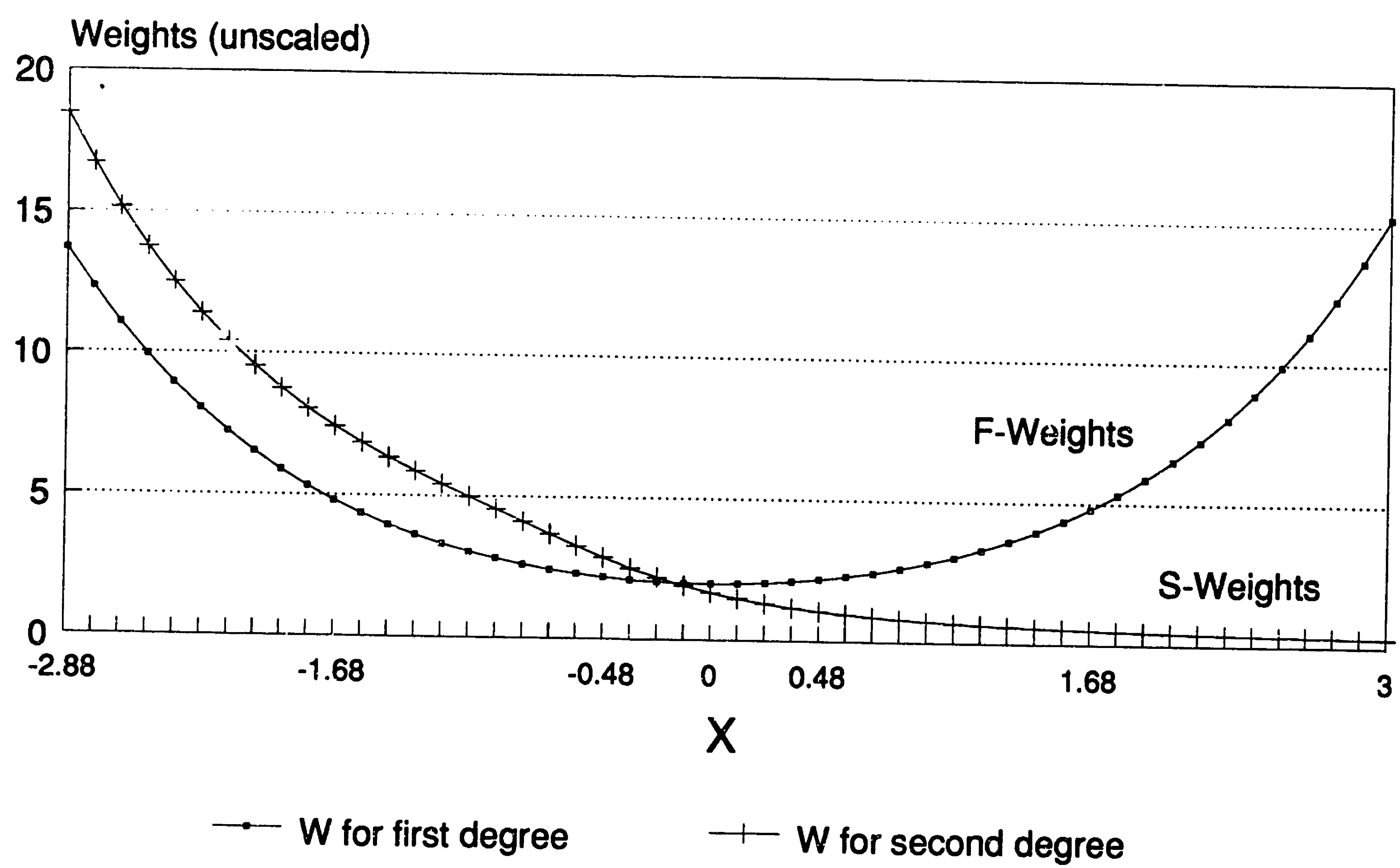


Figure 2.4
Weights for SD Tests
(when dist. is logistic with (0,1))



Chapter III

Econometric Implementation of Trading Strategies: Neural Network and Moving Average Rules

1 Introduction

There have been two kinds of financial analysts: the fundamental analyst who conducts fundamental analysis, and the technical analyst who conducts technical analysis. Fundamental analysis is concerned with economic valuation of financial assets whereas technical analysis is related to chasing trends. In other words, technical analysis is regarded to be useful by investors who believe that previous movements of asset returns (or prices) are valuable information for predicting future movements of the asset returns (or prices). For academic researchers, 'technical analysis' had been meaningless at least until the financial markets were believed to be efficient.[1] If, for example, asset prices follow random walks, it is of no use to try to predict future asset prices with current information of asset prices.

Recently, there has been growing empirical evidence that some financial asset prices do not follow random walks. For example, Lo and MacKinlay (1988) showed that stock prices did not follow random walks in weekly investment horizons. Also Debondt and Thaler (1985, 1987) showed that long horizon stock returns had a tendency of mean-reversion. Although the deviation from the random walk does not necessarily imply the inefficiency of the market or equivalently the effectiveness of technical analysis (i.e. see Leroy (1973)), it certainly provides a good possibility for technical analysis to be useful.

With respect to stock returns, contrarian trading strategies that are defined as strategies which buy previous losers and sell previous winners were shown to generate profits or high returns (i.e. see Lehmann (1990), Lo and MacKinlay (1990)). Shin (1992) provided economic implications for stock return predictability by imposing stochastic dominance criteria on contrarian trading strategies and passive buy-and-hold strategies. Since a contrarian trading strategy is supposed to exploit the serial and cross-sectional autocorrelations of asset returns (i.e. see Lo and MacKinlay (1990)), it can be classified as one of the most simple technical trading rules.

[1] According to Fama's definition, this efficiency is a weak-form efficiency.

Although it is difficult to assess the proper risks of trading strategies, contrarian trading strategies combined with low levels of transaction costs certainly provide a good investment opportunity as shown in Shin (1992).

Whereas the deviation from the random walk hypothesis suggests that technical analysis can be useful, the inability to reject the random walk hypothesis does not necessarily imply that charting is just a waste of time. Highly nonlinear dynamics may not be detected by conventional tests that are usually related to sample means and sample variances. For example, the variance ratio tests examine only the first and the second moments of samples. Therefore, there may well be the case where drawing a chart can actually help predict future movements of random variables, even though conventional statistical tests cannot. Among the highly nonlinear system is a system which seemingly follows random walks, but that is in fact generated by highly nonlinear *deterministic* process. This process is called deterministic 'chaos'. Hsieh (1991) investigated the stock market to see whether or not it is governed by chaotic dynamics. He found no evidence of chaotic behavior in stock returns. A priori, it seems to be difficult to advocate the possibility of chaotic behavior in financial asset returns since asset returns seem more likely to be stochastic rather than to deterministic.

Contrary to the popularity of technical analysis in practice, there are only a few papers that have tried to advocate the usefulness of technical analysis. Among them are Treynor and Furguson (1985) (T-F), Brown and Jennings (1989) (B-J), and Brock, Lakonishok and LeBaron (1992) (BLL). T-F provided a theoretical background on how technical analysis can be useful by using Bayes Rules. B-J advocated the value of previous realization of asset returns in the environment where information is not fully revealing. The main point made by T-F is that once an investor receives new information, he can assess the probability of receiving the information ahead of the market by investigating the previous returns. In the similar context, B-J advocated the value of previous realization of asset returns in the environment where information is not fully revealing. BLL supported the benefit of technical analysis from the empirical perspective by showing that technical trading rules could detect some patterns in stock returns that are not consistent with the four popular null models; random walks, AR(1), GARCH-M, and EGARCH.

In this paper, attempts are made to assess the effectiveness of technical trading rules on foreign exchange rates from the stochastic dominance point of view. Along with the most popular technical trading rules which are moving average trading rules, a relatively new econometric method which is called 'artificial neural network' will be implemented to develop new kinds of technical trading rules.

The concept of stochastic dominance and the method for stochastic dominance test is well explained in Shin (1992). Briefly explaining about the concept, it is as follows. Stochastic dominance is a concept about stochastic ordering among random variables combined with preferences. It is said that a random variable X first-degree stochastically dominates another random variable Y if all insatiable investors prefer X to Y. In a similar way, X is said to second-degree stochastically dominate Y if all insatiable and risk-averse investors prefer X to Y. Detailed description about the theory and the test method of stochastic dominance is contained in Shin (1992). Artificial neural network is a nonparametric regression method (or a model) that originated in cognitive science areas inspired by the brain structure. For example, consider the procedure that human beings can recognize a rose. First, human beings observe objects and learn that they are roses if their appearances belong to a certain category. After a number of repeated learning, human beings can associate the appearance of an object with the name of the object, which is a rose. In artificial neural networks, the appearance of an object is called an 'input' and the name of an object is called an 'output'. The repeated learning is called 'training the network with training samples'. In most cases, it is still not known how human brains work in associating inputs (appearances of objects) with outputs (names of objects). Therefore, the associating procedure is regarded as a black box (see Figure 3.1).

Artificial neural network can be regarded as a highly non-linear statistical model which can imitate the black box in a very general but tractable way. Contrary to the early intuition by Minsky and Papert (1969) that only a certain type of functions would be approximated by artificial neural networks, recent works about artificial neural networks rigorously showed that almost all kinds of functions can be approximated by artificial neural networks (i.e. see Gallant and White (1988), Hornik, Stinchcombe and White (1989)). Due to its general applicability, it has been used in vast areas of science and engineering from cognitive science to artificial intelligence, and even to computer hardware design. It was only recently applied to economics. For example, White (1988) implemented neural network to predict IBM daily stock returns, but unfortunately, found that using neural network did not out-perform the random walk model. Utans and Moody (1991) used neural network to predict corporate bond ratings and showed that it performs much better than a linear bond rating predictor.

