ELSEVIER

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

Volume 8

Cheng Few Lee

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

Series Editor: Cheng Few Lee

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO
MANAGEMENT   VOLUME 8

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

EDITED BY

CHENG FEW LEE

Department of Finance, Rutgers University, USA

2001

JAI
An Imprint of Elsevier Science

Amsterdam – London – New York – Oxford – Paris – Shannon – Tokyo

ELSEVIER SCIENCE Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB, UK

First edition 2001

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

British Library Cataloguing in Publication Data
A catalogue record from the British Library has been applied for.

ISBN: 0-7623-0798-6

∞ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

CONTENTS

LIST OF CONTRIBUTORS

Paul Brockman	Department of Accountancy, Hong Kong Polytechnic University, Hong Kong
Robert Brooks	Royal Melbourne Institute of Technology (RMIT), School of Economics and Finance, Australia
Chuang-Chang Chang	Department of Finance, National Central University, Taiwan
Marc Chopin	Department of Economics & Finance, Louisiana Tech University, USA
San-Lin Chung	Department of Finance, National Central University, Taiwan
Ronnie J. Clayton	Jacksonville State University, USA
Robert Faff	Royal Melbourne Institute of Technology (RMIT), School of Economics and Finance, Austrailia
Tan Pooi Fan	Royal Melbourne Institute of Technology (RMIT), School of Economics and Finance, Australia
Asim Ghosh	Department of Finance, Saint Joseph's University, USA
Jason T. Greene	Department of Finance, J. Mack Robinson College of Business, Georgia State University, USA
Shantaram P. Hedge	Finance Department, The University of Connecticut, USA

Charles W. Hodges	Department of Finance, J. Mack Robinson College of Business, Georgia State University, USA
Lisa Kramer	Department of Economics, University of Toronto, Canada
John C. Lee	JP Morgan and Chase Company, USA
Susan M. Managiero	Assistant Professor of Finance, Pace University, USA
Matthew R. Morey	Department of Finance, Pace University, USA
Yiuman Tse	School of Management, Binghamton University (SUNY), USA
H. D. Vinod	Department of Economics, Fordham University, USA
Chi-Keung Woo	Energy and Environmental Economics, Inc., USA
Yexiao Xu	School of Management, University of Texas at Dallas, USA
Maosen Zhong	Department of Business Administration, The University of Texas at Brownsville, USA

EDITORIAL BOARD

This Page Intentionally Left Blank

PREFACE

This research annual publication intends to bring together investment analysis and portfolio theory and their implementation to portfolio management. It seeks theoretical and empirical research manuscripts with high quality in the area of investment and portfolio analysis. The contents will consist of original research on:

(1) the principles of portfolio management of equities and fixed-income securities,
(2) the evaluation of portfolios (or mutual funds) of common stocks, bonds, international assets, and options,
(3) the dynamic process of portfolio management,
(4) strategies of international investments and portfolio management,
(5) the applications of useful and important analytical techniques such as mathematics, econometrics, statistics, and computers in the field of investment and portfolio management.
(6) Theoretical research related to options and futures.

In addition, it also contains articles that present and examine new and important accounting, financial, and economic data for managing and evaluating portfolios of risky assets. Comprehensive research articles that are too long as journal articles are welcome. This volume of annual publication consists of fifteen papers.

Chapter 1. Marc Chopin and Maosen Zhong re-examine the inverse relationship between stock returns and inflation in the post-World War II period. They find that both real activity and monetary fluctuations generate the contemporaneous correlation between stock returns and inflation.

Chapter 2. Chuang-Chang Chang and San-Lin Chung develop a trinomial lattice approach for American-style lookback and barrier options. First, they construct a trinomial model to price lookback options. Secondly, they incorporate the idea of path function. And lastly, they demonstrate how to use the Pelsser and Vorst (1994) approach to compute the hedge ratios for barrier options in our trinomial model.

Chapter 3. Paul Brockman and Yiuman Tse investigate the information role of portfolio depository receipts (PDRs) by using the common factor models of Gonzalo and Granger (1995) and King et al. (KPSW) (1991). Their results are consistent with the view that PDRs are designed for discretionary liquidity traders while futures are better suited for informed traders.

Chapter 4. Matthew R. Morey and H. D.Vinod discuss the bootstrap methodology to suggest a new "Double" Sharpe ratio that allows an investor to make a tradeoff between risk-adjusted performance and estimation risk using the same weighting for estimation risk as the original Sharpe ratio uses for standard deviation.

Chapter 5. Shantaram P. Hedge and Susan Mangiero investigate the difference in the stock market liquidity of a sample of matched firms with high ("institutional favorites") and low ("neglected firms") institutional equity ownership and analyst coverage. They report on their findings that help to better understand the complex relationship between firm visibility and liquidity.

Chapter 6. Asim Ghosh and Ronnie J. Clayton apply the theory of cointergration in case of European stock markets to investigate whether stock prices are predictable. Empirical evidence suggests that the indices are pairwise conintegrated and hence predictable during the period investigated. They estimate the appropriate error correction model and use it to perform out-of sample forecasting.

Chapter 7. Lisa A. Kramer formally derives that a variety of test statistics which have been employed in the finance and accounting literatures for the purpose of conducting hypothesis tests in event studies do not follow their conventionally assumed asymptotic distribution even for large samples of firms.

Chapter 8. Robert Faff, Robert Brooks, and Tan Pooi Fan develop a simple version of a dynamic CAPM by the inclusion of a lagged dependent variable in the market model framework. They use the multivariate approach developed by Gibbons (1982) applied to Australian industry portfolio returns over the period of 1974 to 1995.

Chapter 9. Yexiao Xu investigates the accuracy of Jensen's Alpha and reveals a potential return measurement bias both theoretically and empirically due to the nonlinear geometric compounding of the return data. They also show that this source of bias can cause problems in the beta estimation of mutual funds.

Chapter 10. Jason T. Greene and Charles W. Hodges examine how the imperfect market timing skill and trading frequency affects the return distribution for market timing strategies in open-end mutual funds. They demonstrate that traders need only a modest level of skill in order to beat the market when employing a daily timing strategy. They suggest market timers who trade mutual funds can have a significant impact on a fund's reported performance.

Chapter 11. Chin-Shen Lee proposes a dynamic hedging rule in which hedge ratio is ex-ante updated upon the next period spot and futures prices forecasts. It is demonstrated that our hedging-with-forecasting rule could attain a multi-period perfect hedge status.

Chapter 12. Steven V. Mann and Pradipkumar Ramanlal present a method to measure the duration and convexity of bonds with embedded options that accounts for representative contractual features and realistic movements in the yield curve. The method unifies two aspects of the literature that have evolved almost independently, namely, specification of the interest-rate process and characterization of how the yeild curve changes shape.

Chapter 13. Shih-Kuo Yeh and Bing-Huei Lin investigate a jump-diffusion process, which is a mixture of an O-University process with mean-reverting characteristics used by Vasicek (1977) and a compound Poisson jump process, for the term structure of interest rates. They develop a methodology for estimating both the one-factor and two-factor jump-diffusion term structure of interest rates models and complete an empirical study for Taiwan money market interest rates.

Chapter 14. Chi-Keung Woo, Ira Horowitz, and Khoa Hoang consider the problem of an electric-power marketer offering a fixed-price forward contract to provide electricity purchased from a fledging spot electricity market that is unpredictable and potentially volatile.

Chapter 15. John C. Lee's paper tries to do two things. It first demonstrates the power of Microsoft Excel in that it is possible to create large Decision Trees for the Binomial Pricing Model. The second thing the paper tries to do is present Binomial Option model in a less mathematical matter and make it so that the reader will not have to keep track of many things at one time by using Decision Trees to price call and put options.

STOCK RETURNS, INFLATION AND THE MACROECONOMY: THE LONG- AND SHORT-RUN DYNAMICS

Marc Chopin and Maosen Zhong

ABSTRACT

We re-examine the inverse relationship between stock returns and inflation in the post-World War II period. Fama (1981) theorizes that the inverse inflation-stock return correlation is a proxy for the negative relationship between inflation and real activity. Geske and Roll (1983) argue that the inflation-stock return correlation reflects changes in government expenditures, real economic conditions and monetization of budget deficits. We test these hypotheses simultaneously using a multivariate Vector-Error-Correction Model (VECM) proposed by Johansen and Juselius (1992, 1994). We find that both real activity and monetary fluctuations generate the contemporaneous correlation between stock returns and inflation. However, the Federal Reserve bank seems not to monetize Federal deficits, nor do government deficits appear to drive changes in real economic activity during the period examined. Thus, our results appear to be more compatible with Fama's explanation than that of Geske and Roll.

INTRODUCTION

The observed negative relationship between common stock returns and various measures of expected and unexpected inflation during the post-World War II

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 1–18.

ISBN: 0-7623-0798-6

period is "troublesome" because it appears to contradict Fisher's (1930) prediction of a positive relationship between expected inflation and nominal asset returns, and the received wisdom that common stocks are hedges against inflation. The inflation-stock return correlation has been subjected to extensive study.[1]

Among the most influential of these studies are those of Fama (1981) and Geske and Roll (1983). Based on the notion that money demand is procyclical, Fama (1981) theorizes that the inflation-stock return correlation is essentially a proxy for the negative relationship between inflation and real activity [see also Fama and Gibbons (1982)]. On the other hand, Geske and Roll (1983) emphasize the countercyclical impact of monetization of government deficits and argue that the asset returns-inflation correlations are due to changes in government expenditures in response to changes in real economic conditions and the monetization of budget deficits. In this paper we examine these hypotheses empirically using a multivariate Vector-Error-Correction Model (VECM).

Previous studies of the inflation-stock return relationship focus primarily on short-run dynamics, ignoring possible long-run equilibrium relationships (cointegration) among variables.[2] As Engle and Granger (1987) demonstrate, inferences from VARs are likely to be biased if no allowance is made for long-run cointegrating relationships when they exist. Paying particular attention to the long-run structures, this paper re-examines the stock return-inflation relationship using a multivariate Vector-Error-Correction model (VECM) proposed by Johansen and Juselius (1992, 1994). This framework facilitates examination of both the short- and the long-run Granger-causal ordering of stock returns and inflation in a broad macroeconomic context.

When specifying the VECM, we allow both Fama's (1981) and Geske and Roll's (1983) hypotheses to be tested simultaneously. To summarize the implications of our estimation and testing procedure, we find that Fama's explanation appears to be more consistent with the data in the post-war period than that of Geske and Roll. Further, we provide an alternative interpretation of the "spurious" stock returns-inflation correlation based on the long- and short-run dynamics of the system. In addition to Fama's (1981) proxy effect, we find that as the economy moves toward long-run equilibrium, short-run changes in real activity and the money supply may induce the observed "spurious" correlation.

The rest of the paper is structured as follows. In Section 2, we briefly review the literature. Section 3 formulates the model and provides an outline of the data. Section 4 describes the identification and estimation of the long-run cointegrating relationships. Section 5 describes the short-run dynamics with emphasis on the relationship between stock returns and inflation. The last section provides a brief summary and our conclusions.

LITERATURE OVERVIEW

Contrary to the perception that common stocks provide a hedge against inflation, many studies find a negative relation between stock returns and inflation [e.g. Lintner (1975), and Fama and Schwert (1977)]. In their attempt to explain this "spurious" relation, Fama (1981), and Fama and Gibbons (1982) rely on a combination of money demand theory and the quantity theory of money, contending that an increase (decrease) in real activity is expected to coincide with a decrease (increase) in inflation. Participants in the stock market anticipate the changes in real activity, so that stock prices appear to move inversely with inflation.

In their "reverse causality" explanation, Geske and Roll (1983) argue that a reduction in real activity leads to an increase in fiscal deficits. As the Federal Reserve bank monetizes a portion of fiscal deficits the money supply increases, which in turn boosts inflation. Stock market returns reflect the changes in these macroeconomic variables, resulting in an inverse relationship between stock returns and inflation.