Since the main purpose of this paper is to test the usefulness of technical trading rules with the stochastic dominance criteria, attempts are not made to refine the theory of neural network, or to exploit all kinds of complicated neural network models.[2] Instead, one of the simplest neural network model — 'single hidden layer feedforward models' — is used to predict the following movements of exchange rates. Several issues concerning about the proper network architectures will also be briefly discussed.

Before investigating technical trading rules, a simple specification test will be conducted to see if the null hypothesis of random walks for exchange rates can be rejected by the variance ratio test method by Lo and MacKinlay (1988).[3] Briefly explaining the results, they are as follows. The null hypothesis that foreign exchange

[2]Exploiting all kinds of possible models to see if there is any extraordinary profit opportunity will inevitably face a severe selection bias. Nonetheless, the bias cannot be completely avoided if only the past data are involved in tests.

[3]The test method by Lo and MacKinlay is adopted since it is also robust to certain kinds of heteroskedasticity.

rates follow random walks is usually not rejected at 5 percent significance level for all exchange rates. Therefore, the effectiveness of linear auto regressive models in predicting returns are implicitly denied. Both neural network models and moving average trading rules are found to be effective in predicting returns in the sense that i) 'buy' returns have higher means than 'sell' returns with standard deviations and skewnesses comparable, ii) predicted returns by the neural network models are positively correlated with actual returns, and iii) 'buy' returns usually first-degree stochastically dominate 'sell' returns.

This paper is organized as follows. Section 2 will briefly describe about the data and will show the random characteristics of the returns for exchange rates. Section 3 contains a brief overview of neural network and a single hidden layer feedforward model. Estimation of the models and the empirical assessment of the estimated models will also be reported in section 3. Section 4 will describe the Monte-Carlo experiments for comparing the finite sample properties of two estimation methods, which are the non-linear least square method and the back-propagation method. Section 5 will describe the moving average trading rules and empirical assessment of the trading rules. Economic evaluation of the neural network models and the moving average trading rules will be performed with stochastic dominance criteria in section 6. Summary and conclusions will follow in section 7. Finally, appendix will follow section 7.

2 Random Characteristics of Exchange Rates

2.1 Data Summary

The data used in this paper are from the EHRA macro data tape from the Federal Reserve Bank. The data contain daily closing exchange rates for 6 currencies of Canadian Dollar, French Franc, German Mark, Japanese Yen, Swiss Franc, and British Pound for the period of 1971.1.5 to 1992.7.7. Continuously compounded returns are computed from the data. Weekly returns are formed from Wednesday to next Tuesday daily returns. Missing data are replaced with the same exchange rates on the previous date (or zero returns).

2.2 Variance Ratio Test

Let's denote P_t as a dollar denominated exchange rate at time t and $R_t = \log P_t - \log P_{t-1} = X_t - X_{t-1}$ as a return from holding a foreign currency from time $t-1$ to t ignoring the interest payment for the foreign currency. Consider the following discrete time exchange rate process:

$$X_t = X_{t-1} + \mu + \epsilon_t. \qquad (1)$$

Under the null hypothesis that X_t follows random walks (or ϵ_t is random over time), Lo and MacKinlay (1988) developed a variance ratio test which is robust to some kinds of heteroskedasticity. The test is based on the fact that if two random variables are random, *the variance of the sum* of two random variable should be the same with *the sum of the variances* of two random variables. For more interested readers, the null hypothesis and the theorem by Lo and MacKinlay (1988) are restated without proof in the appendix.

A nice interpretation can be made about the variance ratio test statistics. If VR(q) denotes variance ratio statistics of q periods ((Variance of q period returns)/(Sum of variances of q one period return)), then VR(2q)/VR(q) can be interpreted as a first-order autocorrelation coefficient of q period returns. For example, VR(2) can be interpreted as a first-order autocorrelation coefficient of one period returns. Therefore, if VR(2q)/VR(q) is greater than one, it means that q period returns are positively first-order autocorrelated.

Table 3.1 describes the random characteristics of returns from the exchange rates. Daily returns show slightly positive autocorrelations for all exchange rates, but most of the autocorrelations are not statistically significant. Weekly returns show similar characteristics. Except for the Canadian $, returns are positively autocorrelated for all different horizons (1 week, 2 weeks, 4 weeks, 8 weeks), but most of the autocorrelations are not statistically significant. Thereforc, the random walk hypothesis is not generally rejected for all exchange rates, which implies that linear AR models would hardly be effective in predicting returns.

The variance ratio test, however, considers finite sample moments. In other words, the test utilizes the characteristics of the returns only up to the second moments. If the returns are generated by some highly non-linear process (extreme case would be a deterministic chaos), then it may well be the case that the test cannot detect the deviation from the random walks. To see if a highly non-linear econometric model can catch the predictability of the returns, neural network models will try to be implemented in the next section.

3 Artificial Neural Network

Inspired by the structure of the brain, artificial neural network models were originally developed by cognitive scientists. Like the brain structure with many neurons, a neural network model is built by adequately connecting different nodes (or perceptrons), and also by properly defining the interaction among the nodes. Neural network models are generally concerned with input-output pairs. In other words, neural network models usually try to identify the relationships between inputs and outputs from the samples of input-output pairs for making out-of-sample predictions.[4] For the application purpose, the essential question concerning artificial neural network models has been concentrated on "what kind of functions can artificial neural network models approximate?" Unfortunately, the initial response concerning the applicability of the models was negative, as shown by Minsky and Papert (1969). Their judgment was

[4]Precisely speaking, this kind of neural network models can be categorized as models with supervised learning. Models with unsupervised learning, on the other hand, uses samples of inputs only for identifying the groups which out- of-samples belong to.

that only a narrow class of functions can be approximated by simple artificial neural network models, and furthermore, they conjectured that the expansion to complicated neural network models would also be sterile. Recently, it was shown by several authors that the judgment and the intuition by Minsky and Paper (1969) was incorrect. In fact, a simple single layer feedforward neural network model has been shown to have the ability to approximate a wide class of functions. Detailed description of the theorems which contain the proof for the universal approximating property of single hidden layer feedforward neural networks is out of scope in this paper, but more mathematically interested readers can see Gallant and White (1988) and Hornik, Stinchcombe and White (1989).

Due to the universal approximation properties of artificial neural network models, the models started to re-emerge, and to be successfully applied in many different areas of science. Different kinds of artificial neural network models have been developed. Classifying the models into two categories — depending on whether nodes in the model are memoryless or not — one is a group of static models (memoryless nodes) and the other is a group of dynamic models (nodes with memory). In other words, in static models, outputs are determined by current inputs only, whereas in dynamic models, outputs are determined by current and previous inputs. For example, each node may have a differential equation in dynamic models.

The model that will be used in this paper is a simple single hidden layer feedforward neural network model (static model). Single hidden layer means that models are composed of three layers: input layer, output layer, and hidden layer in between. Feedforward means that the interaction between a node in one layer and a node in another layer is in one direction. In other words, there is not a feedback, but a feedforward interaction between any two nodes in different layers (see Figure 3.2).

3.1 Model

To facilitate the explanation of the model, let's define some notations as follows.

- x_{jp}: jth element of input vector in the pth training sample where $j = 1, \ldots, I_n$ $p = 1, \ldots, P$.

- X_p: Input vector in the pth training sample

- γ_{ij}: Parameter for jth node in the first layer and ith node in the second layer where $j = 1, \ldots, I_n$ and $i = 1, \ldots, H_n$ ($\widehat{\gamma_{ij}}$ is the estimated value for γ_{ij})

- β_i: Parameter for ith node in the second layer where $i = 1, \ldots, H_n$. ($\widehat{\beta_i}$ is the estimated value for β_i.)

- θ: Parameter vector. ($\widehat{\theta}$ is the estimated value for θ.)

- $\widehat{y_p}$: estimated value of output in the pth training sample ($\equiv f(X_p, \widehat{\theta})$)

- y_p: actual output value in the pth training sample

- $\Phi(z)$: logistic function (squasher) — $\Phi(z) = \frac{1}{1+\exp -z}$

The model for predicting time series event in this paper is as follows:

$$\widehat{y_p} = \widehat{\beta_0} + \sum_{i=1}^{H_n} \Phi(\sum_{j=1}^{I_n} \widehat{\gamma_{ij}} x_{jp}) \widehat{\beta_i} \equiv f(X, \theta). \tag{2}$$

There are three issues concerning the architecture and the estimation of the model. The first issue is how to determine the number of inputs and the number of hidden units in the hidden layer (I_n and H_n). This issue is often called as the issue of 'generalization of the model'. In theory, standard multilayer feedforward neural network models can approximate virtually any class of functions to any degree of accuracy, provided that sufficiently large number of hidden units are available (see Hornik, Stinchcombe and White (1989)). In practice, however, it is difficult to determine how many hidden units are the optimum given the dimension of input spaces. More important, the risk of overfitting increases as the number of hidden units increase. In other words, neural network models are expected to show improved in-sample performance as the number of hidden units increase. However, the improved in-sample performance does not necessarily imply the improved out-of-sample performance. On the contrary, it is often observed that the out-of-sample performance of neural network models is worsen after a certain number of hidden units as the number of hidden units increases.