Kaul (1987, 1990) also suggests that, in one way or another, changes in real activity and the money supply underlie the observed negative relationship between inflation and stock returns. To examine the Granger-causal[3] relationships among real activity, inflation, the money supply and stock returns, James, Koreisha and Partch (1985) use a Vector Autoregressive Moving Average model and report evidence of a strong link between stock returns and the growth rate of the money base. Using a Vector Autoregression (VAR) model, Lee (1992) documents that stock returns explain substantial variation in real activity but little variation in inflation. Lee also finds that inflation explains little of the variation in real activity, which responds negatively to shocks in inflation for the post-war period. Thus, Lee suggests that stock prices anticipate changes in real activity. Lee's findings appear to be more compatible with Fama's (1981) hypothesis rather than that of Geske and Roll (1983).

Marshall (1992) argues that the negative inflation-asset returns correlation may be generated by real economic fluctuations, by monetary fluctuations, or possibly changes in both real an monetary variables. Marshall also suggests that the inflation-stock return correlation is more strongly negative when inflation is caused by fluctuations in real economic activity (as predicted by Fama) than when it is caused by monetary fluctuations.

In contrast to the relatively short time periods covered by most of the work examining the stock returns-inflation relationship, Boudoukh and Richardson (1993) employ data sets covering the periods from 1802 through 1990 and 1820–1988 for the U.S. and Britain, respectively. Inconsistent with the proxy hypothesis of Fama, these authors report a positive relationship between inflation and nominal stock returns over long horizons.

In a recent contribution, Hess and Lee (1999) account for the stock returns-inflation relation with two types of shocks to the economy: supply shocks and demand shocks. Supply shocks are characterized as shocks to real activity, while the demand shocks originate largely through shocks to the money supply. Their results lend some support to Marshall's contention (1992) that fluctuations in real economic activity assume more responsibility for the inflation-asset returns correlation than monetary fluctuations.

Although the "proxy" hypothesis of Fama (1981) seems more consistent with the data than the "reverse-causality" hypothesis of Geske and Roll (1983), the evidence is still not conclusive. To date, studies of inflation and stock returns focus on short-run dynamics and ignore possible long-run equilibrium (cointegration) relationships among variables. As discussed above, omitting the long-run relations when they are indeed present may bias the results. Investigation of the long-run equilibria, and deviations from those equilibria, may help explain more thoroughly the inflation-stock returns relationship. This study incorporates the implications of possible long-run relationships in re-examining the two competing hypotheses [Fama (1981) and Geske and Roll (1983)] using the VECM. Below, we formulate the VECM to estimate both the long- and short-run macroeconomic relationships underlying the observed negative correlation between stock returns and inflation in the post-war period.

MODEL FORMULATION

Following Geske and Roll (1983), James, Koreisha and Partch (1985), and Lee (1992), our model includes stock returns, inflation, real activity, interest rates, and the money base. Geske and Roll's reverse causality hypothesis suggests that an adverse shock to real activity, signaled by stock prices, leads to an increase in Federal deficits and the money supply as the Federal Reserve monetizes part of the deficit. Thus, we also include Federal deficits and the money base in the model.

We obtain monthly data for the Standard and Poor's composite stock price index (S), the consumer price index (P), the industrial production index (X), the three-month Treasury bill rate (R), the money base (M) and Federal budget deficits (F) from the DRI/Citibase data tape. The industrial production index proxies for real activity, and the money base proxies for Federal Reserve Bank actions. The sample period covers from January 1968, the beginning of our deficit data on the DRI data CD, to November 1996. With the exception of Federal deficits, we use natural logarithms of all variables.

Prior to constructing the model, we test each data series for unit roots using the Augmented Dickey-Fuller, Phillips and Perron, and Weighted Symmetric tests. All three tests indicate that each of the six variables included in the model is integrated of order one, i.e. each of the variables is stationary in first differences,[4] therefore each variable enters the model in first differences.

When variables are integrated of the same order, there may be a long-run equilibrium relationship among the variables (Engle & Granger, 1987). Under these conditions, the variables may share a common trend in a lower integrating order, say, $\sim I(0)$ and are said to be cointegrated. Granger (1988) reports that if a VAR does not include the error-correction term(s) when the series are indeed cointegrated, then causality may not be detected when it is, in fact, present. Therefore, we test for the presence of cointegration among S, P, X, R, M, and F using the method of Johansen and Juselius (1992).[5]

A vector autoregressive (VAR) representation in a generalized error-correction form

$$\Delta Z_t = C_t + \Gamma_1 \Delta Z_{t1} + \Gamma_2 \Delta Z_{t2} + \ldots + \Gamma_k \Delta Z_{tk} + \Pi Z_{t1} + \varepsilon_t \tag{1}$$

(hence a VECM) may be presented as follows:

where Δ is the first-difference operator, Z_t is a vector of the six $I(1)$ potentially endogenous variables, S, P, X, R, M, and F; Γ_i is a six by six matrix of parameters reflecting the models' short-run structure; Π is defined to be the long-run equilibrium parameter matrix, which can be partitioned into two parts, i.e. $\Pi = \alpha\beta'$, where α is a six by six matrix of the error-correction term coefficients, representing the speed of adjustment from disequilibrium; β is a six by six matrix of long-run coefficients such that the $\beta' Z_{t-1}$ embedded in the above VECM model represents up to $n - 1$ cointegrating relationships, where n is the number of potentially endogenous variables; C_t is a vector of unrestricted constant terms; and ϵ_t stands for a six-dimensional vector of white-noise disturbance terms.

THE UNDERLYING LONG-RUN RELATIONSHIP

In the process of constructing the VECM, we identify the underlying long-run equilibrium relationships, which we will use to interpret the inflation-stock returns correlation. As shown below, the identification of these long-run equilibria entails extensive and careful empirical testing. Since Z_t is a vector of non-stationary $I(1)$ variables, ΠZ_{t-1} must contain $r \leq (n-1)$ stationary long-run error-correction relations for $\epsilon_t \sim I(0)$ to be white-noise. Testing the number of cointegrating relationships amounts to testing the rank of Π, or to finding r stationary vectors in β. Johansen (1988) provides a maximum likelihood approach for estimating β, and a likelihood ratio test with a non-standard distribution for estimating the number of stationary cointegrating vectors. We use both the trace statistic and the maximal eigenvalue to detect the presence of cointegration, combined with Reimers' (1992) adjustment to correct the test statistics for small sample bias.

The number of cointegrating vectors suggested by the Johansen procedure may be sensitive to the number of lags in the VAR model (Banerjee et al., 1993). Thus, we rely on the Akaike Information Criterion (Hall, 1994) and Sims' likelihood ratio test to select the appropriate lag length in the VAR. Testing up to a maximum of 18 lags, the minimum AIC occurs with 11 lags in the VAR model. In addition, Sims LR tests consistently yield insignificant LR statistics with 11 lags versus 12 lags.[6] Therefore, when testing for cointegration and estimating the cointegrating vectors, we choose 11 as the VAR lag length in Johansen procedure.

When specifying the deterministic components (e.g. intercept or trend) in the short- and/or long-run model, Johansen (1992) suggests testing the joint hypotheses of the rank order and the deterministic components based on the so-called "Pantula Principle."[7] The model selected using the Pantula Principle does not include a linear trend, and the intercept is restricted to the cointegration space. Table 1 presents the trace statistics (λ_{trace}) and maximal eigenvalues (λ_{max}) for this model. Both the λ_{trace} and the λ_{max} statistics indicate the presence of three significant nonzero cointegrating vectors ($r = 3$).

It is important to note that inclusion of a stationary variable in the cointegrating space will increase the number of cointegrating vectors found by the Johansen test. Thus, one or more of the cointegrating vectors found may be the result of a stationary variable in the cointegrating space, rather than the presence of a long-run relationship among the variables. Before estimating the VECM, we use the likelihood ratio (LR) test proposed by Johansen and Juselius (1992) to test the null hypothesis that each variable in the cointegration space is stationary. The test statistic, calculated in combination with the determination of the cointegration rank, follows the chi-square (χ^2) distribution with the degrees of freedom equal to number of non-stationary β-vectors, i.e. $(n - r)$. Results of this test confirm the findings of the unit root tests in the preceding section, indicating that each of the six variables is non-stationary in levels.[8]

Table 1. Cointegration Tests and Long-run Exclusion Tests.

Panel A: Cointegration Tests

λ_{Trace}	C.V. (0.95)	H_0	λ_{max}	C.V. (0.95)
158.35*	102.14	r = 0	60.03*	40.30
98.32*	76.07	r = 1	35.98*	34.40
62.35*	53.12	r = 2	28.82*	28.14
33.52	34.91	r = 3	15.58	22.00
17.94	19.96	r = 4	11.93	15.67
6.01	9.24	r = 5	6.01	9.24

Panel B: Long-run Exclusion (χ^2) Test

r	S	P	X	R	M	F
3	8.17*	11.82*	23.73*	9.44*	6.62**	51.79*

Notes: The critical values are taken from Osterwald-Lenum (1992, Table 1*). An *(**) indicates significance at the 5% (10%) level.

Next we test the model to determine whether all six variables should remain in each of the three cointegrating vectors. The null hypotheses are defined as follows:

H_0: $\beta_{ij} = 0$ i = 1, 2, 3 (the cointegrating vectors) and j = 1, 2, . . . , 6 (the variables).

The LR test statistic derived by Johansen and Juselius (1992) is asymptotically distributed as $\chi 2(r)$. Table 1, panel B includes results of the long-run exclusion tests. The test statistics suggest that none of these *I*(1) variables should be excluded from the cointegration space.

Johansen and Juselius (1994) suggest that although direct interpretation of the unrestricted cointegrating vectors may be possible, the results should only be considered indicative and should not replace a more general procedure for identifying structural relations. Further, since any linear combination of the three stationary vectors is also a stationary vector, the estimates produced for any particular vector in β are not necessarily unique. To further refine the specification of the VECM, we impose restrictions, motivated by theory, on β and test whether the three vectors of β are uniquely identified in the cointegration space.

Johansen and Juselius (1994) define a set of restrictions that will identify the unique vectors. Specifically, let H_i be the restriction matrix for the *i*th cointegrating vector of dimension $n \times s_i$ such that H_β: $\beta = (H_1\varphi_1, \ldots, H_r\varphi_r)$, where s_i is the number of freely estimated parameters in β_i, and each φ_i is an $(s_i \times 1)$ vector of parameters to be estimated in the *i*th cointegrating relation.

Let $M_{i.jm} = H_j'H_m - H_j'H_i(H_i'H_i)^{-1}H_i'H_m$ (for $i \neq j$, m; and i, j, $m = 1, 2, 3$), then the rank condition is necessary and sufficient for identification of the first vector ($i = 1$) and requires:

$$\text{rank}(M_{1.22}) \geq 1,\ \text{rank}(M_{1.33}) \geq 1,\ \text{rank}\begin{bmatrix} M_{1.22} & M_{1.23} \\ M_{1.32} & M_{1.33} \end{bmatrix} \geq 2 \tag{2}$$

Similar expressions test whether the restrictions identify the other two vectors, $i = 2, 3$, in β.[9]

We test restrictions on β implied by the hypotheses in H_i jointly using a likelihood ratio (LR) test.[10] Johansen and Juselius (1994) provide a numerical optimization procedure based on successive reduced rank regressions concentrating the likelihood on α and β. The LR test is χ^2-distributed with $v = \Sigma_i(n + 1 - r - s_i)$ degrees of freedom. Assuming identification, this procedure produces the unique cointegrating vectors.

Unless economic theory is particularly uninformative about the hypotheses to be tested, testing joint restrictions on each β_i spanning the cointegration space is not usually the best way to start. Following the testing procedure suggested by Johansen and Juselius (1992), we propose two long-run relationships and formulate the test as H: $\beta = (H_1\varphi_1, H_2\varphi_2)$, where H_1 and H_2 include two sets of variables, the first set comprising the $r_1 = 2$ 'known' vectors and the second set comprising $r_2 = 1$ unrestricted vector. If we fail to reject both 'known' vectors, then we impose restrictions on variables in the third cointegrating vector and test the three restricted vectors jointly. Failure to reject these joint restrictions on the three vectors suggests that, given the rank condition is satisfied, the restrictions identify the three cointegrating vectors.