The most intuitive method for model selection would be a method of cross-validation. For each observation in the training sample, mean squared prediction error is calculated (i.e. MSE_i), and then all MSEs are averaged out across all the observations ($\sum_{i=1}^{N} \text{MSE}_i/N$). Then the model with the minimum sum of MSEs is selected. One drawback of the cross-validation method is that it requires a huge amount of computations. One version of this method which is less computationally demanding is a method of v-fold cross-validation. Instead of computing MSE for all observations, v-fold cross-validation method computes several MSEs for each group of observations, where there are v randomly selected disjoint groups in the sample. Unfortunately, the v-fold cross-validation method is still computationally much demanding. To reduce the cost of a model architecture, parametric estimates of Akaike's final prediction error (FPE) for each model will be compared, and the model with the minimum FPE will be selected. Akaike's FPE is defined as

$$\text{FPE}(\theta) = \text{MSE}(\theta)\left(\frac{1+\frac{S(\theta)}{N}}{1-\frac{S(\theta)}{N}}\right) \approx \text{MSE}(\theta)\left(1+2\frac{S(\theta)}{N}\right), \qquad (3)$$

where θ is a set of parameters in the model (model-specific), $S(\theta)$ is the number of parameters in the model, and N is the total number of observations in the training sample.[5]

The second issue concerns the overall usefulness of the model. In the standard linear regression model, the statistical significance of the model is often determined by nR^2 which is asymptotically χ^2 distributed. In the neural network model, it is difficult to measure the statistical significance of the model by nR^2 since it is not χ^2 distributed anymore because certain parameters are not identified under the null

[5]Moody (1991) showed that FPE could be a valid estimate of the prediction risk for neural network models under several assumptions.

hypothesis (see White (1988)). For example, in the model in this paper, if the null hypothesis is that $\beta_i = 0$, then $\gamma_{ij}, j = 1, \ldots, I_n$ are not identified. Instead of assessing the statistical significance of the model, in this paper, economic usefulness of the model will be evaluated by stochastic dominance criteria. For detailed description of stochastic dominance criteria and the test method, see Shin (1992).

The last issue concerns the method for estimating the parameters. There are basically two kinds of methods: back-propagation method and non-linear least square (NLS) method. These are described in the next sub-section.

3.2 Methods for Estimation

The most frequently used is the method of 'back-propagation' which was developed from the method of 'stochastic approximation' by Robins and Monro (1951). Starting from arbitrary *random* parameter values,[6] each parameter value is updated according to the gradient and the estimation error of the model until some termination conditions are satisfied.[7] Let's denote $\widehat{f}_{pk}$ as $f(X_p, \widehat{\theta}(k))$ and $\bigtriangledown$ as the gradient operator (with respect to parameters). k denotes the kth iteration. Then the back-propagation method works as follows.

$$\widehat{\theta}(k+1) = \widehat{\theta}(k) + \eta_k \bigtriangledown \widehat{f}_{pk}(y_p - \widehat{f}_{pk}), p = 1, \ldots, P \tag{4}$$

As shown in the above equation, the estimation error of a final output is fed into the procedure for updating parameters. This is why the method is called 'back-propagation' method. In the finite sample, passing through the sample once is not usually enough to get convergence results. Multiple pass-throughs are often required. That is why k and p are distinguished in the above equation. For the initial pass-through, k and p are the same.

[6]If the starting values are the same for all parameters, then parameter values for each input in different hidden units become the same ($\gamma_{lj} = \gamma_{mj}$, for all l, m).

[7]One stopping criteria, for example, is to stop the updating when the magnitude of the gradient is sufficiently small.

The back-propagation method for the model (single hidden layer neural network) in this paper can be more specifically described as follows.

$$\widehat{\gamma_{ij}}(k+1) = \widehat{\gamma_{ij}}(k) - \eta_k \frac{d(\widehat{y_p} - y_p)^2}{d\gamma ij} = \widehat{\gamma_{ij}}(k) - \eta_k(\widehat{y_p} - y_p)\widehat{\beta_i}\widehat{\Phi_i(.)}(1 - \widehat{\Phi_i(.)})\widehat{x_j} \quad (5)$$

$$\widehat{\beta_i}(k+1) = \widehat{\beta_i}(k) - \eta_k(\widehat{y_p} - y_p)\widehat{\Phi_i(.)},\ i = 1, 2, \ldots \quad (6)$$

$$\widehat{\beta_0}(k+1) = \widehat{\beta_0}(k) - \eta_k(\widehat{y_p} - y_p). \quad (7)$$

White (1987a,b) provided several conditions for the learning rate to have the method converge to the local minimum. They are i) $\sum_{k=1}^{\infty} \eta_k = \infty$, ii) $\lim_{k \approx \infty} \sup(\eta_k^{-1} - \eta_{k-1}^{-1}) < \infty$, and iii) $\sum_{k=1}^{\infty} \eta_k^d < \infty$ for some $d > 1$. One necessary condition for the learning rate to satisfy the above conditions is that the learning rate should be a decreasing function of k. If the learning rate is set constant, the method fails to converge for the following reason. The random component of each observation causes the approximation error by which coefficients are updated, and if the learning rate is constant, the impact of the approximation error becomes transient. The leading case for the learning rate is to set η_k to be proportional to k^{-1}. This satisfies the conditions for convergence. In this paper, the learning rate will be set as proportional to k^{-1}.

Among NLS methods, the most representative method is the Gauss-Newton method. Other methods such as Newton-Rhapson method, steepest gradient method, or Marquardt method are variant forms of the Gauss-Newton method which utilize the gradient of a function in searching for an optimum. Therefore, only the Gauss-Newton method will be briefly described.

Consider the following non-linear model.

$$Y = F(\beta_0, \beta_1, \ldots, \beta_k, x_1, x_1, \ldots, x_n) + \epsilon \quad (8)$$

Let θ and g denote a vector of parameters and a gradient of F with respect to θ respectively. The Gauss-Newton method is based on the Taylor expansion of $F(\theta_{k+1})$ around the previous parameters θ_k which is

$$F(\theta_{k+1}) = F(\theta_k) + g'(\theta_{k+1} - \theta_k) + \cdots. \quad (9)$$

Now consider the normal equation, which is

$$gY = gF(\theta_{k+1}) \tag{10}$$

By ignoring all terms except the first two in the Taylor series and by substituting the Taylor series into $F(\theta)$,

$$gY = gF(\theta_k) + gg'(\theta_{k+1} - \theta_k). \tag{11}$$

Then the following updating rule is derived.

$$\theta_{k+1} = \theta_k + (gg')^{-1} g\,(Y - F(\theta_k))\,. \tag{12}$$

White (1987a,b) showed that both non-linear least square (NLS) and back-propagation with $\eta_k \propto k^{-1}$ converge stochastically to the same limit. These two methods share the common methodological aspect that both of them use the local gradient descent in finding the (local) minimum (of objective function). Nonetheless, there is one critical difference between these two methods, which is that the NLS method (i.e. Newton-Rhapson method) updates the coefficients with information from all samples at a time, whereas the back-propagation method updates the coefficients with information from one observation at a time (see Figure 3.3). Therefore, the NLS method is more efficient but more computationally demanding, and the back-propagation method is computationally less demanding but less efficient than the NLS method. The back-propagation method tends to have slower convergence rate than the NLS method. Detailed comparisons between these two methods will be illustrated in section 4 via Monte-Carlo experiments. With the exchange rate data, it is found that the method of back propagation is found to be difficult to get convergence results in the finite sample.[8] Therefore, NLS will be used for estimating the coefficients.

[8]When the method of back propagation is applied to the exchange returns, convergence is not achieved with 1,000 multiple pass-throughs of 1,300 observations .

3.3 Empirical Assessment of the Models

Before explaining about the empirical results, one practically subtle problem in estimating the model has be mentioned. Unlike linear regression models, estimating procedure for neural network models with logistic squashing functions is sensitive to the scaling of inputs. In other words, the value of $f(x) = \frac{1}{1+\exp(-x)}$ is sensitive around $x = 0$ and insensitive if x is very large or very small. In theory, this is not a problem since the learning rate and the parameter values can be adjusted so that proper rescaling is achieved for all the inputs and outputs. However, this often cause a trouble in practice. For example, $\exp(-x)$ is usually defined only up to $x > -70$ for most computers. If $x < -70$, the exponential function blows out, or $f(x) = \frac{1}{1+\exp(-x)}$ has to be set as 0. Then the trouble begins—parameters are never updated. Therefore, some kind of normalization of inputs and outputs is often necessary in implementing neural network models. In this paper, all inputs and outputs in the training sample are normalized by dividing the data by the standard deviation after subtracting the mean. These mean and standard deviation are used in normalizing the out-of-sample data.

Two kinds of models are implemented depending on a set of inputs. The first one (Model 1) has 5 inputs which are 5 lagged returns, and the second one (Model 2) has 5 lagged returns and 4 moving averages, as well as one-period lagged equally weighted average of returns from all exchange rates. For daily returns, both models are considered with the training sample of the first 1,300 observations. For weekly returns, only Model 1 is considered with the training sample of the first 50C weekly returns.

In determining the number of hidden units for each model and for each exchange rate, the method of Akaike's FPE is used with a number of hidden units from two to eight. Table 3.2 shows the results from the model selection (or number of hidden units selection) procedures with FPE criteria. [9]

[9] Although it is not shown in the table, the problem of local minima was serious in estimating the

Table 3.3 shows the return characteristics of non-training samples when predictions are made by neural network models. 'Buy' returns are the returns when the predicted returns are positive, and 'sell' returns are the returns which result when the predicted returns are negative. It is clear that 'buy' returns have higher means than 'sell' returns with comparable standard deviations and skewnesses (except for Model 1 for daily returns from Swiss Franc). Therefore, it seems that neural network models are useful in predicting returns.

Table 3.4 compares predicted returns with actual returns. One interesting feature of predicted returns is that they have much smaller standard deviations than actual returns for all models, and also for all exchange rates. Another important observation is that the correlation coefficients between the predicted returns and the actual returns are all positive, although the magnitudes are usually small. This also implies that neural network models are somehow successful in predicting returns, especially weekly returns.

4 Back-Propagation Method vs. NLS

In this section, the back-propagation method and the non-linear least square (NLS) method will be compared via Monte-Carlo experiments. As briefly mentioned in the previous section, both methods utilize the local gradient descent to find the (local) minimum. The back-propagation method pass through each observation in the sample one by one, and the NLS uses information from all observations in the sample for each updating incidents. Therefore, for each updating incident, the back-propagation method is usually less computationally demanding, but less efficient the NLS in the sense that the NLS utilizes more information from the data. However, this does not necessarily mean that the total computing time for the back-propagation method is

model. One way to partially avoid the problem is to try using different sets of initial parameters. Unfortunately, this requires a heavy computation. For example, if each model is estimated with 5 different sets of initial parameters, the total number of models to be estimated would be $5 \times 7 \times 6 \times 3 = 630$. Therefore, it was assumed that the attained minimum was global.

shorter than the NLS to achieve convergence. It may well be the case where a large number of repeated passes through a sample is necessary for the method of back-propagation to satisfy any convergence criterion whereas the NLS method achieves the criterion quickly.

Asymptotically, both methods are shown to have the same limit by White (1987a,b) as long as the learning rate in the back-propagation method is set properly ($\eta_k \propto k^{-1}$). However, it is not certain how fast both methods can achieve the convergence criterion in the finite sample. Simple Monte-Carlo experiments will compare the finite sample properties of the two methods.

4.1 Data Generating Process

Data are assumed to have two components: one is a random component and the other is a deterministic component. The deterministic component is assumed to be chaotic. Before describing the actual data generating process, a brief explanation about chaotic processes will precede since understanding chaotic processes is essential in understanding the data generating process which will be used for the Monte-Carlo comparison of the two estimating methods (NLS and back propagation).

A system i) which has deterministic rules in governing its behavior, and ii) which has highly erratic (or seemingly random) resulting outcomes is called a chaotic process. Among many chaotic processes, the simplest one is the Tent map which is expressed as

$$0 < x_0 < 1,$$

$$x_t = 2x_{t-1} \text{ if } 0 < x_{t-1} < 0.5,$$

$$x_t = 2(1 - x_{t-1}) \text{ if } 0.5 \leq x_{t-1} < 1.0. \tag{13}$$

As shown in the above equations, the sequence of x_t is completely determined by the initial value of x_0. However, it has the following three important properties:

1. x_t's are uniformly distributed in the unit interval [0,1] as t grows without bounds.

2. Misspecifying the initial value x_0 has an fatal effect in identifying the whole sequence of x_t.

3. kth order autocorrelation coefficients are asymptotically zero for all k.[10]

Most chaotic processes share the above three properties although their asymptotic distributions are different. Another good examples of chaotic processes are 'Pseudo Random Number Generators' which are frequently used for generating random numbers by computers. The whole sequence of numbers are completely determined by the initial value or 'seed' which is given by the users.[11]

There are many other complex chaotic processes. Among them are Logistic map, Henon Map, Lorenz map, and Mackey-Glass Equation (see Hsieh (1991). Generally, all these chaotic processes are expressed as

$$x_t = f(x_{t-1}, x_{t-2}, \cdots). \tag{14}$$

Detailed explanation of the each process is out of scope in this paper. Only the Henon map will be explained since it is the process which is chosen for the Monte-Carlo comparison of the estimating methods (NLS and back propagation).

The Henon map is a bivariate chaotic system, which is described by a second order difference equation as

$$x_t = 1 - 1.4x_{t-1}^2 + 0.3x_{t-2}. \tag{15}$$

Although the whole sequence is determined by randomly selected initial values of x_{-1} and x_0, the resulting series are highly erratic (see White (1989)).

Now let's describe the actual data generating process. If Z_t is denoted as a variable which is generated, Z_t is assumed to have the following system.

$$Z_t = X_t + Y_t, \text{ where}$$

[10]This implies the failure of the variance ratio tests in detecting the non-randomness of the process.
[11]This is why it is called 'Pseudo' not 'True'.

$$X_t = x_t \times (k-1) * 0.5, \ x_t \sim N(0,1), \ k = 1, 2, \cdots, 6$$

$$Y_t = 1 - 1.4Y_{t-1}^2 + 0.3Y_{t-2}, \quad Y_{-1} = Y_0 = 0.5. \tag{16}$$

In the above equations, $k = 0$ means that Z_t is completely deterministic, and Z_t becomes more random as k increases.

4.2 Monte-Carlo Comparisons

Several sets of 2,000 data points are generated by the process which was described in the previous sub-section. For each set of data, both the NLS method and the back propagation method with different initial parameter values are applied to estimate the single hidden layer neural network model with 5 hidden units and the two inputs which are Z_{t-1} and Z_{t-2} (inputs are assumed to be correctly specified). Denoting SSE_i as sum of squared errors in the ith iteration, the convergence criterion is as follows for both the NLS and the back propagation method.

$$\frac{\text{SSE}_{i-1} - \text{SSE}_i}{\text{SSE}_i + 10^{-6}} < 10^{-6} \tag{17}$$

For the NLS method, the maximum number of iterations is set as 500, and for the back propagation method, the maximum number of pass-throughs is set as 1,000.

The sample mean and variance of the chaotic component (X_t) are calculated as approximately 0.2612 and 0.5161. Since the two components (X_t and Y_t) in the random variable (Z_t) are assumed independent, the ratio of $\text{Var}(X_t)/(\text{Var}(X_t)+(n-1)^2 \times 0.25)$ can be interpreted as a measure for the ratio of deterministic component in the random variable (Z_t) when $k = n$. For $k = 1, 2, 3, 4, 5, 6$, these ratios are as follows:

1.0, 0.67, 0.34, 0.19, 0.11, 0.08.

Table 3.5 illustrates the comparisons between the NLS method and the back-propagation method. This table shows the best 4 results out of 10 trials for each k. The numbers in the parenthesis denote the number of iterations (or pass-throughs) made for getting the results. First, it should be noted that the convergence criterion is not always satisfied with 1,000 number of pass- throughs for the back propagation method. When $k = 1$, or equivalently when the process is completely deterministic, the NLS method has much higher R^2 than the back propagation method. For all other cases, the NLS method also has slightly higher R^2s than the back propagation method.[12] Except for a few cases, the NLS method is quick in satisfying the convergence criterion. Therefore, the NLS method seems to be more efficient than the back propagation method in finite samples.

For both methods, R^2s are decreasing dramatically as the ratio of deterministic components in the process decrease. For example, the best R^2 from the back propagation method drops from 0.29 to 0.10 as the ratio of deterministic components decrease from 1.0 ($k = 1$) to 0.67 ($k = 2$). The best R^2 from the NLS method drops more dramatically from 0.89 to 0.15. The presence of random components seems to have a great impact on the effectiveness of neural network models. One drawback of the NLS method is that it seems to be more sensitive to starting parameter values than the back propagation method. Therefore, it is important to have enough trials when the NLS method is applied to neural network models.

5 Moving Average Trading Rules

Among several commonly used technical trading rules, one of the simplest forms of moving average trading rules will be discussed in this section.

5.1 Moving Average Rules

Let's denote P_t an exchange rate at time t. If the interest payment for holding a foreign currency is ignored, then the one period return from the foreign currency is

[12]The R^2s are not significantly improved with 5,000 iterations for the back propagation method. Moreover, the convergence criterion is still not satisfied with 5,000 pass-throughs.

defined as

$$R_t = \log P_t - \log P_{t-1}. \tag{18}$$

Also, let's denote $M_t(q)$ a q period moving average of an exchange rate which is defined as

$$M_t(q) = \frac{1}{q}\sum_{i=0}^{q-1} P_{t-i}. \tag{19}$$

Then the trading rule is according to the sign of $P_t - M_t(q)$. In other words, if $P_t - M_t(q)$ is positive, buy the foreign currency. On the other hand, if $P_t - M_t(q)$ is negative, sell the foreign currency.

5.2 Empirical Assessment

Moving averages of four different periods are considered for daily and weekly trading. For daily trading, moving averages of 50, 100, 200, and 300 days are considered. For weekly trading, moving averages of 20, 30, 50, and 100 weeks are considered. In calculating moving averages for weekly trading, exchange rates on every Wednesday are regarded as the exchange rates of the week (exchange rate movement within a week is ignored).

Table 3.6a and 3.6b show the return characteristics of 'buy' returns and 'sell' returns according the moving average trading rules. As shown by LeBaron (1992), the moving average trading rules seem to be effective in the sense that the means of 'Buy' returns are usually higher than those of 'Sell' returns with the standard deviations and skewnesses comparable for both daily and weekly returns, and for all moving averages of different periods.

6 Evaluation by Stochastic Dominance Criteria

In this section, the economic significance of the artificial neural network models and the moving average trading rules will be evaluated with stochastic dominance criteria.

For more details about the stochastic dominance criteria (and about Kolmogorov Smirnov test), please see Shin (1992).