Following Geske and Roll (1983) and in the spirit of the St. Louis model of real activity, we begin by specifying a long-run relationship between real activity, Federal budget deficits and the money base, $\beta_1'Z = f(X, F, M)$. Rather than imposing the constraint suggested by the St. Louis model that only changes in the money base have a long-run impact on output, along with the industrial production index we include both the money base and Federal deficits in the first cointegrating vector to be tested. By normalizing this vector on X (real activity), the error correction term represents the deviation of industrial production, the money supply and fiscal deficits from their long-run equilibrium relationship.

The next hypothesized cointegrating vector represents a long-run equilibrium between real activity, inflation and interest rates: $\beta_2'Z = f(X, R, P)$. Hendry and Mizon (1993) identify a similar cointegrating relationship which they call 'excess demand'. Fama (1981), and Fama and Gibbons (1982) rely on a combination of the quantity theory of money and money demand theory to motivate an inverse relationship between real activity and inflation. In contrast, Geske and Roll (1983) use a Phillips curve model to argue that inflation and real activity will be positively correlated. In addition, the Fisher hypothesis suggests a positive relationship between inflation and interest rates. Therefore, we include X, R, and P in the second proposed cointegrating vector.

Before attempting to identify the third cointegrating vector, we test the specification of the first two vectors. To satisfy the rank condition for identification and to facilitate testing, we impose *just identifying* restrictions on the third cointegrating vector. Johansen and Juselius (1994) document that it is always possible to impose $s_i = r - 1$ *just identifying* restrictions on any cointegrating vector (β_i) without changing the likelihood function. Given the first two *over-identified* vectors, we randomly impose $r - 1 = 2$ zero restrictions on the beta coefficients of X and P in the third vector, and leave the coefficients on S, F, M, and R unrestricted. The resulting LR statistic tests whether the two 'known' vectors hold jointly. The LR statistic, reported as case 1 in Table 2 is 1.00 with 2 degrees of freedom, and is not significant at even the 10% level.[11] Therefore, with the third vector *just identified*, we fail to reject the first two *over-identified* vectors, the policy-real activity relation (X, F, M) and the excess demand relation (X, R, P).

In order to better understand the long-run structure of the model, we follow Hendry and Doornik's (1994) general-to-specific modeling approach and principle of parsimony, and impose *over-identifying* restrictions on the third vector. Thus, we further impose *over-identifying* restrictions, i.e. more than $r - 1 = 2$, on the third cointegrating vector and rely on econometric testing to specify this long-run equilibrium relationship. Recall that exclusion tests indicate that none of the variables should be omitted from the cointegrating space, and that the stock price variable is not included in either of the first two identified cointegrating vectors. These results imply that the third vector should include stock prices. When testing the third cointegrating vector (β_3), vectors one and two (β_1 and β_2) remain as specified above.

In Table 2, we report test results for ten possible over-identifying restrictions on the third vector in the cointegration space (see cases 2–11).[12] Case 4, in which S, F and R remain in the third cointegrating vector, represents the only set of over-identifying restrictions that we fail to reject.[13] This last cointegrating vector represents a long-run equilibrium relationship between stock prices,

Table 2. Tests of Identification Restrictions in the Cointegrating Vectors.

Test Cases		S	P	X	R	M	F	LR statistics
				Restricted β				
	β_1	0	0	1	0	–	–	
	β_2	0	–	1	–	0	0	
1	β_3	1	0	0	–	–	–	$\chi^2(2) = 1.00$
2	β_3	1	0	–	–	0	0	$\chi^2(3) = 7.04$**
3	β_3	1	–	0	–	0	0	$\chi^2(3) = 7.04$**
4	β_3	1	0	0	–	0	–	$\chi^2(3) = 1.11$
5	β_3	1	0	0	–	–	0	$\chi^2(3) = 10.21$*
6	β_3	1	–	–	0	0	0	$\chi^2(3) = 7.04$**
7	β_3	1	0	–	0	0	–	$\chi^2(3) = 15.75$*
8	β_3	1	0	–	0	–	0	$\chi^2(3) = 12.67$*
9	β_3	1	–	0	0	0	-	$\chi^2(3) = 8.02$*
10	β_3	1	–	0	0	–	0	$\chi^2(3) = 14.46$*
11	β_3	1	0	0	0	–	–	$\chi^2(3) = 15.77$*

Notes: The degrees of freedom are calculated by $v = \Sigma_i(n + 1 - r - s_i)$. The "–" in the restricted β column indicates no restriction on that coefficient. An * in the LR statistics column indicates the rejection of the null at the 5% significance level, while ** indicates rejection at 10% level. The rank conditions are satisfied for all cases.

interest rates and Federal deficits, providing evidence that government borrowing impacts the relationship between the stock market and interest rates as suggested by the crowding out hypothesis.

Sensitivity test results show that the three *uniquely* identified cointegrating vectors are invariant to changes in the lag length in the VAR model. Further, the three identified cointegrating relationships are stable using the October 1987 stock market crash and changes in Federal Reserve operating procedures of October of 1979 and October of 1982 as break points.[14] The coefficient estimates, with standard errors in parentheses, for each of the three cointegrating vectors are shown below.

$$\hat{\beta}_1' Z_t = X_t - \underset{(0.010)}{0.286} M_t - \underset{(0.001)}{0.004} F_t - \underset{(0.052)}{2.887} \tag{3}$$

$$\hat{\beta}_2' Z_t = X_t + \underset{(0.009)}{0.096} R_t - \underset{(0.010)}{0.416} P_t - \underset{(0.047)}{2.791} \tag{4}$$

$$\hat{\beta}_3' Z_t = R_t + \underset{(0.141)}{1.366} S_t - \underset{(0.012)}{0.127} F_t - \underset{(0.799)}{9.395} \tag{5}$$

Using the Johansen procedure, we propose to test the relevance of these long-run relationships in each of the six equations included in the VECM Eq. (1).

Next, we describe the specification of the short-run VECM and argue that short-run adjustments to departures from these long-run equilibria are responsible for the "spurious" inflation-stock returns relation.

THE SHORT-RUN DYNAMICS

Before we estimate the VECM, we subject our model to careful specification tests to ensure reliable statistical inferences. We begin with a series of over- and under-fitting tests of the VECM's lag specification, followed by a battery of diagnostic tests.

First, imposing common lags across all variables in a VAR (or VECM) model may produce biased results (Ahking & Miller, 1985). Therefore, we allow the lag length for each explanatory variable to range from one to twelve lags, using Akaike's minimum Final Prediction Error (FPE) criterion in conjunction with the "specific gravity" rule of Caines, Keng, and Sethi (1981) to specify the lag structure of each equation. Thornton and Batten (1985) report that the FPE criterion is superior to many other alternative lag-selection procedures. The FPE procedure minimizes a function of the one-step-ahead prediction error. The specific-gravity criterion aids in selecting the sequence in which the explanatory variables enter each of the equations. In the final specification, each equation includes only those variables that reduce the FPE when added to the regression model.[15] These procedures yield six equations whose dependent variables are explained by their own lags and, when they meet the FPE criterion, lags of the other five variables included in the model.

After specifying each equation, we estimate the system of six equations using an iterative version of Zellner's Seemingly Unrelated Regression (SUR) technique to improve the efficiency of our estimates. Next, we apply the Wald likelihood ratio test to the SUR variance-covariance matrix to assess the model's adequacy against over- and under-fitted versions of the system.[16] This test of the robustness of the lag structure of the estimated system (Hafer & Sheehan, 1991) results in some minor modifications. In Table 3 we report the final specification of the VECM model and the results of the likelihood ratio tests of Granger-causality run using the SUR estimation technique. At the risk of placing too much emphasis on signs and sizes of coefficients in a reduced-form model, we also report the sum of coefficients of all lags of each endogenous variable in the VECM.[17] Diagnostic test results generally confirm the adequacy of each of the six equations in the VECM, failing to find evidence of parameter instability and serial correlation, for example (results are available upon request).

Table 3. Granger-causality Test Results from VECM (Seemingly Unrelated Regression Estimates).

	Independent Variables								
Equations	$\Sigma\Delta S(L)$	$\Sigma\Delta P(L)$	$\Sigma\Delta X(L)$	$\Sigma\Delta R(L)$	$\Sigma\Delta M(L)$	$\Sigma\Delta F(L)$	$\beta_1' Z_{t-1}$	$\beta_2' Z_{t-1}$	$\beta_3' Z_{t-1}$
ΔS_t	$\Sigma = 0.21$	0	0	$\Sigma = -0.12$	$\Sigma = -1.07$	0	0.04	–0.08	–0.001
	23.05(2)*			38.15(10)*	3.02(1)		0.15	1.17	0.21
ΔP_t	0	$\Sigma = 0.78$	$\Sigma = 0.04$	0	$\Sigma = 0.33$	0	–0.006	0.02	0.0002
		181.80(11)*	10.13(3)*		26.58(5)*		1.05	23.89*	0.68
ΔX_t	$\Sigma = 0.11$	0	$\Sigma = 0.29$	$\Sigma = 0.08$	0	0	0.04	–0.06	–0.001
	31.92(6)*		28.83(1)*	21.80(5)*			3.89*	14.87*	3.83*
ΔR_t	$\Sigma = 0.20$	$\Sigma = -0.85$	$\Sigma = 1.85$	$\Sigma = 0.18$	$\Sigma = 5.02$	$\Sigma = 0.004$	0.49	–0.18	0.004
	17.91(3)*	38.15(7)*	21.44(5)*	57.74(6)*	12.45(2)*	11.54(4)*	6.25*	1.27	0.33
ΔM_t	$\Sigma = 0.01$	$\Sigma = 0.06$	0	$\Sigma = -0.01$	$\Sigma = 0.40$	0	0.04	–0.03	–0.001
	10.18(2)*	9.58(5)*		5.65(3)	19.41(5)*		20.63*	17.17*	20.42*
ΔF_t	0	0	0	0	$\Sigma = -158.17$	$\Sigma = -8.76$	23.77	–58.07	3.39
					35.91(11)*	458.87(12)*	0.65	5.17*	7.73*

Notes: Σ denotes the sum of estimated coefficients of each endogenous variables with lags in the VECM. The number underneath the sum of coefficients in each cell is the likelihood ratio statistic estimated by seeming unrelated regression (SUR). The number in parenthesis (L) reports the lag structure based on FPE and over- and under-fitting diagnostic tests. A zero cell denotes the exclusion of the lagged variables in the equation suggested by diagnostic tests. An * indicates the rejection of the null hypothesis of non-causality at the 5% significance level.

We first discuss the short-run Granger-causality implications obtained from the VECM. Consistent with most previous findings, we do not observe a direct causal relationship between inflation and stock returns. We do observe a causal link between stock returns and real activity growth ($\Delta S_t \rightarrow \Delta X_t$, LR(6) = 31.92*, see Table 3),[18] suggesting that stock market returns anticipate changes in real economic activity. Changes in real activity appear to lead to changes in the inflation rate ($\Delta X_t \rightarrow \Delta P_t$, LR(3) = 10.13*). These results are consistent with Fama's (1981) theory that the observed inverse relationship between stock returns and inflation is indeed a proxy for changes in real economic activity.

Interestingly, there appears to be no direct causality between real activity and policy actions (see Table 3). Thus, monetary policy and fiscal policy proxies may not respond to changes in economic conditions. Moreover, we fail to find significant causality from changes in deficits to changes in the money base, suggesting that the Federal Reserve bank does not monetize Federal deficits, at least in the short-run. These results fail to confirm Geske and Roll's (1983) conjecture that stock prices decline in response to an anticipated decrease in economic activity and increases in government expenditures and the money supply that are expected to follow.

Thorbecke (1997) suggests that monetary policy actions are more likely to impact stock returns than are changes in fiscal policy. As indicated in Table 3, stock returns lead (signal) changes in the money base ($\Delta S_t \rightarrow \Delta M_t$, LR(2) = 10.18*), while changes in the money base lead inflation ($\Delta M_t \rightarrow \Delta P_t$, LR(5) = 26.58*). This finding corresponds with Marshall's (1992) argument that inflation-asset return correlation may be generated by fluctuations in both the money supply and real activity. Below, we suggest that fluctuations in both the money supply and real activity are related to an underlying long-run relationship.