Table 3.7 shows the stochastic dominance relationships between 'buy' returns and 'sell' returns from neural network models and from moving average trading rules. For the neural network models, either 'buy' returns first-degree stochastically dominate 'sell' returns or they are empirically indistinguishable. For moving average trading rules, 'buy' returns usually first-degree stochastically dominate 'sell' returns. More stochastic dominance relationships are attained for the moving average trading rules than the neural network models. Overall, it seems that both neural network models and moving average trading rules are economically useful in predicting returns if market frictions (i.e. transaction cost) are negligible.[13] The relative performance of the neural network models and the moving average trading rules are described in table 3.8. For 'buy' returns, in many cases, returns from moving average trading rules first- or second-degree stochastically dominate returns from neural networks. There are also many cases where 'buy' returns from the moving average trading rules, and from the neural network models, do not stochastically dominate each other, or they're empirically indistinguishable. There is only one case in which 'buy' returns stochastically dominate 'buy' returns from moving average trading rules.

The opposite relationships are usually attained for 'sell' returns from the moving average trading rules and from the neural network models. In many cases, 'sell' returns from the neural network models first- or second-degree stochastically dominate 'sell' returns from the moving average trading rules. There are also many cases where they do not stochastically dominate each other, or where they are empirically indistinguishable. In no case do 'sell' returns from the moving average trading rules stochastically dominate 'sell' returns from the neural network models.

Overall, table 3.8 implies that the moving average trading rules seem to perform better than the neural network models in predicting returns. Unfortunately, a decisive

[13]Considering that foreign exchange markets are highly liquid, market frictions in foreign exchange markets seem insignificant.

statement about the relative usefulness of neural network models in comparison with moving average trading rules is hard to make from these results for several reasons. First, there is an implicit selection bias involved in moving average trading rules. In other words, moving average trading rules are already proven to be useful in predicting returns, which is why they are widely used for technical analysis.[14] If hundreds of different types of neural network models were tried, it would be possible for some neural network models to perform better than most moving average trading rules. Second, as mentioned earlier, in estimating neural network models, the problem of local minima was not resolved because of heavy computational requirements. If extensive trials were made to achieve global minima in estimating the models, then the models considered in this paper could possibly perform better than the moving average trading rules.

The bottom line about neural network models and the moving average trading rules is that predictions made by neural network models and by the moving average trading rules are economically significant if market frictions are not serious.

7 Summary and Conclusions

In this paper, the predictability of returns for foreign exchange rates was investigated. The heteroskedasticity robust variance ratio test showed that the null hypothesis of random walks for the exchange rates could not usually be rejected at a 5% significance level. This implies that linear auto regressive models are not effective in predicting returns since the variance ratio statistics can be interpreted as a linear combination of autocorrelation coefficients. However, the inability to reject the random walk hypothesis does not exclude the usefulness of highly non-linear models, since the variance ratio test exploits the characteristics of sample returns only up to the second moments.

[14]There must have been thousands of trading rules which were tested if they were useful in predicting returns by technical analysts.

Two kinds of non-linear predictions were considered: one by neural network models and the other by simple moving average trading rules. The implementation of the neural network models was successful in the sense that i) 'buy' returns have higher means than 'sell' returns with comparable standard deviations and skewnesses, ii) predicted returns from the neural network models are positively correlated with the actual returns, and iii) 'buy' returns usually first-degree stochastically dominate 'sell' returns. The moving average trading rules were also found to work well. 'Buy' and 'sell' returns using the moving average trading rules have characteristics similar to those using the neural network models.

Comparing the relative performance of the neural network models and the moving average trading rules, the moving average trading rules seemed to perform better than the neural network models in terms of stochastic dominance relationships. However, this does not imply that neural network models are not useful in comparison with moving average trading rules for the following two reasons: 1) the selection bias is implicitly involved in moving average trading rules because it is already proven to be very effective out of thousands of possible trading rules, and 2) the problem of local minima is still present in the estimated models because only one set of initial parameters was used in estimating the models to avoid heavy computation requirements.

Two estimating methods — NLS and back propagation — were compared by the Monte-Carlo simulations. The NLS method usually had quicker convergence results than the back propagation method. In terms of R^2, the NLS method also had a better performance. The impact of random components in the process on the effectiveness of neural network models seemed to be significant in the sense that R^2s decreases by a large degree. Although the NLS method and the back propagation method with a certain learning rate have the same limit, the NLS method seemed to be more efficient than the back propagation method in finite samples.

Overall, the results from the neural network models can be regarded as promising. In spite of the problem of local minima, the simple models were successful in predicting returns. More extensive investigations into various kinds of neural network models would increase the predictability of returns for foreign exchange rates. Finally, since they are able to capture the highly non-linear patterns of asset returns in a very systematic way, it is expected that artificial neural networks will become indispensable for most technical analyses in the near future.

References

[1] Brock, W., J. Lakonishok, and B. LeBaron (1992), "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," Journal of Finance 47, 1731-1764.

[2] Brown, D.P. and R.H. Jennings (1989), "On Technical Analysis," Review of Financial Studies 2, 527-551.

[3] Debondt, W. and R. Thaler (1985), "Does the Stock Market Overreact?," Journal of Finance 40, 793-805.

[4] Debondt, W. and R. Thaler (1987), "Further Evidence on Investor Overreaction and Market Seasonality," Journal of Finance 42, 557-582.

[5] Gallant, A.R. and H. White (1988), "There Exists a Neural Network That Does Not Make Avoidable Mistables," Proceedings of the *Second Annual IEEE Conference of Neural Networks*, IEEE Press, New York, NY.

[6] Hornik, K., M. Stinchcombe and H. White (1989), "Multilayer Feedforward Networks Are Universal Approximators," Neural Networks, 2, 359-366.

[7] Hsieh, David A. (1991), "Chaos and Nonlinear Dynamics: Application to Financial Markets," Journal of Finance 46, 1839-1877.

[8] Hush, D.R. and B.G. Horne (1993), "Progress in Supervised Neural Networks," IEEE Signal Processing Magazine, 10, 8-39.

[9] LeBaron, B. (1992), "Do Moving Average Trading Rule Results Imply Nonlinearities in Foreign Exchange Markets?," Working Paper, University of Wisconsin - Madison.

[10] Leroy, S.F. (1973), "Risk Aversion and the Martingale Property of Stock Returns," International Economic Review 14, 436-446.

[11] Lo, A. and C. MacKinlay (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," Review of Financial Studies 1, 41-66.

[12] Lo, A. and C. MacKinlay (1990), "When Are Contrarian Profits Due To Stock Market Overreaction?," Review of Financial Studies 3, 175-205.

[13] Minsky, M. and S. Papert (1969), *Perceptrons*, MIT Press, Cambridge, MA.

[14] Scheinkman, J. and B. LeBaron (1989), "Nonlinear Dynamics and Stock Returns," Journal of Business 62, 311-337.

[15] Shin, S.H. (1992), "Economic Implications of Asset Return Predictability: Stochastic Dominance Comparisons," Working Paper, M.I.T.

[16] Treynor, J. and R. Ferguson (1985), "In Defense of Technical Analysis," Journal of Finance 40, 757-775.

[17] Utan, J. and J. Moody (1991), "Selecting Neural Network Architectures via the Prediction Risk: Application to Corporate Bond Rating Prediction," Proceedings of the *First International Conference on Artificial Intelligence Applications on Wall Street*, IEEE Computer Society Press, Los Alamitos, CA.

[18] White, H. (1988), "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns," Working Paper, University of California, San Diego.

[19] White, H. (1989), "Some Asymptotic Results for Learning in Single Hidden-Layer Feedforward Network Models," Journal of the American Statistical Association 84, 1003-1013.

Appendix

Let's denote $R_t(q)$ as a q period return at time t and np as a total number of one period returns in the sample. Then consider the following unbiased estimators for variances of one period and q period returns.

$$\bar{\sigma}(1) = \frac{1}{nq-1}\sum_{k=1}^{nq}(R_t(1)-\mu)^2$$

$$\bar{\sigma}(q) = \frac{1}{m}\sum_{k=q}^{nq}(R_t(q)-q\mu)^2$$

$$m = q(nq-q+1)\left(1-\frac{q}{nq}\right)$$

and the statistics

$$\bar{M}_r(q) = \frac{\bar{\sigma}(q)}{\bar{\sigma}(1)} - 1$$

Consider the following null hypothesis.

1. For all t, $E(\epsilon_t) = 0$, and $E(\epsilon_t\epsilon_{t-\tau}) = 0$ for any $\tau \neq 0$.

2. $\{ \epsilon_t \}$ is ϕ-mixing with coefficients $\phi(m)$ of size $r/(2r-1)$ or α-mixing with coefficients $\alpha(m)$ of size $r/(r-1)$, where $r > 1$, such that for all t and for any $\tau \geq 0$, there exists some $\delta > 0$ for which

$$E|\epsilon_t\epsilon_{t-\tau}|^{2(r+\delta)} < \Delta < \infty$$

3. $\lim_{nq\to\infty} \frac{1}{nq}\sum_{t=1}^{nq} E(\epsilon_t^2) = \sigma_0^2 < \infty$

4. For all t, $E(\epsilon_t\epsilon_{t-j}\epsilon_t\epsilon_{t-k}) = 0$ for any nonzero j and k where $j \neq k$.

Theorem (LM) 1 *Denote $\theta(q)$ the asymptotic variance of $\bar{M}_r(q)$. Also denote $\hat{\delta}(j)$ as*

$$\hat{\delta}(j) = \frac{\sum_{k=j+1}^{nq}(R_k(1)-\hat{\mu})^2(R_{k-j}-\hat{\mu})^2}{[\sum_{k=1}^{nq}(R_k(1)-\hat{\mu})^2]^2}.$$

Then

1. *$\bar{M}_r(q)$ converges almost surely to zero for all q as n increases without bound.*

2. *$z^*(q) \equiv \sqrt{nq}\bar{M}_r(q)/\sqrt{\hat{\theta}}$ is distributed as asymptotically standard normal, where*

$$\hat{\theta}(q) \equiv \sum_{j=1}^{q-1} \left[\frac{2(q-j)}{q}\right]^2 \hat{\delta}(j).$$

Table 3.1
Random Characteristics of Returns for Exchange Rates

Sample period is from 1971.1.5 to 1992.7.7. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. V-Ratio (q) is the statistics from the variance ratio test with return horizon of q days (or weeks). * means that the number is statistically different from 1 at 5% significance level.