First, recall that the cointegrating relationship, $\beta_1' Z_{t-1}$, represents departures from long-run equilibrium in the public sector (between real activity and policy actions). Disequilibrium in the public sector (e.g. insufficient or excess money supply or government expenditures) affects the growth rate of real activity ($\beta_1' Z_{t-1} \rightarrow \Delta X_t$: LR(1) = 3.89*), which precedes changes in the inflation rate ($\Delta X_t \rightarrow \Delta P_t$: LR(3) = 10.13*). This sequence appears to induce the observed negative correlation between real activity and inflation, *a la* Fama (1981). Departures from long-run equilibrium in the public sector also appears to result in changes in the money base ($\beta_1' Z_{t-1} \rightarrow \Delta M_t$: LR(1) = 20.63*), suggesting that the Federal Reserve bank responds quickly to disequilibrium in the public sector. Changes in the money base then directly impact the rate of inflation ($\Delta M_t \rightarrow \Delta P_t$: LR(5) = 26.58*). To the extent that stock market returns anticipate both changes in real economic activity and changes in the money base, stock returns exhibit a contemporaneous negative correlation with inflation.

Next, when the economy expands (contracts), excess demand (supply) in the private sector, represented by $\beta_2' Z_{t-1}$, may result. In this case, excess demand (supply) leads to increases (decreases) in the growth of real economic activity ($\beta_2' Z_{t-1} \rightarrow \Delta X_t$: LR(1) = 14.87*), which further alter the rate of inflation. This adjustment will continue until equilibrium in the private sector is restored. The Federal Reserve also responds to the disequilibrium in private sector ($\beta_2' Z_{t-1} \rightarrow \Delta M_t$: LR(1) = 17.17*), giving rise to changes in base money and inflation. Again, this sequence of events appears to be anticipated (signaled) by the stock market. Therefore, the short-run adjustments to the disequilibrium in the private sector also contributes to the inverse relationship between stock returns and inflation.

The third cointegrating vector represents long-run equilibrium among short-term interest rates (R), stock prices (S), and Federal budget deficits (F). This may be considered a yield equilibrium relationship. When disequilibrium occurs, represented by the error-correction term ($\beta_3' Z_{t-1}$), an interest rate above long-run equilibrium results in an increase in the cost of capital and hence a reduction in private investment and reduced growth in real activity ($\beta_3' Z_{t-1} \rightarrow \Delta X_t$: LR(1) = 3.83*). The reduction in real activity growth will reduce inflation ($\Delta X_t \rightarrow \Delta P_t$: LR(3) = 10.13*). Further, an interest rate above the rate suggested by stock prices and the budget deficits is associated with a decrease in the money base ($\beta_3' Z_{t-1} \rightarrow \Delta M_t$, LR(1) = 20.42*), which moves the interest rate toward the long-run equilibrium rate ($\Delta M_t \rightarrow \Delta R_t$: LR(2) = 12.45**). Changes in the interest rate impact the rate of return for common stocks ($\Delta R_t \rightarrow \Delta S_t$: LR(10) = 38.15**). These results suggest that observed Federal Reserve actions may be in response to changes in interest rates, rather than monetization activities. Thus, our results appear to be inconsistent with Geske and Roll's (1983) deficit-monetization explanation of the inflation-stock returns correlation.

SUMMARY AND CONCLUSION

In this paper we re-examine the "troubling" relationship between stock returns and inflation, using Johansen and Juselius' (1992, 1994) cointegration technique and a vector-error-correction model. Unlike most previous studies, we account for both long-run and short-run structures in testing the validity of Fama's (1981) proxy hypothesis and Geske and Roll's (1983) reverse causality hypothesis of the negative correlation between inflation and stock returns. A battery of tests of cointegration, structural hypotheses and identification indicate the presence of three economically meaningful long-run relationships. We identify cointegrating vectors representing: (i) a policy-real activity equilibrium, (ii) a real activity equilibrium in the private sector, and (iii) a yield equilibrium relationship between the three-month Treasury bill rate, stock prices and Federal deficits.

Our short-run causality test results favor Fama's explanation rather than that of Geske and Roll. More importantly, the inverse relationship between stock returns and inflation during the post-war period appears to have its roots in the long-run equilibria among real activity and security yields. When disequilibria occur, both real economic conditions and monetary activities adjust so as to return the system to equilibrium. While inflation and stock returns respond to these adjustments, the "spurious" negative correlation arises as participants in the stock market anticipate the changes in real activity, the money supply and inflation. In summary, our evidence suggests that the three long-run relations and the short-run adjustments as the economy moves toward equilibrium offer a plausible explanation of the negative correlation between stock returns and inflation.

NOTES

1. See, for example, Fama (1981), Fama and Gibbons (1982), Geske and Roll (1983), James, Koreisha and Partch (1985), Stulz (1986), Kaul (1987, 1990), Lee (1992), Marshall (1992), Domain, Gilster and Louton (1996), Hess and Lee (1999). The next section provides a brief literature review.

2. Work focusing on short-run relationships between inflation and stock returns includes that of James, Koreisha and Partch (1985), Lee (1992), Hess and Lee (1999), and others.

3. Throughout the remainder of this paper *causality* refers to Granger-causality. Thus, X *causes* Y if including lagged values of X, in addition to lagged values of Y, improves the accuracy of predictions of Y.

4. While Pantula et al. (1994) argue that the weighted symmetric test is more powerful than several alternative uni root tests, we conduct all three tests to ensure the robust inferences. The unit root test results are available upon request.

5. The notion that Johansen and Juselius' (1992) system estimates of cointegration outperform the Engle-Granger (1987) single equation method is widely acknowledged. For a detailed comparison, see McDonald and Taylor (1993) p. 95, Harris (1995, p. 72).

6. In general, we also find the residuals from each equation to be free from serial correlation with 11 lags in the VAR model, with the deficits equation being the exception. To conserve space, we do not report the AIC and Sims LR test statistics, butthey are available upon request.

7. The test procedure involves moving from the most restrictive model to the least restrictive model and comparing the trace statistic (or λ_{max}) to its critical value at each stage. The progression stops the first time the null hypothesis is not rejected.

8. The LR test statistics are 25.54 (S), 21.93 (P), 22.61 (X), 20.51 (R), 23.06 (M), 22.67 (F), which are compared to $\chi^2_{0.95}(3)$ statistic of 7.81.

9. For a detailed description, refer to Johansen and Juselius (1994) and Harris (1995, p. 111).

10. Here, all restrictions on the cointegrating vectors refer to zero restrictions.

11. The degrees of freedom are equal to $v = \Sigma_i(n - r + 1 - s_i) = (6 - 3 + 1 - 3) + (6 - 3 + 1 - 3) + (6 - 3 + 1 - 4) = 2$. The H-matrices used to identify these cointegrating vectors are available upon request.

12. The degrees of freedom $v = \Sigma_i(n - r + 1 - s_i) = (6 - 3 + 1 - 3) + (6 - 3 + 1 - 3) + (6 - 3 + 1 - 3) = 3$.

13. To test the robustness of this result, we further restrict four of six beta coefficients in the third cointegrating vector, rejecting all combinations that include stock prices and meet the rank condition for identification. Detailed test results are available upon request.

14. The stability tests of the cointegrating relationships are proposed by Hansen and Johansen (1993) and equivalent to the trace tests in different sub-samples, and the results are available upon request.

15. See Darrat and Brocato (1994) for details.

16. The FPE method may result in over-fitting. We thank an anonymous referee for alerting us this issue.

17. For example, the sum of coefficients of all seven lags of inflation variables in the interest rate equation is –0.85 (the Fisher elasticity). However, if we count only significant coefficients, the Fisher elasticity is 1.76 (not reported here), similar to Crowder and Wohar's (1999) estimates of the range of 1.33 to 1.49. This suggests that our empirical results are quite consistent with previous literature of Fisher theory of interest rate and inflation.

18. An * indicates significance at the 5% level.

ACKNOWLEDGMENTS

We wish to thank Dwight Anderson, Ali Darrat, John Barkoulas, and two anonymous referees for their helpful comments and suggestions. The usual caveat applies.

REFERENCES

Ahking, F. W., & Miller, S. M. (1985). The Relationship Between Government Deficits, Money Growth, and Inflation. *Journal of Macroeconomics*, *7*, 447–467.

Banerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. F. (1993). *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.

Boudoukh, J., & Richardson, M. (1993). Stock Returns and Inflation: A Long Horizon Perspective. *American Economic Review*, *83*(5), 1346–1355.

Caines, P. E., Keng, C. W., & Sethi, S. P. (1981). Causality Analysis and Multivariate Autoregressive Modelling with an Application to Supermarket Sales Analysis. *Journal of Economic Dynamics and Control*, *3*, 267–298.

Crowder, W. J., & Wohar, M. E. (1999). Are Tax Effects Important in the Long-run Fisher Relationship? Evidence from the Municipal Bond Market. *Journal of Finance*, *54*, forthcoming.

Darrat, A. F., & Brocato, J. (1994). Stock market Efficiency and the Federal Budget Deficit: another Anomaly? *The Financial Review*, *29*, 49–75.

Domian, D. L., Gilster, J. E., & Louton, D. A. (1996). Expected Inflation, Interest Rates, and Stock Returns. *The Financial Review*, *31*, 809–830.

Engle, R. F., & Granger, C. W. J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*, **55**, 251–276.

Fama, E. F. (1981). Stock returns, Real Activity, Inflation and Money. *American Economic Review*, *71*, 545–565.

Fama, E. F., & Gibbons, M. R. (1982). Inflation, Real Returns and Capital Investment. *Journal of Monetary Economics*, *9*, 297–323.

Fama, E. F., & Schwert, G. W. (1977). Asset Returns and Inflation. *Journal of Financial Economics*, *5*, 115–146.

Fisher, I. (1930). *The Theory of Interest*. Macmillan, New York.

Geske, R., & Roll, R. (1983). The Monetary and Fiscal Linkage Between Stock Returns and Inflation. *Journal of Finance*, *38*, 1–33.

Granger, C. W. J. (1988). Causality, Cointegration, and Control. *Journal of Economic Dynamics and Control*, *12*, 551–559.

Hafer, R. W., & Sheehan, R. G. (1991). Policy Inferences Using VAR Models. *Economic Inquiry*, *29*, 44–52.

Hall, A. (1994). Test for a Unit Root in Time Series with Pretest Data-Based Model Selection. *Journal of Business and Economics Statistics*, *12*, 461–470.

Hansen, H., & Johansen, S. (1993). Recursive Estimation in Cointegrated VAR-Models, working paper. Institute of Mathematical Statistics, University of Copenhagen.

Harris, R. (1995). *Using Cointegration Analysis in Econometric Modelling*. London: Prentice Hall Publisher.

Hendry, D. F., & Doornik, J. A. (1994). Modelling Linear Dynamic Econometric Systems. *Scottish Journal of Political Economy*, *43*, 1–33.

Hendry, D. F., & Mizon, G. E. (1993). Evaluating Dynamic Econometric Models by Encompassing the VAR. In: P. C. P. Phillips (Ed.), *Models, Methods and Applications of Econometrics*. Oxford: Basil Blackwell Publisher.

Hess, P. J., & Lee, B.-S. (1999). Stock Returns and Inflation with Supply and Demand Disturbances. *Review of Financial Studies*, *12*, 1203–1218.

James, C., Koreisha, S., & Partch, M. (1985). A VARMA Analysis of the Causal Relations among Stock Returns, Real Output, and Nominal Interest Rates. *Journal of Finance*, *40*, 1375–1384.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S. (1992). Determination of Cointegration Rank in the Presence of a Linear Trend. *Oxford Bulletin of Economics and Statistics*, *54*, 383–397.

Johansen, S., & Juselius, K. (1992). Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for U.K. *Journal of Econometrics*, *53*, 211–244.

Johansen, S., & Juselius, K. (1994). Identification of the Long-Run and the Short-Run Structure: An Application to the ISLM Model. *Journal of Econometrics*, *63*, 7–36.