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
(Daily)						
Mean	-0.000031	0.000017	0.000165	0.000196	0.000216	- 0.000041
Std	0.00231	0.00637	0.00654	0.00601	0.00746	0.00609
V-Ratio (2)	1.06064	1.03516	1.03573	1.04448	1.02727	1.0660*
(Weekly)						
Mean	-0.000152	0.000066	0.000779	0.000933	0.001021	- 0.000205
Std	0.00528	0.01413	0.01425	0.01388	0.01634	0.01374
V-Ratio (2)	1.11390	1.05846	1.08648	1.07886	1.04444	1.05320
V-Ratio (4)	1.14986	1.23043	1.2773*	1.2548*	1.17649	1.13080
V-Ratio (8)	1.10215	1.28143	1.3556*	1.39566	1.25750	1.27429
V-Ratio (16)	1.06037	1.44841	1.52160	1.48534	1.40296	1.43166

Table 3.2
Cross Validation By Final Prediction Errors

Training samples are the first 1,300 observations (or 1,300 daily returns) for models of daily returns and the first 500 observations (or 500 weekly returns) for models of weekly returns. Sample period starts from 1971.1.5. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. Neural network model 1 is the model with 5 inputs (5 lagged returns). Neural network model 2 has 10 inputs (5 lagged returns, previous day's equally weighted returns for all 6 exchange rates, and 4 moving averages (moving averages of 50, 100, 200, 300 days or 20, 30, 50, 100 weeks) of the exchange rates. For weekly returns, only model 1 is considered. Final prediction error is defined as

$$\text{FPE} \approx \text{MSE} \times \left(1 + 2\frac{S(\lambda)}{N}\right),$$

where MSE is the mean squared error of the model, $S(\lambda)$ is the number of parameters in the model, and N is the number of observations. Trials are conducted for each model with number of nodes in the hidden layer from 2 to 8 hidden units and the model with the minimum FPE is selected.

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
Model 1 (Daily) No. of hidden units	6	4	2	6	8	6
Model 2 (Daily) No. of hidden units	4	5	5	3	3	2
Model 1 (Weekly) No. of hidden units	6	2	3	6	6	2

Table 3.3
Actual Return Characteristics From Neural Network Models

Sample period is from 1971.1.5 to 1992.7.7. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. 'Buy' returns are the returns on days (or weeks for weekly returns) when the predicted returns from the neural network models are positive. Similarly, 'Sell' returns are the returns when the predicted returns are negative. Neural network model 1 is the model with 5 inputs (5 lagged returns). Neural network model 2 has 10 inputs (5 lagged returns, previous day's equally weighted returns for all 6 exchange rates, and 4 moving averages (i.e. moving averages of 50, 100, 200, 300 days or 20, 30, 50, 100 weeks) of the exchange rates. Models are estimated from the first 1,300 observations for daily models and 500 (weekly) observations for a weekly model and applied to the rest of the sample period. For weekly returns, only model 1 is considered.

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
[Daily]						
Model 1						
Mean (buy)	0.000055	0.000223	0.000252	0.000529	0.000110	0.000428
Mean (sell)	-0.000123	-0.000404	-0.000126	-0.000154	0.000218	-0.000342
Std (buy)	0.002473	0.006420	0.006575	0.006363	0.007317	0.006346
Std (sell)	0.002481	0.006582	0.006512	0.006005	0.007854	0.006683
Skew (buy)	0.001439	-0.003809	-0.003495	0.004592	0.003742	0.004244
Skew (sell)	-0.002333	-0.003766	0.002756	0.005040	-0.005071	-0.004544
Model 2						
Mean (buy)	0.000078	0.000216	0.000162	0.000457	0.000321	0.000443
Mean (sell)	-0.000157	-0.000336	0.000085	-0.000093	-0.000125	-0.000312
Std (buy)	0.002376	0.006233	0.006619	0.006341	0.007617	0.006158
Std (sell)	0.002573	0.006794	0.006472	0.006026	0.007409	0.006783
Skew (buy)	-0.001723	0.003873	-0.004034	0.005304	0.003759	-0.001963
Skew (sell)	-0.001704	-0.005768	0.003312	0.004066	-0.005330	-0.002974
[Weekly]						
Model 1						
Mean (buy)	0.000036	0.001195	0.001588	0.002168	0.000921	0.001014
Mean (sell)	-0.000110	-0.002291	-0.001654	-0.000453	-0.000807	-0.001472
Std (buy)	0.004767	0.015720	0.015713	0.015077	0.016440	0.016499
Std (sell)	0.006288	0.015698	0.015349	0.014393	0.017885	0.015434
Skew (buy)	-0.003905	-0.010765	-0.008287	0.011685	0.011476	-0.005125
Skew (sell)	-0.004290	-0.004421	0.011637	0.012134	0.007098	0.008104

Table 3.4
Actual vs. Predicted Returns From Neural Network Models

Sample period is from 1971.1.5 to 1992.7.7. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. Neural network model 1 is the model with 5 inputs (5 lagged returns). Neural network model 2 has 10 inputs (5 lagged returns, previous day's equally weighted returns for all 6 exchange rates, and 4 moving averages (i.e. moving averages of 50, 100, 200, 300 days) of the exchange rates. Models are estimated from the first 1,300 observations for daily models and 500 (weekly) observations for a weekly model and applied to the rest of the sample period. For weekly returns, only model 1 is considered. All the statistics for actual returns are from the returns during the post-training sample period. 'Act-M1 (Act-M2)' stands for the correlation between actual and predicted returns from model 1 (model 2).

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
[Daily]						
(Mean)						
Actual	-0.000038	-0.000028	0.000129	0.000208	0.000154	-0.000012
Model 1	-0.000011	0.000131	0.000245	0.000100	0.000554	-0.000251
Model 2	-0.000001	0.000316	0.000363	0.000083	0.000464	-0.000163
(Std)						
Actual	0.002479	0.006493	0.006557	0.006207	0.007543	0.006552
Model 1	0.000390	0.001098	0.001100	0.000816	0.002143	0.000985
Model 2	0.000258	0.001157	0.001502	0.000826	0.001070	0.000525
(Corr)						
Act-M1	0.030280	0.062205	0.036630	0.057087	0.004469	0.064905
Act-M2	0.051146	0.054494	0.024307	0.050650	0.033821	0.070665
[Weekly]						
(Mean)						
Actual	-0.000054	-0.000354	0.000247	0.000944	0.000284	-0.000343
Model 1	-0.000252	0.000516	0.001300	0.001191	0.001759	-0.000585
(Std)						
Actual	0.005758	0.015806	0.015645	0.014820	0.017007	0.015974
Model 1	0.001180	0.003514	0.002383	0.003258	0.003990	0.003147
(Corr)						
Act-M1	0.050777	0.132527	0.124673	0.119544	0.091706	0.074280

Table 3.5
Monte-Carlo: NLS vs. Back-Propagation

For each k, 10 sets of 2,000 data points are generated by the following process:

$$Z_t = X_t + Y_t, \text{ where } X_t \sim N(0,1) \times (k-1) * 0.5, \; k = 1, 2, \cdots, 6$$

$$\text{and } Y_t = 1 - 1.4Y_{t-1}^2 + 0.3Y_{t-2}, \; Y_{-1} = Y_0 = 0.5.$$

$k = 1$ implies that the process is completely chaotic (or deterministic). As k increases, the random component in the process also increases.
The model is a single hidden layer neural network with 5 hidden units and 2 lagged inputs. Although 10 trials with different initial parameters are made in estimating the model, only the best 4 results are reported here. Convergence criterion is that

$$(\text{SSE}_{i-1} - \text{SSE}_i)/(\text{SSE}_i + 10^{-6}) < 10^{-6}.$$

The maximum number of interations for the NLS is set as 1,000, and the maximum number of pass-throughs for the back-propagation is set as 500. The numbers in the parenthesis denote the number of iterations (or pass-throughs) to get the convergence results.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
Best						
R^2-BP	0.2986	0.1095	0.0307	0.0093	0.0056	0.0034
	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)
R^2-NLS	0.8942	0.1511	0.0437	0.0143	0.0178	0.0117
	(500)	(31)	(39)	(30)	(116)	(32)
2nd Best						
R^2-BP	0.2986	0.1049	0.0222	0.0092	0.0048	0.0016
	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)
R^2-NLS	0.8620	0.1503	0.0366	0.0138	0.0115	0.0114
	(481)	(45)	(86)	(17)	(27)	(21)
3rd Best						
R^2-BP	0.2543	0.0997	0.0187	0.0087	0.0047	0.0009
	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)
R^2-NLS	0.7133	0.1435	0.0365	0.0105	0.0114	0.0096
	(500)	(18)	(96)	(37)	(22)	(82)
4th Best						
R^2-BP	0.2150	0.0990	0.0179	0.0066	0.0038	0.0001
	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)	(1,000)
R^2-NLS	0.5022	0.1430	0.0282	0.0102	0.0105	0.0086
	(62)	(23)	(29)	(41)	(18)	(15)

Table 3.6a
Actual Daily Return Characteristics From Moving Average Rules

Sample period is from 1971.1.5 to 1992.7.7. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. 'Buy' returns are the returns on days when the rate is higher than the moving average. Similarly, 'Sell' returns are the returns when the rate is lower than the moving average. 'p=N' stands for the case where N days of moving averages are used for return prediction.