Kaul, G. (1987). Stock Returns and Inflation: The Role of the Monetary Sector. *Jounal of Financial Economics*, *18*, 253–276.

Kaul, G. (1990). Monetary Regimes and the Relation between Stock Returns and Inflationary Expectations. *Journal of Financial and Quantitative Analysis*, *25*(3), 307–321.

Lee, B-S. (1992). Causal Relations Among Stock Returns, Interest Rates, Real Activity and Inflation. *Journal of Finance*, *47*, 1591–1603.

Lintner, J. (1975). Inflation and Security Return. *Journal of Finance*, *30*, 259–280.

Marshall, D. A. (1992). Inflation and Asset Returns in a Monetary Economy. *Journal of Finance*, *47*, 1315–1342.

McDonald, R., & Tayalor, M. P. (1993). The Monetary Approach to the Exchange Rate. *IMF Staff Papers*, *40*, 89–107.

Osterwald-Lenum, M. (1992). A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics*, *50*, 361–377.

Pantula, S. G., Gonzalez-Farias, G., & Fuller, W. A. (1994). A Comparison of Unit-Root Test Criteria. *Journal of Business and Economic Statistics*, *12*, 449–459.

Reimers, H.-E. (1992). Comparisons of Tests for Multivariate Cointegration. *Statistical Papers*, *33*, 335–359.

Stulz, R. M. (1986). Asset Pricing and Expected Inflation. *Journal of Finance*, *41*, 209–223.

Thorbecke, W. (1997). On stock market returns and monetary policy. *Journal of Finance*, *52*, 635–654.

Thornton, D. L., & Batten, D. S. (1985). Lag-Length Selection and Tests of Granger Causality Between Money and Income. *Journal of Money, Credit, and Banking*, *17*, 164–178.

VALUATION AND HEDGING OF AMERICAN-STYLE LOOKBACK AND BARRIER OPTIONS

Chuang-Chang Chang and San-Lin Chung*

ABSTRACT

This study developed a trinomial lattice approach for American-style lookback and barrier options. The contributions of this study are threefold. First, we constructed a trinomial model to price lookback options by extending the Hull and White (1993) binomial model. We also investigated the properties of the behavior in the difference between American and European lookback option values. Second, we incorporated the idea of path function, proposed by Hull and White (1993) with the trinomial tree approach developed by Ritchken (1995) to construct an efficient procedure which is able to value both European and American barrier options (especially for discretely observed options). Third, we demonstrated how to use the Pelsser and Vorst (1994) approach to compute the hedge ratios for barrier options in our trinomial model.

1. INTRODUCTION

The option markets have developed prosperously since Black and Scholes (1973) made a breakthrough on deriving the closed-form pricing formula for European options. Most of the option contracts traded in the market were plain vanilla options in the era of the 1970s. However, the over-the-counter markets

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 19–37.

ISBN: 0-7623-0798-6

throughout the world have provided a variety of so-called exotic options to meet their customers' demand for investment and hedging purposes since the 1980s. The trade volumes of these options have increased in recent years. Two important extensions of the plain vanilla options are barrier and lookback options. In some cases analytical valuation formulas can be found for these options, while in others numerical methods must be employed.

Hull and White (1993) extended the Cox, Ross and Rubinstein (hereafter CRR) (1979) binomial model to develop efficient procedures for valuing European and American exotic options. They incorporated a Markov path function into the CRR model and used lookback and Asian options as examples to illustrate the computational efficiency and accuracy of their model. However, continuously observed lookback options are not the usual cases. Most lookback option contracts use discrete observations to record the maximum or minimum values of the underlying assets. Hence, it is important to take observation frequency into consideration when one constructs model for pricing lookback options. Cheuk and Vorst (1997) succeeded in constructing a path-independent binomial model which could deal with the observation frequency problem encountered in pricing lookback options. From their simulation results, they demonstrated that the observation frequency will affect the lookback option value significantly.

As pointed out by Boyle and Lau (1994), the standard CRR binomial model can also be used to value barrier options, but the convergence of this method is very slow, and the results tend to have a persistent bias. They observed that the source of the problem arises from the allocation of the barrier with respect to adjacent layers of nodes in the lattice. They found that if the layers of the lattice were constructed so that the barrier falls between the layers of the lattice, the pricing bias may be quite significant. To avoid this pricing bias, they set up an algorithm to choose the binomial time step so that the constructed lattice has layers that are as close as possible to the barrier. Although Boyle and Lau's approach can reduce the size of the pricing bias, their method may not necessarily obtain more accurate results and may be difficult to implement for the cases of options with multiple or time-varying barriers.

Ritchken (1995) provides a simple and efficient trinomial model that can be used to price and hedge options that have constant or time-varying single barriers as well as the options that are subject to multiple barriers. He exploits the extra freedom to choose another parameter to put the nodes of the trinomial tree in the optimal position with respect to the barrier. In this way, he gets a very accurate approximation for barrier option values, even for the trees with a small time step, provided the underlying asset price is not too close to the barrier. However, he still needs a large number of steps to obtain accurate values when the underlying asset price is very close to the barrier. To solve the problem encountered in Ritchken's model, Cheuk and Vorst (1996) introduce a shift parameter to change the geometry of the tree. This technique enables them not only to compute accurate barrier option values with a small number of steps, but also to value options with arbitrary time-varying barriers.

This article combines the idea of path function, proposed by Hull and White (1993) with the trinomial tree approach developed by Ritchken (1995) to price discretely observed American-style barrier options. This is a further contribution in the literature because the above mentioned can only value the continuously observed, but not the discretely observed, American barrier options. For continuously observed American knock-in put option, typically the algorithm for pricing this option on a tree goes as follows: The price of the American knock-in put on the knock-in boundary is set to be equal to the price of the American put (which can be easily computed on the tree). The terminal payoff is equal to zero below the knock-in boundary. The discounting backward is performed only in the region below the knock-in boundary. At each node a decision of whether to exercise or not is made in the usual way. For details of this procedure, we refer the reader to Chriss (1997). However, the above procedure can not be applied to price discretely observed American knock-in put option because the price of the American knock-in put on the knock-in boundary is not equal to the price of the American put for the non-observed time points. In contrast, since path function is recorded in our method, it is very flexible to price discretely observed American-style complex (such as time-varying) barrier options.

The plan of this paper is as follows. In Section 2, we construct a trinomial model to value lookback options by extending the Hull and White (1993) binomial method. We also investigate the properties of the behavior in the difference between American and European lookback options. In Section 3, we incorporate the idea of path function proposed by Hull and White (1993) with the trinomial method developed by Ritchken (1995) to value both European and American barrier options by illustrations of knock-in options as examples. In Section 4, we show how to implement the Pelsser and Vorst (1994) approach to calculate the hedge ratios for barrier options in our trinomial model. We then draw conclusions in Section 5.

2. PRICING AMERICAN-STYLE LOOKBACK OPTIONS

We construct a trinomial model to value lookback options by extending the Hull and White (1993) binomial method in this section. One reason that we use a trinomial model is because the convergence in a trinomial model is faster than that in a binomial model. A trinomial model needs more time to compute than a binomial model with the same number of time steps. However, it can be shown that a trinomial model involves less computing time than a binomial model with twice as many time steps.[1]

Adopted from Hull and White (1993), we assume that the value of an option at time t is a function of t, the price of the underlying asset, S, and some function of the path followed by the asset price between time zero and time t, $F(S, t)$. We summarize the notation as follows:

$S(t)$: the asset price at time t;
$F(S, t)$: path function followed by the asset between time zero and time t;
$v(S, F, t)$: the option value at time t when the underlying asset price is S and the path function has a value of F;
r: risk-free interest rate (assume constant);
T: time to the maturity of the option;

The asset price, S, is assumed to follow, under the risk-neutral measure, a geometric Brownian motion:

$$dS = rSdt + \sigma Sdz \tag{1}$$

where r denotes the risk-free interest rate, σ the return volatility, and z a standard Brownian motion. We divide the time to maturity of the option into n equal time steps of length $\Delta t(\Delta t = T/n)$. Followed from Boyle (1988), the asset price, at any given time, can move into three possible states, up, down, or middle, in the next period. If S denotes the asset price at time t, then at time $t + \Delta t$, the price will be Su, Sd, or Sm. The parameters are defined as follows:

$$u = e^{\lambda\sigma\sqrt{\Delta t}}, \tag{2}$$

$$d = e^{-\lambda\sigma\sqrt{\Delta t}}, \tag{3}$$

and

$$m = 1 \tag{4}$$

Matching the first two moments (i.e. mean and volatility) of the risk-neutral returns distribution leads to the probabilities associated with these states as

$$P_u = \frac{u(V + M^2 - M) - (M - 1)}{(u - 1)(u^2 - 1)}, \tag{5}$$

$$P_d = \frac{u^2(V + M^2 - M) - u^3(M - 1)}{(u - 1)(u^2 - 1)}, \tag{6}$$

and

$$P_m = 1 - P_u - P_d, \tag{7}$$

with

$$M = e^{r\sqrt{\Delta t}}, \tag{8}$$

and

$$V = M^2[e^{\sigma^2\sqrt{\Delta t}} - 1]. \tag{9}$$

In our model, we chose the value of $\lambda = \sqrt{\frac{\pi}{2}}$, as suggested by Omberg (1988).

After setting the parameters, we can compute the asset price at any node (i, j) using the following equation.

$$S(i, j) = S(0) \times u^j \times d^{2i-j}, \tag{10}$$

As pointed out by Hull and White (1993), to value a path-dependent option, one can value the option at each node in the tree for all alternative values of the path function $F(t, S)$ that can occur. In our case, we set $F(t, S)$ equal to the maximum (minimum) asset price realized between time zero and time t as the strike price for lookback put (call) options. However, this approach must meet two requirements. First, the path function has to be Markov. Second, the number for $F(t, S)$ must not grow too fast with the size of the tree.

We let the kth value of F at node (i, j) as $F_{i,j,k}$ and define the corresponding option value as $v_{i,j,k}$. The option value at maturity date, $v_{n,j,k}$ is known for all j and k. To compute its value at node (i, j) where $i < n$, we illustrate that the asset price has a probability P_u of moving up to the node $(i+1, j+2)$, a probability P_m of moving to the node $(i+1, j+1)$, and a probability P_d of moving down to the node $(i+1, j)$.

We suppose that the kth value of F at node (i, j) moves to the k_uth value of F at node $(i+1, j+2)$ when there is an up movement in the asset price, and to the k_mth value of F at node $(i+1, j+1)$ when there is no change in the asset price, and to the k_dth value of F at node $(i+1, j)$ when there is down movement in the asset price. For a European lookback option, this means that

$$v_{i,j,k} = e^{-r\Delta t}[P_u v_{i+1, j+2, k_u} + P_m v_{i+1, j+1, k_m} + P_d v_{i+1, j, k_d}, \tag{11}$$

For an American lookback option, the value in Eq. (11) must be compared with the early exercise value, and $v_{i,j,k}$ is equal to the maximum of the two values.

We illustrate our trinomial model using a three-month American lookback put option on a stock without paying dividends as shown in Fig. 1. In this case, we assume that $S(0)$ is 50, σ is 40% annually, and r is 10% per year. Since we let the time step (n) equal 3, hence Δt equals 0.0833 year. Substituting this data into Eqs (2) to (10), we obtain $u = 1.1224$, $d = 0.8909$, $P_u = 0.3282$, $P_m = 0.3545$, and $P_d = 0.3173$.