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
p=50 days						
Mean (buy)	0.000034	0.000297	0.000463	0.000529	0.000533	0.000268
Mean (sell)	-0.000092	-0.000275	-0.000214	-0.000200	-0.000192	-0.000343
Std (buy)	0.002121	0.006182	0.006508	0.006079	0.007739	0.005802
Std (sell)	0.002396	0.006331	0.006325	0.005692	0.006796	0.006153
Skew (buy)	-0.001639	0.003661	-0.003365	0.007330	-0.005150	-0.003009
Skew (sell)	-0.001462	-0.004183	0.004321	0.003040	0.004581	-0.002445
p=100 days						
Mean (buy)	0.000006	0.000267	0.000424	0.000506	0.000502	0.000208
Mean (sell)	-0.000063	-0.000268	-0.000244	-0.000221	-0.000225	-0.000288
Std (buy)	0.002136	0.006311	0.006471	0.005914	0.007689	0.005917
Std (sell)	0.002393	0.006256	0.006418	0.005949	0.006854	0.006104
Skew (buy)	-0.001880	0.002359	-0.001978	0.007262	-0.005032	-0.002532
Skew (sell)	-0.001127	-0.003450	0.003725	0.002940	0.004627	-0.003033
p=200 days						
Mean (buy)	0.000004	0.000293	0.000329	0.000393	0.000420	0.000256
Mean (sell)	-0.000066	-0.000296	-0.000152	-0.000133	-0.000174	-0.000345
Std (buy)	0.002127	0.006186	0.006260	0.005857	0.007550	0.005855
Std (sell)	0.002420	0.006513	0.006881	0.006082	0.007191	0.006259
Skew (buy)	-0.001879	0.003170	-0.002068	0.007018	0.004696	-0.003342
Skew (sell)	-0.001200	-0.003879	0.004087	0.002204	0.004521	-0.001624
p=300 days						
Mean (buy)	0.000017	0.000326	0.000337	0.000315	0.000375	0.000204
Mean (sell)	-0.000075	-0.000341	-0.000142	-0.000059	-0.000070	-0.000274
Std (buy)	0.002229	0.006100	0.006313	0.005966	0.007589	0.006012
Std (sell)	0.002355	0.006632	0.006891	0.006033	0.007324	0.006202
Skew (buy)	-0.001680	0.003480	-0.002383	0.006969	-0.004594	-0.004067
Skew (sell)	-0.001555	-0.004569	0.004142	0.002774	0.004143	0.002540

Table 3.6b
Actual Weekly Return Characteristics From Moving Average Rules

Sample period is from 1971.1.5 to 1992.7.7. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Weekly returns are formed from Wednesday to Tuesday daily returns. Returns do not include the interest payment for the foreign currencies. 'Buy' returns are the returns on days when the rate is higher than the moving average. Similarly, 'Sell' returns are the returns when the rate is lower than the moving average.'p=N' stands for the case where N weeks of moving averages are used for return prediction. Each Wednesday exchange rates are considered as those of the week.

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
p=20 weeks						
Mean (buy)	-0.000092	0.001226	0.002059	0.002068	0.002198	0.001120
Mean (sell)	-0.000210	-0.001207	-0.001158	-0.000474	-0.000701	-0.001501
Std (buy)	0.004741	0.013971	0.014173	0.013141	0.016735	0.013361
Std (sell)	0.005778	0.014438	0.014403	0.014882	0.015828	0.014181
Skew (buy)	-0.003971	-0.007738	0.005209	0.010109	0.006155	-0.007762
Skew (sell)	-0.005325	-0.008650	-0.008657	0.016206	0.006970	0.005952
p=30 weeks						
Mean (buy)	-0.000005	0.001324	0.001655	0.001827	0.001999	0.001071
Mean (sell)	-0.000278	-0.001407	-0.000708	-0.000242	-0.000715	-0.001585
Std (buy)	0.004753	0.013816	0.013960	0.012962	0.016509	0.013768
Std (sell)	0.005822	0.014735	0.014989	0.015356	0.016340	0.013931
Skew (buy)	-0.003972	-0.007113	0.003300	0.010084	-0.002749	0.006311
Skew (sell)	-0.005352	-0.008968	-0.007486	0.016489	0.009717	-0.008386
p=50 weeks						
Mean (buy)	-0.000016	0.001534	0.001454	0.001681	0.001700	0.000746
Mean (sell)	-0.000319	-0.001637	-0.000422	-0.000416	-0.000208	-0.001230
Std (buy)	0.005220	0.013663	0.014024	0.013566	0.016761	0.014365
Std (sell)	0.005502	0.014887	0.015176	0.014746	0.016346	0.013630
Skew (buy)	-0.005479	-0.004112	-0.005819	0.014759	0.004771	-0.006062
Skew (sell)	-0.003948	-0.010691	-0.002570	0.010504	0.008888	-0.004981
p=100 weeks						
Mean (buy)	0.000137	0.000713	0.001358	0.001023	0.001643	0.000742
Mean (sell)	-0.000381	-0.000736	-0.000172	0.000599	-0.000120	-0.000990
Std (buy)	0.005625	0.014344	0.014538	0.014206	0.016628	0.014374
Std (sell)	0.005363	0.014977	0.015246	0.014664	0.017593	0.014113
Skew (buy)	-0.005999	-0.009221	-0.010153	0.013415	-0.008555	-0.010755
Skew (sell)	-0.003836	-0.008509	0.010260	0.013681	0.012860	0.009277

Table 3.7
Buy vs. Sell Daily Returns

Sample period is from 1971.1.5 to 1992.7.7. Neural network models are estimated from the first 1,300 daily observations and applied to the rest of the sample period. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Neural network model 1 (N 1) is the model with 5 inputs (5 lagged returns). Neural network model 2 (N 2) has 10 inputs (5 lagged returns, previous day's equally weighted returns for all 6 exchange rates, and 4 moving averages (i.e. moving averages of 50, 100, 200, 300 days) of the exchange rates. 'p=N' stands for the case where N days of moving averages are used for return prediction. 'Buy' returns are the returns when the rate is higher than the moving averages (for MA) or when the predicted returns are positive (for NN Models). Similar interpretation applies to 'Sell' returns. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$||$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

(Moving Average Rules)

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
p=50	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$
p=100	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$
p=150	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$
p=300	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \doteq S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$

(Neural Network Models)

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
N 1	$B \doteq S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \succeq_{FSD} S$	$B \doteq S$	$B \succeq_{FSD} S$
N 2	$B \doteq S$	$B \succeq_{FSD} S$	$B \doteq S$	$B \doteq S$	$B \doteq S$	$B \succeq_{FSD} S$

Table 3.8
Moving Average Rules vs. Neural Network Models

Sample period is from 1971.1.5 to 1992.7.7. Neural network models are estimated from the first 1,300 daily observations and applied to the rest of the sample period. Returns are continuously compounded returns. For days with missing data, zero returns are assumed. Neural network model 1 is the model with 5 inputs (5 lagged returns). Neural network model 2 has 10 inputs (5 lagged returns, previous day's equally weighted returns for all 6 exchange rates, and 4 moving averages (i.e. moving averages of 50, 100, 200, 300 days) of the exchange rates. 'p=N' stands for the case where N days of moving averages are used for return prediction. 'M' stands for returns from moving average predictions and '1(2)' stands for returns from neural network model 1(2). 'Buy' returns are the returns when the rate is higher than the moving averages (for M) or when the predicted returns are positive (for 1, 2). Similar interpretation applies to 'Sell' returns. All decisions are made at the 5 percent significance level.

Notations: '$\succeq_{FSD}$' stands for first- (and second-) degree stochastic dominance, '$\succeq_{SSD}$' stands for second-degree stochastic dominance but *not* for first-degree stochastic dominance, '$\|$' stands for *not* second- (and not first-) degree stochastic dominance, and '$\doteq$'stands for indistinguishable distributions. '$\doteq$' supersedes all the above three relations.

(Buy Returns)

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B.Pound
p=50	$M \doteq 1$	$M \doteq 1$	$M \succeq_{FSD} 1$	$M \succeq_{FSD} 1$	$M \| 1$	$M \succeq_{FSD} 1$
p=100	$M \doteq 1$	$M \doteq 1$	$M \| 1$	$M \succeq_{SSD} 1$	$M \| 1$	$M \succeq_{SSD} 1$
p=200	$M \doteq 1$	$M \doteq 1$	$M \succeq_{SSD} 1$	$M \succeq_{SSD} 1$	$M \| 1$	$M \succeq_{FSD} 1$
p=300	$M \doteq 1$	$M \succeq_{FSD} 1$	$M \succeq_{SSD} 1$	$M \succeq_{SSD} 1$	$M \doteq 1$	$M \doteq 1$
p=50	$M \doteq 2$	$M \doteq 2$	$M \succeq_{FSD} 2$	$M \succeq_{FSD} 2$	$M \| 2$	$M \doteq 2$
p=100	$2 \succeq_{FSD} M$	$M \doteq 2$	$M \succeq_{FSD} 2$	$M \succeq_{SSD} 2$	$M \| 2$	$M \doteq 2$
p=200	$M \doteq 2$	$M \doteq 2$	$M \succeq_{SSD} 2$	$M \succeq_{SSD} 2$	$M \| 2$	$M \doteq 2$
p=300	$2 \doteq M$	$M \succeq_{FSD} 2$	$M \succeq_{FSD} 2$	$M \succeq_{SSD} 2$	$M \| 2$	$M \succeq_{FSD} 2$