Figure 1 shows the results of the backward calculations. The top number at each node is the stock price. The numbers in the second row present the possible values of F, the maximum stock price realized between time zero and time $i\Delta t$. The numbers in the third row show the corresponding option values. We demonstrate the calculation procedures using the nodes A, B, C and D as an

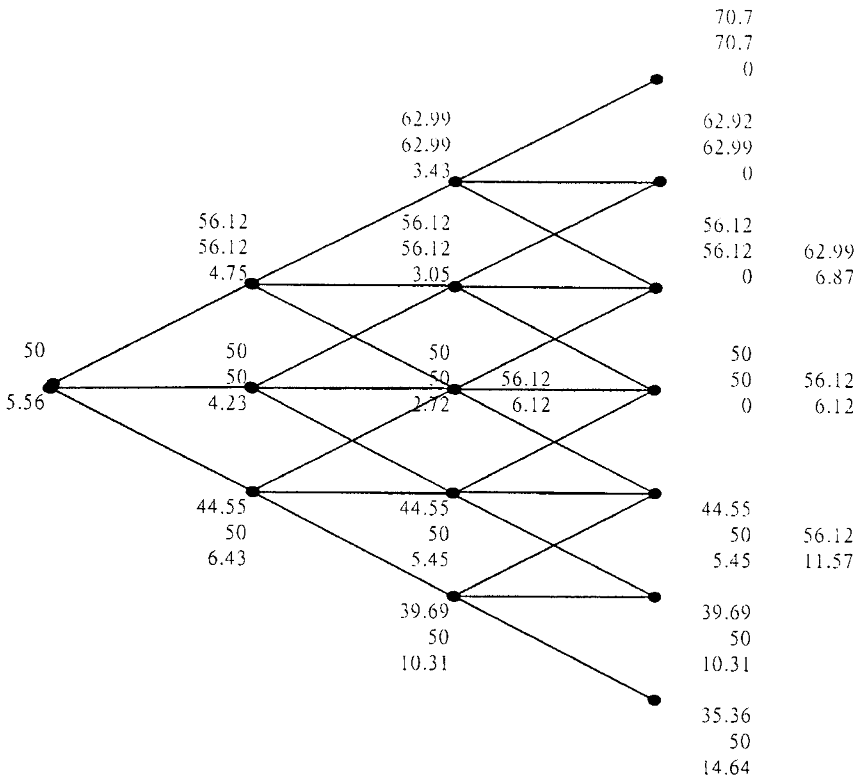


Fig. 1. Trinomial Tree for Valuing an American Lookback Put Option on a Stock Price.

example. Look at nodes *A*, *B*, *C* and *D* in Fig. 1 to see the way that the tree is constructed. At node *A*, there are two possible values for *F*. They are

$$F_{2,2,1} = 57.79; \qquad \mathrm{F}_{2,2,2} = 50.00$$

Similarly, for nodes *B*, *C*, and *D*, we have

$$F_{3,4,1} = 66.78; \qquad \mathrm{F}_{3,4,2} = 57.79$$

$$F_{3,3,1} = 57.79; \qquad \mathrm{F}_{3,3,2} = 50.00$$

$$F_{3,2,1} = 57.79; \qquad \mathrm{F}_{3,2,2} = 50.00$$

The payoff at node *B* when $F = F_{3,4,1}$ is given as

$$v_{3,4,1} = \max(66.78 - 57.79, 0) = 8.99,$$

Similarly, $v_{3,4,2} = 0$. At nodes *C* and *D*, we obtain

$$v_{3,3,1} = 7.79; \qquad \mathrm{v}_{3,3,2} = 0.00$$

$$v_{3,2,1} = 14.52; \qquad \mathrm{v}_{3,2,2} = 6.74$$

We then consider the state at node *A* when $F = 50.00$ (i.e. $k = 2$). If the stock price goes up, so that node *A* moves to node *B*, *F* will change from 50 to 57.79. From the notation expressed in Eq. (11), we know that $k_u = 2$. Further, if the stock price does not change, so that node *A* moves to node *C*, *F* remains at 50, and $k_m = 2$. Finally, if the stock price goes down, so that node *A* moves to node *D*, *F* remains at 50, and $k_d = 2$.

Using Eq. (11), we can obtain the option value at node *A* when $F = 50$ as follows:

$$\begin{aligned}&[v_{3,4,2} \times 0.3282 + v_{3,3,2} \times 0.3545 + v_{3,2,2} \times 0.3173]e^{-0.1 \times 0.0833}\\ &= 0 \times 0.3282 + 0 \times 0.3545 + 6.74 \times 0.3173]e^{-0.1 \times 0.0833}\\ &= 2.12\end{aligned}$$

Since the early exercise value at node *A* when $F = 50$ is zero, the option value at this situation is equal to 2.12.

A similar procedure for the state where $F = 56.12$ at node *A* resulted in $k_u = 2$, $k_m = 1$ and $k_d = 1$. In this case, the holding value of the option is given

$$\begin{aligned}&[v_{3,4,2} \times 0.3282 + v_{3,3,1} \times 0.3545 + v_{3,2,1} \times 0.3173]e^{-0.1 \times 0.0833}\\ &= 0 \times P_u + 7.79 \times P_m + 14.52 \times P_d]e^{-0.1 \times 0.0833}\\ &= 7.31\end{aligned}$$

Because the early exercise value (= 7.79) is greater than the holding value (7.31) in this case, early exercise is optimal. Hence, the option value equals

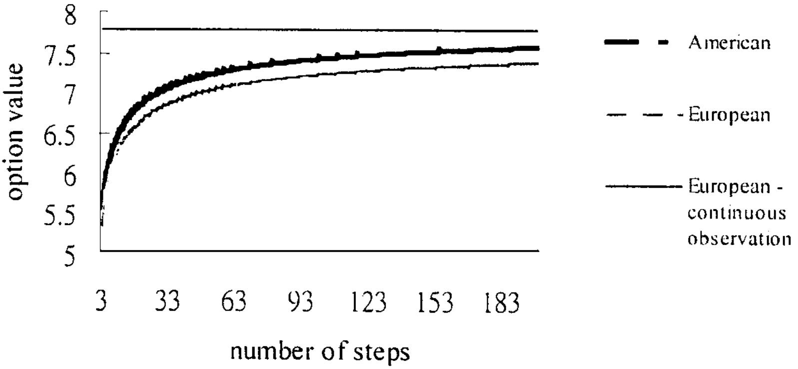


Fig. 2. Pricing European and American Lookback Options Using Trinomial Method.

7.79. Working back through the tree, repeating the above computational procedures at each node, we obtain the option value at time zero equals 4.87.

It is difficult to calculate the American lookback option values with continuous observation frequency since the convergence of numerical methods for pricing these options is very slow as pointed out by Cheuk and Vorst (1997). Hence it is worth finding a way for approximating the American option contracts with continuous observation frequency. As shown in Fig. 2, the difference between the European and American lookback option values with discrete observation frequency seems to be very stable for different numbers of time steps in our trinomial model.

To gain a deep insight into the pattern of the difference between the European and American lookback option values, we use different volatility to carry out simulations and report the results in Fig. 3. From Fig. 3, we find that the difference between the European and American lookback option values remains constant for both cases of low and high volatilities when the number of time steps of the tree is greater than 150. Hence we suggest that the difference between the European and American lookback option values with continuous observation frequency also has similar pattern. Therefore, to obtain the American lookback option values with continuous observation frequency, we can use our trinomial model with 150 time steps to compute the difference between the European and American lookback option values. Then we add the above difference value to the European lookback option values using the closed-form solutions derived using the continuous-time framework.[2]

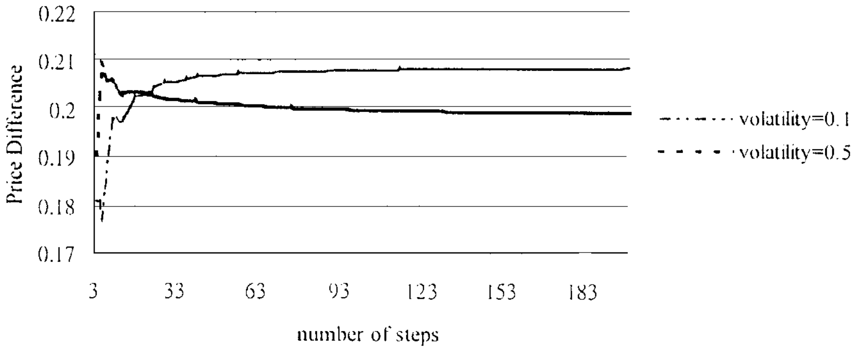


Fig. 3. Price Difference Between American and European Lookback Put Options.

3. PRICING AMERICAN-STYLE BARRIER OPTIONS

Although Boyle and Lau (1994), Ritchken (1995) and Cheuk and Vorst (1996) constructed models to value barrier options, they did not show how to price American barrier options. The American-style knock-in option is the toughest contract for academic to develop models for the valuation and hedging of these options. We will incorporate the path function idea proposed by Hull and White (1993) with the trinomial method developed by Ritchken (1995) to propose a procedure which is able to value both European and American knock-in options.

As pointed out by Ritchken (1995), it is important to put the trinomial tree nodes in the optimal position with respect to the barrier. The extra parameter λ allows us to decouple the time partition from the state partition, and hence to let the layers of the tree hit the barrier exactly. Therefore, to choose λ correctly is a crucial procedure in our trinomial model. We adopted the technique developed by Ritchken (1995) to choose λ.

To illustrate our approach, we use an up-and-in American put option as an example. With the parameter, λ, tentatively set at one, we compute the number of consecutive up movements that leads to the lowest layer of nodes above the barrier, b. This value is denoted as n_0 which is the largest value smaller than η, where η is given as follows:

$$\eta = \frac{ln(b/S(0))}{\sigma\sqrt{\Delta t}}, \tag{12}$$

Suppose that η is an integer, then we should set λ at the level of one. Otherwise, redefine the parameter λ as

$$\lambda = \frac{\eta}{n_0}, \tag{13}$$

Under the above construction, the value of λ will be equal to or greater than one but less than 2. With this specification for λ, the trinominal approach will result in a lattice which can lead to a layer of nodes that hit the barrier exactly. Successive movements of n_0 will take the price to the boundary.

When we use the backwardation approach to value American knock-in options, we must know whether the option is active or not at each note (i, j) in the tree. This can be done using the path function illustrated in the previous section. In our example, the path function will record the maximum value of the asset price between time zero and time $i\Delta t$. If the value of the path function is equal to or greater than the barrier, the option will be in and active, otherwise, the option will not be valid. The first crucial step in our approach is to decide the number and the possible value of the path function at each node (i, j). We summarize the computational algorithm[3] as follows:

$$k(i, j) = i - int\left[\frac{\mathrm{j}+1}{2}\right] + 1 \quad \text{if} \quad j \geq i, \tag{14}$$

and

$$F_{i,j,k} = S(0) \times u^{k-1} \tag{15}$$

where $k(i, j)$ is the number of path function at node (i, j) and $F_{i,j,k}$ is the kth value of the path function. Further,

$$k(i, j) = int\left[\frac{\mathrm{j}}{2}\right] + 1 \quad \text{if} \quad j < i, \tag{16}$$

and

$$F_{i,j,k} = S(0) \times u^{k-1}, \tag{17}$$

Another important step in our approach is to trace the regular rule for the relationship between $F_{i,j,k}$ at time $i\Delta t$ and the possible corresponding values, $F_{i+1,\,j+2,\,k_u}$, $F_{i+1,\,j+1,\,k_m}$ and $F_{i+1,\,j,\,k_d}$ at time $(i+1)\Delta t$. The computational algorithm can be summarized as follows:

$$k_u = \max(k-1,\, 1), \quad k_m = k \quad k_d = k+1, \quad \text{if} \quad j > i \tag{18}$$

$$k_u = \max(k-1,\, 1), \quad k_m = k_d = k, \quad \text{if} \quad j = i \tag{19}$$

and

$$k_u = k_m = k_d = k, \quad \text{if} \quad j < i \tag{20}$$

After setting the above computational algorithms, it is quite straightforward to compute the American knock-in put option values. The option value with the path function $F_{n,j,k}$ at the maturity date, $v_{n,j,k}$ will equal $\max(0, X - S(n, j))$ if $F_{n,j,k} \geq b$, otherwise $v_{n,j,k} = 0$, where X is the strike price of the option. To compute its value at node (i, j) where $i < n$, we have to compare the holding value with early exercise value, and $v_{i,j,k}$ will equal the greater of the two values. The holding value can be obtained using Eq. (11). However, the early exercise value depends on whether the option is active or not at node (i, j). If $F_{i,j,k} \geq b$, the early exercise value will equal $\max(0, X - S(i, j))$, otherwise the early exercise value will be equal to zero. We use an American up-and-in put option as example to show the computing procedures. In this case, we assume that $S(0)$ is 90, σ is 25% annually, r is 10% per year, the time to maturity of option equals 1 year, the strike price equals 90, the barrier price equals 105, and the number of time steps equals 4. Figure 4 shows the results. The top number at each node is the stock price. The numbers in the second row present the possible values of F, the maximum stock price realized between time zero and time $i\Delta t$. The numbers in the third row show the corresponding option values.

As mentioned earlier, Tian (1993) demonstrates that convergence in a trinomial model is faster than in a binomial model for pricing plain vanilla options. We validate that this statement is also true in the case of barrier options by carrying out simulations. The results are shown in Fig. 5. We note that the convergence performance of our trinomial model is much better than that of the binomial model. The value computed by our trinomial model converges to the value computed by Merton's (1973) closed-form solution even for very small number of time step.

To have more insights into barrier options, we also carry out simulations to investigate the relationship between European and American barrier option values using down-and-in and down and out call options as examples. We report the simulation results in Table 1. The results show that the European and American down-and-in call option have the same value if the underlying asset pays no dividend over the life of the option. The reason for this is as follows. Before touching the barrier, both options are not in and thus have the same value. Once the barrier is reached, they become the standard European and American call options respectively. As Merton (1973) had shown, both options must have the same value if the underlying asset pays no dividend. The same argument can be applied to the up-and-in call options.

For down-and-out call options, the European and American-styles have the same value if the strike price is larger than or equal to the barrier. The explanation is straightforward. When the stock price is very close to the barrier in the future, the value of the European down-and-out call option will be very small. Similarly, the value of the American down-and-out call option is also very small if the strike price is greater than or equal to barrier since the option is out-of-the-money in this case. In contrast, if the strike price is lower than the barrier, the American down-and-out call is in-the-money and it is optimal to early exercise this option. Thus the American down-and-out call option is more valuable than its counterpart European option if the strike price is smaller than the barrier.

It's well known that the value of a standard European put option equals the value of a European up-and-in put plus the value of a European up-and-out put.

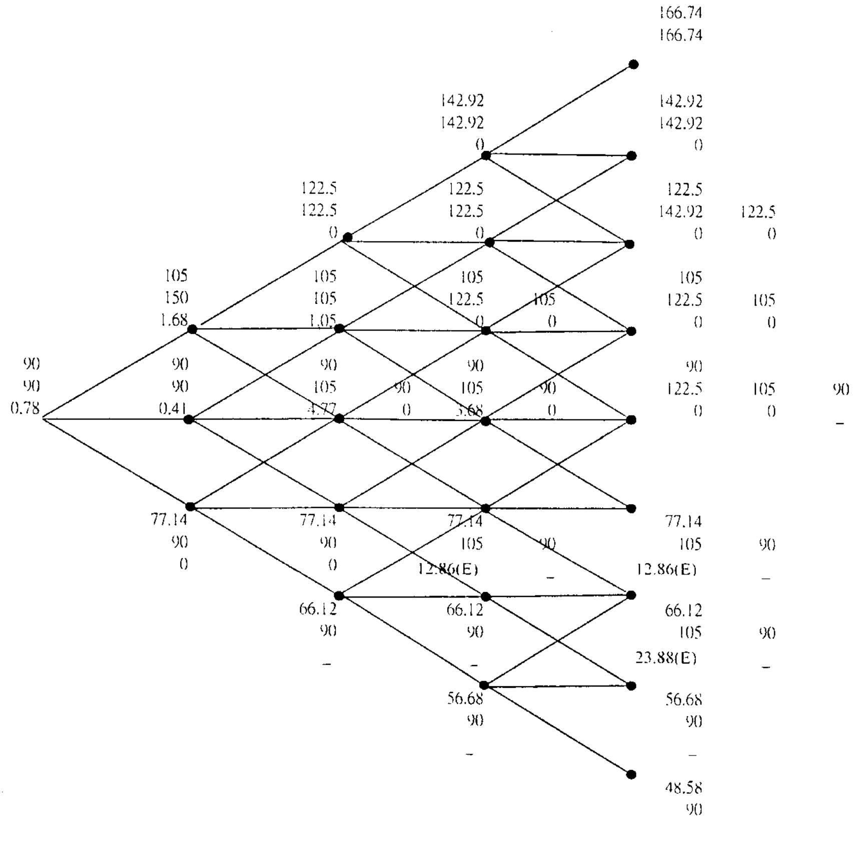


Fig. 4. Trinomial Tree for Valuing an American Up-and-In Put Option on a Stock Price.

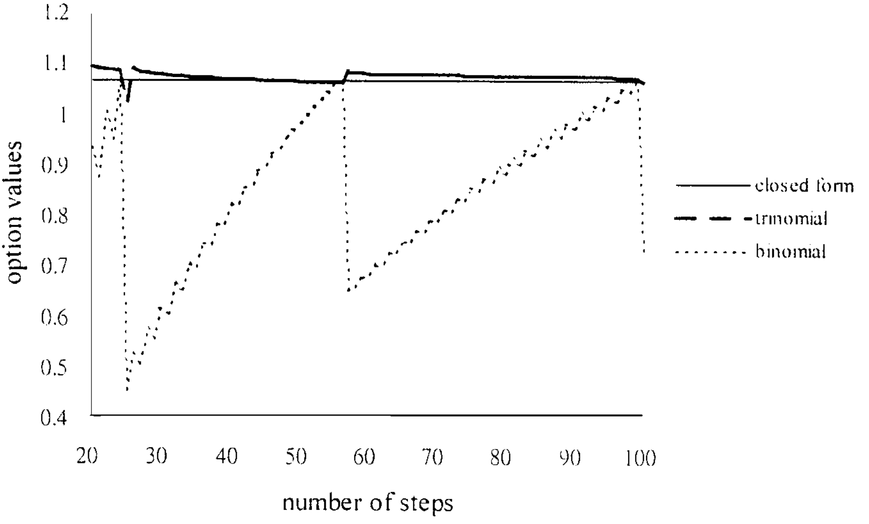


Fig. 5(a). The Convergence Performance of Using Binomial and Trinomial Methods to Price European Up-and-In Put Options.

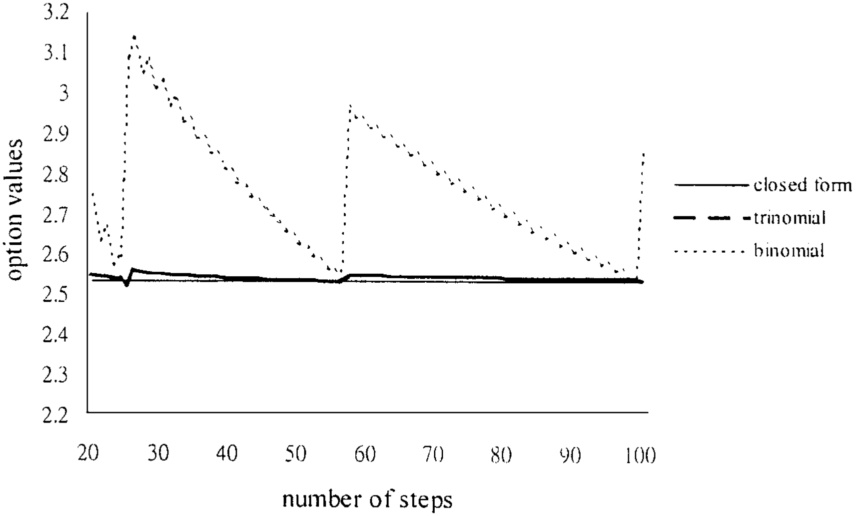


Fig. 5(b). The Convergence Performance of Using Binomial and Trinomial Methods to Price European Up-and-Out Put Options.

Table 1. Prices for the European and American Barrier Options.

	down-and-in call		down-and-out call	
strike price	European	American	European	American
105	4.3727	4.3727	5.0888	5.0888
100	5.6668	5.6668	6.0005	6.0005
95	7.2611	7.2611	6.9719	6.9719
90	9.1913	9.1913	7.9758	7.9758
85	11.4739	11.4739	8.9839	11.5340
80	14.0817	14.0817	9.9919	16.4437
75	16.9672	16.9672	11.0000	21.3535
70	20.0824	20.0824	12.0080	26.2633
65	23.3655	23.3655	13.0161	31.1730

$S(0) = 95$, $b = 90$, $r = 0.1$, $T = 1.0$, and $\sigma = 0.25$.

However, this is no longer true for the American-style options. Thus it is of interest to investigate the difference between the sum of two American barrier options' values and the standard American option value.

Figure 6 shows that the difference is positively correlated to the depth-in-the-money of the option and the distance between the stock price and the barrier.

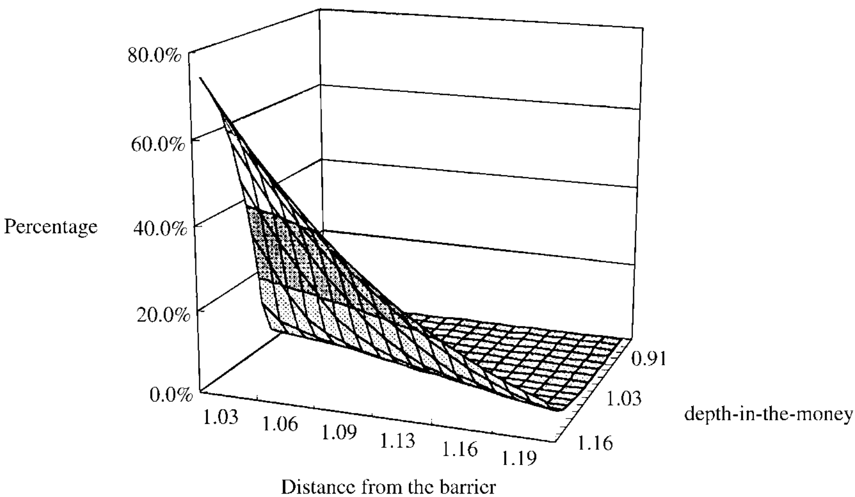


Fig. 6. The Percentage Difference between the Sum of American up-and-in and up-and-out Call Options' Values and the Standard American Call Option Value.

This phenomenon can be discussed as follows. When the barrier is far away from the stock price, the American up-and-out put is almost identical to the standard American put. The American up-and-in put has little value, and thus the difference is small. Concerning the depth-in-the-money, when the stock price is very close to the barrier in the future, the American up-and-in put option value is very close to the standard American option value. In the mean time, if the strike price is smaller than the barrier, the American up-and-out put is in-the-money and it would be optimal to exercise this option early. Thus the difference will increase as the depth-in-the-money increases. The results also indicate that if the option is out-of-the money when the stock price is very close to the barrier, the difference is very small (less than 2.5%) for the chosen parameters. In that case, the value of an American up-and-in put can be approximated as the standard American put value minus the counterpart American up-and-out put value.

4. HEDGING

Pelsser and Vorst (1994) develop a more efficient and accurate method to calculate hedge ratios by extending the binomial tree instead of by numerical differentiation. Their method can be also be applied to the trinomial tree as shown in Cheuk and Vorst (1996). In this Section, we show how to apply Pelsser and Vorst's method to calculate the hedge ratios for American-style lookback and barrier options. To apply their approach in our cases, we have to extend the tree one period back, with the central node at period -1 with the same underlying asset price as the central node at period zero.

In Fig. 7, the solid lines represent the original tree, while the dotted and solid lines together show the extended tree. In the case of barrier options, if the asset

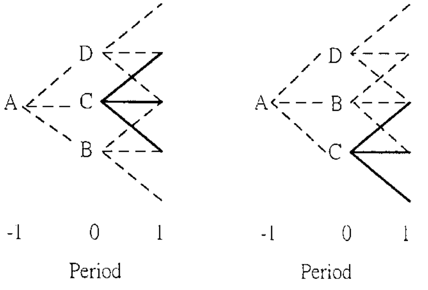


Fig. 7. Extended Trinomial Tree for Hedge Ratios.

price at node C in Fig. 7 is not very close to the barrier, the hedge ratios can be computed using the following equations.

$$\Delta = \frac{O(D) - O(B)}{S(D) - S(B)},$$

$$\Gamma = 2\frac{\dfrac{O(D) - O(C)}{S(D) - S(C)} - \dfrac{O(C) - O(B)}{S(C) - S(B)}}{S(D) - S(B)}, \tag{21}$$

where Δ and Γ are the delta and gamma of the options respectively, and $O(X)$ and $S(X)$ are the option value and the underlying asset price at node X, respectively. We can also use Eq. (21) to calculate the delta and gamma for lookback options.