(Sell Returns)

	Can.$	Fra.Fr	Ger.DM	Jap.Y	Swi.Fr	B. Pound
p=50	$1 \doteq M$	$1 \succeq_{FSD} M$	$1 \succeq_{FSD} M$	$1 \succeq_{FSD} M$	$1 \| M$	$1 \doteq M$
p=100	$1 \doteq M$	$1 \| M$	$1 \succeq_{FSD} M$	$1 \succeq_{FSD} M$	$1 \| M$	$1 \doteq M$
p=200	$1 \doteq M$	$1 \succeq_{FSD} M$	$1 \succeq_{SSD} M$	$1 \succeq_{SSD} M$	$1 \| M$	$1 \succeq_{FSD} M$
p=300	$1 \doteq M$	$1 \succeq_{FSD} M$	$1 \succeq_{SSD} M$	$1 \doteq M$	$1 \| M$	$1 \| M$
p=50	$2 \doteq M$	$2 \doteq M$	$2 \succeq_{FSD} M$	$2 \succeq_{FSD} M$	$2 \| M$	$2 \succeq_{FSD} M$
p=100	$2 \doteq M$	$2 \doteq M$	$2 \succeq_{FSD} M$	$2 \succeq_{FSD} M$	$2 \| M$	$2 \doteq M$
p=200	$2 \doteq M$	$2 \doteq M$	$2 \succeq_{SSD} M$	$2 \succeq_{SSD} M$	$2 \| M$	$2 \succeq_{FSD} M$
p=300	$2 \doteq M$	$2 \succeq_{FSD} M$	$2 \succeq_{SSD} M$	$2 \doteq M$	$2 \doteq M$	$2 \succeq_{FSD} M$

Figure 3.1
Artificial Neural Network

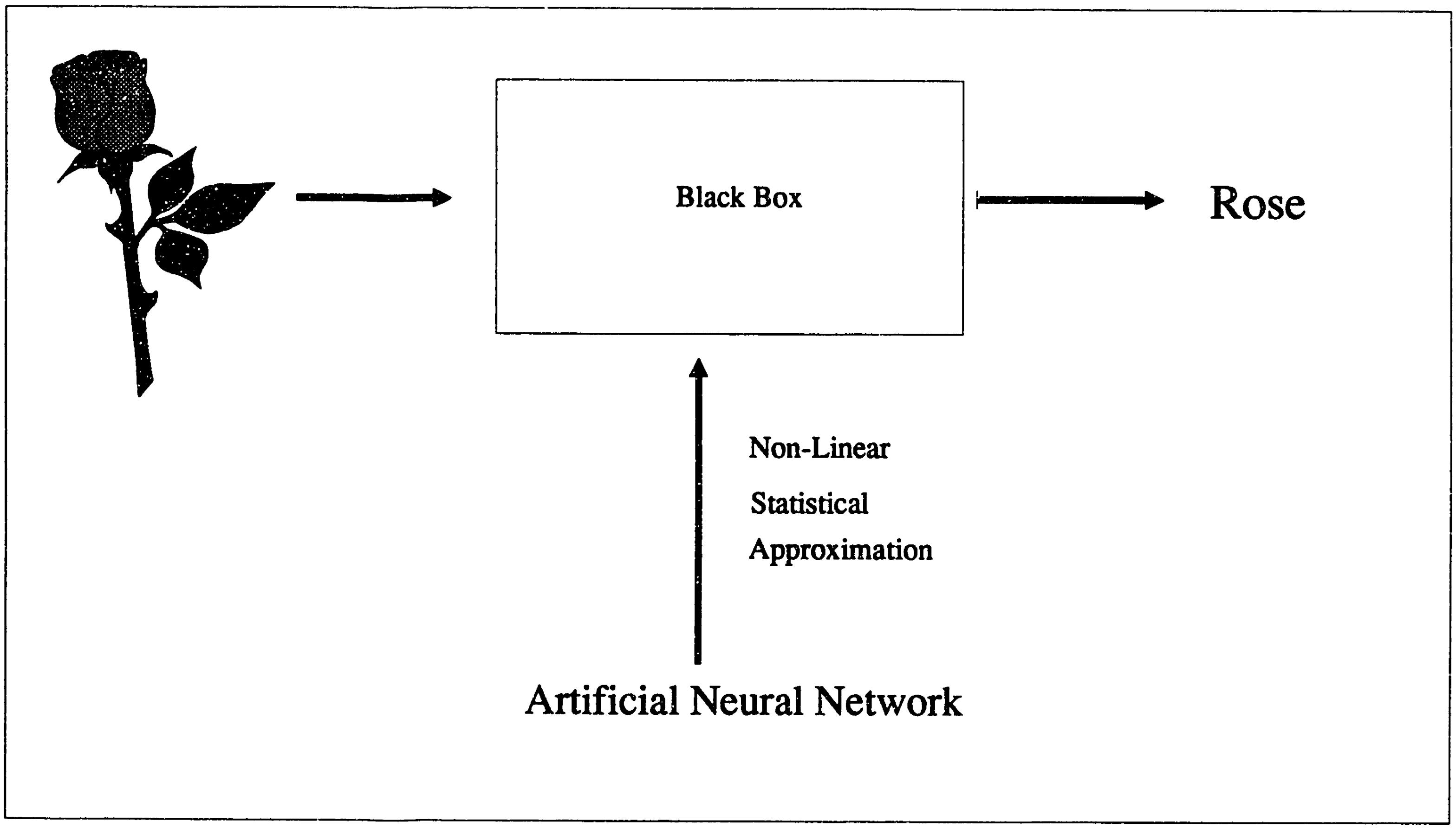
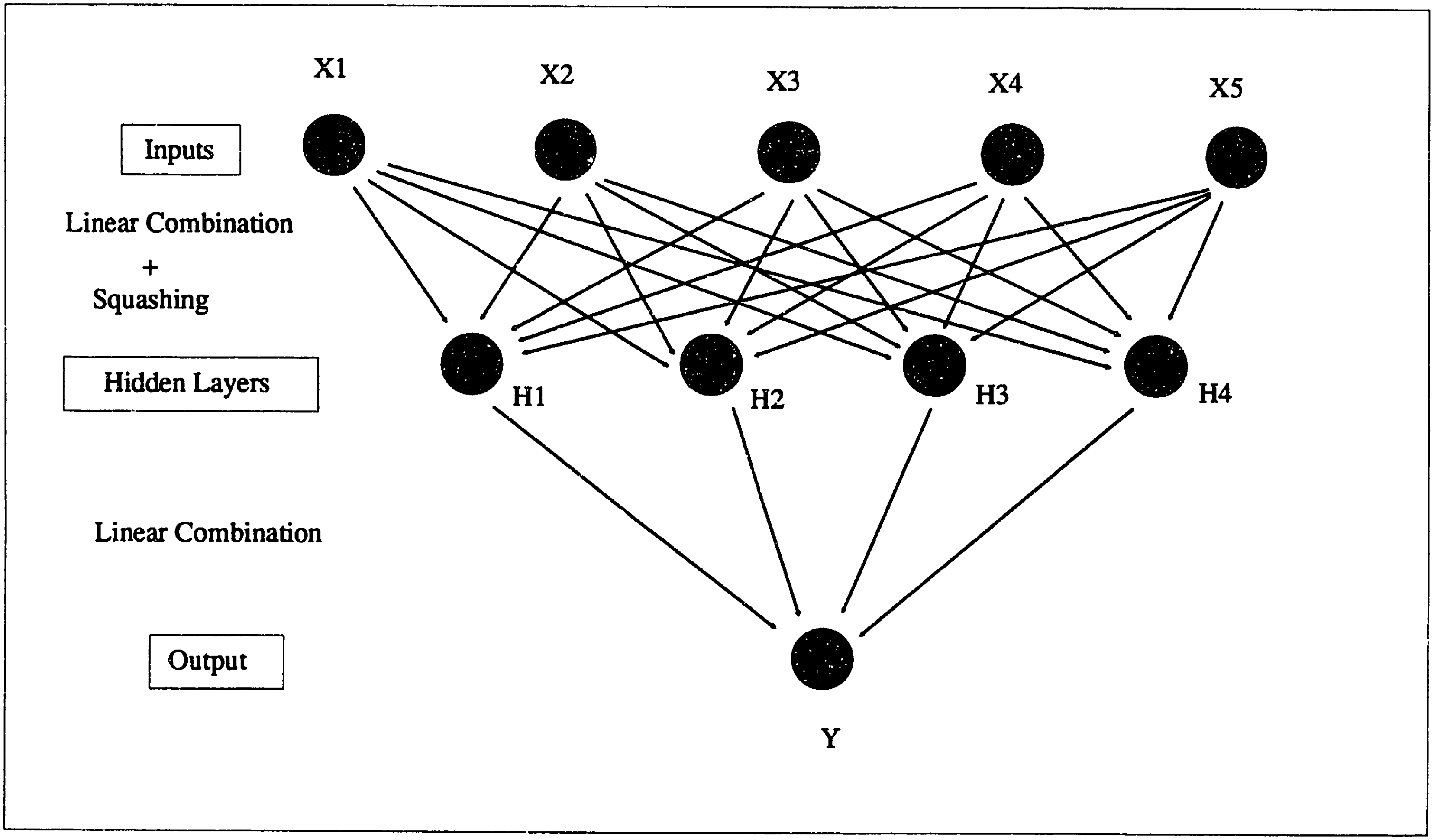


Figure 3.2
Single Hidden Layer Neural Network



See the main text for more details.

Figure 3.3
Back Propagation vs. NLS