However, if the underlying asset at node C, $S(C)$, is very close to the barrier, we must use the original node representing the current asset price and the one neighboring node that is inside the barrier to compute Δ. To compute the Γ of the option, we have to extend the tree with two new states in the direction away from the barrier as shown on the right hand side of Fig. 7. In this case, the hedge ratios are given as

$$\Delta = \frac{O(B) - O(C)}{S(B) - S(C)},$$

$$\Gamma = 2\frac{\dfrac{O(D) - O(B)}{S(D) - S(B)} - \dfrac{O(B) - O(C)}{S(B) - S(C)}}{S(D) - S(C)}, \tag{22}$$

We use European and American up-and-in and up-and-out options to demonstrate the accuracy of hedge ratios computed by the extended Pelsser and Vorst's approach. Employing the closed-form solution derived by Renier and Rubinstein (1993) as benchmarks, Table 2 shows that the bias of the delta for European up-and-in and up-and-out options is very large when we use traditional numerical differentiation in the original tree. However, the extended Pelsser and Vorst's approach can compute the delta of these options with a high degree of accuracy. Further, the delta and gamma of the American up-and-in and up-and-out options computed by traditional approach are quite different from those computed using the extended Pelsser and Vorst's approach. We believe that the hedge ratios computed using the extended Pelsser and Vorst's approach are more accurate than those computed using traditional numerical differentiation approach.

Table 2. Hedge Ratios for Barrier Options.

		European Put		American Put	
		up-and-in	up-and-out	up-and-in	up-and-out
delta	closed form	0.1380	− 0.5126	N.A.	N.A.
	traditional	0.1634	− 0.5075	0.1551	− 0.7168
	Pelsser and Vorst	0.1385	− 0.5127	0.1667	− 0.7288
gamma	closed form	0.0017	0.0148	N.A.	N.A.
	traditional	4.8912	0.9014	2.2979	0.2101
	Pelsser and Vorst	0.0176	0.0164	0.0126	0.0267

$S(0) = 95$, $K = 100$, $b = 105$, $r = 0.1$, $T = 1.0$, and $\sigma = 0.25$.

5. CONCLUSIONS

We have developed a trinomial lattice for pricing and hedging American-style lookback and barrier options. Our approach can calculate the discretely observed American-style barrier option values efficiently and accurately. Hence our model fills in the gap existing in the literature which is unable to provide methods for pricing the discretely observed American-style complex barrier options.

We find that the difference between the European and American lookback option values remains constant when the number of time steps in the tree is greater than 150. Therefore, we can use our trinomial model with 150 time steps to compute the difference in value between these two options. Then we add the difference in value to the European lookback option value using the closed-form solutions derived by continuous-time framework to obtain the American lookback option values with continuous observation frequency.

We also find that the European and American down-and-in call options have the same value if the underlying asset pays no dividend over the life of the option. Further, for down-and-out call options, the European and the American-styles have the same value if the strike price is larger than or equal to the barrier. Another interesting finding is that the difference between the sum of the two American barrier option values and the standard American option value is positively correlated to the depth-in-the-money of the option and the distance between the stock price and the barrier. Finally, we also show demonstrate to apply the Pelsser and Vorst's method to calculate the hedging ratios for American-style lookback and barrier options.

The approach developed in this paper could be extended to the case of options with double or time-varying barriers. Further, our approach can also be applied to price warrants with a reset strike price. However, we leave these interesting research topics to future research.

ACKNOWLEDGMENT

We thank the helpful comments from the participants at the Seventh Conference on Pacific Basin Finance, Economics and Accounting.

NOTES

* The earlier version of this paper was presented at the Seventh Conference on Pacific Basin Finance, Economics and Accounting.

1. Tian (1993) illustrates that the CRR binomial model needs more than twice the number of time steps to obtain the same accuracy of Boyle's (1998) trinomial model with pricing error within 1% of the true value for plain vanilla options. Thus, a trinomial model for pricing plain vanilla options has faster convergence than a binomial model, if we compare the computing time.
2. The closed-form solutions for European lookback options can be found in Goldman, Sosin and Gatto (1979) and Conze and Viswanathan (1991).
3. For different types of barrier options, we have to modify the computational algorithm reported in this section.

REFERENCES

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *18*, 637–659.

Boyle, P. (1988). A Lattice Framework for Option Pricing with Two State Variables. *Journal of Financial and Quantitative Analysis*, *23*, 1–12.

Boyle, P., & Lau, S. H. (1994). Bumping Up Against the Barrier with the Binominal Model. *Journal of Derivatives*, *Summer*, 6–14.

Cheuk, T. H. F., & Vorst, T. C. F. (1996). Complex Barrier Options. *Journal of Derivatives*, *Fall*, 8–22.

Cheuk, T. H. F., & Vorst, T. C. F. (1997). Currency Lookback Options and Observation Frequency: A Binominal Approach. *Journal of International Money and Finance*, *2*, 173–187.

Chriss, N. A. (1997). *Black-Scholes and Beyond: Option Pricing Models*. McGraw-Hill, New York.

Cox, J., Ross, S., & Rubinstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics*, *7*, 229–264.

Conze, A., & Viswanathan, R. (1991). Path Dependent Options: The Case of Lookback Options. *Journal of Finance*, *46*, 1893–1907.

Goldman, B. W. M., Sosin, H. B., & Gatto, M. A. (1979). Path Dependent Options: Buy at the Low, Sell at the High. *Journal of Finance*, *34*, 401–414.

Hull, J., & White, A. (1993). Efficient Procedures for Valuing European and American Path-Dependent options. *Journal of Derivatives*, *Fall*, 21–31.

Merton, R. C. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, *4*, 141–183.

Omberg, E. (1988). Efficient Discrete Time Jump Process Models in Option Pricing. *Journal of Financial and Quantitative Analysis*, *23*, 161–174.

Pelsser, A., & Vorst, T. C. F. (1994). The Binominal Model and the Greeks. *Journal of Derivatives*, *Spring*, 45–49.

Ritchken, P. (1995). On Pricing Barrier Options. *Journal of Derivatives*, *Winter*, 19–28.

Rubinstein, M., & Reiner, E. Exotic Options, Working Paper, University of California at Berkeley.

Tian, Y. (1993). A Modified Lattice Approach to Option Pricing. *Journal of Futures Markets*, *13*, 563–577.

This Page Intentionally Left Blank

THE INFORMATION ROLE OF PORTFOLIO DEPOSITORY RECEIPTS

Paul Brockman and Yiuman Tse

ABSTRACT

The purpose of this paper is to investigate the information role of portfolio depository receipts (PDRs) by using the common factor models of Gonzalo and Granger (1995) and King et al. (KPSW) (1991). PDRs are exchange-traded securities representing a claim on an underlying basket of stocks held by an investment trust. This study shows that the stock index, stock index futures, and PDRs are cointegrated with a common stochastic trend. The common economic/market factor is predominantly derived from the futures market and that PDRs appear to be the least important in the context of price discovery. The results are consistent with the view that PDRs are designed for discretionary liquidity traders while futures are better suited for informed traders. These findings are also useful to hedgers and arbitrageurs and have implications for price discovery and optimal security design.

INTRODUCTION

Recent advances in financial economic theory and practice clearly demonstrate the benefits of constructing and trading broad-based stock market portfolios. To meet this demand, several cash index alternatives (e.g. index funds, stock index futures, stock index options, and index participations (IPs)) have been designed and tested in the marketplace. Few financial innovations have proven more

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 39–56.

ISBN: 0-7623-0798-6

successful in terms of trading volume, liquidity, or arbitrage and hedging opportunities than the creation of these stock index alternatives. More recently, another stock index derivative known as a portfolio depository receipt (PDR) has been successfully introduced onto U.S. and Canadian stock exchanges.[1] While considerable research has focused on the information roles of more traditional derivatives (e.g. futures and options), little is known of the information role played by PDRs. The purpose of this paper is to investigate the information role of these increasingly-popular instruments using the common factor techniques of Gonzalo and Granger (1995) and King et al. (KPSW) (1991).

More specifically, this study ranks the stock index futures and PDRs in terms of their significance to the price discovery process and shows that this ranking is consistent with the security design preferences of informed versus uninformed traders. To a large degree, identifying the instrument that leads the others in terms of price discovery is synonymous with identifying the instrument preferred by informed traders.

To the extent that derivatives play an active role in the price discovery process, asset prices will more accurately and rapidly reflect changing production costs and consumption utilities. Derivatives add to the operational efficiency of risk allocation by lowering the transaction costs of trading an equivalent bundle of underlying assets. As noted by Manaster and Rendleman (1982, p. 1044), "In the long run, the trading vehicle that provides the greatest liquidity, the lowest trading costs, and the least restrictions is likely to play the predominant role in the market's determination of equilibrium values of the underlying stocks".

To date, most research on derivatives has focused on futures and options due to their trading success and subsequent data availability. Although many stylized facts have been generated about these two basic derivatives, much less is known about the information role of newer instruments such as PDRs. Ongoing financial innovation has led to an increasing number of derivative securities that trade simultaneously on the same underlying asset(s). The price discovery role of these new derivative securities is an empirical issue that can be addressed only through ongoing financial research. The results of such research are useful in identifying the underlying characteristics of successful derivative securities design.

The results of this study add support to the notion of liquidity trader preference for baskets of securities such as PDRs. We show that one type of composite security (i.e. the futures) is better designed for the informed trader while another type of composite (i.e. the PDR) is better designed for the discretionary liquidity trader. Informed traders in possession of private information will tend to trade in the futures instrument, thus making it the dominant price-discovery market. PDRs, on the other hand, appeal primarily to discretionary liquidity traders seeking refuge from potential losses to informed traders. This analysis is built on previous theoretical work by Admati and Pfleiderer (1988), Bhushan (1991) and Chowdry and Nanda (1991). These studies show that discretionary liquidity traders will tend to concentrate trading in specific time periods, specific assets, or on specific markets, respectively, in order to reduce the probability of transacting with informed traders. The results of our study are consistent with similar trading preferences with respect to alternative derivative designs (i.e. futures versus PDRs).

This study contributes to the literature by analyzing the pricing behavior of PDRs in order to determine the information role they play in the market. Similar to futures and options lead-lag studies, this paper asks the question: Are PDRs "active" instruments in the aggregation of market-wide information or are they "passive" followers of underlying index prices?[2] Because the underlying stock index and its derivative securities complex are driven by the same underlying factor(s), all of the related return series are expected to display positive covariation in the long run (i.e. cointegration). In the short run, however, different securities may impound information at different rates due to design characteristics favored by informed traders.

Using the maximum likelihood estimation technique of Johansen (1988, 1991), we show that the stock index, stock index futures, and PDRs are part of a cointegrated system with two cointegrating vectors. Subsequent results, based on Gonzalo and Granger (1995), demonstrate that the common economic/market factor is predominantly derived from the futures market. Next, variance decomposition (VDC) and impulse response function (IRF) analysis is used to explain the price movements in each of the three instruments due to movements in the common factor. This analysis also measures the amount and speed with which innovations to the common factor are transmitted to each individual security. The results show that futures price movements closely track innovations to the common factor and that information travels rapidly among the underlying index and its derivative securities.

The remainder of the paper is designed as follows. Section 2 contains a brief overview of the related literatures, including security design and previous lead-lag empirical results. Section 3 describes the data set after providing some background information on the creation and structure of PDRs. Section 4 discusses the method of analysis and explains the empirical findings. Section 5 summarizes and concludes the paper.

RELATED RESEARCH

Security Design and Informed versus Uninformed Trading

Boot and Thakor (1993) claim that it may be optimal to partition the cash flows of an asset into distinct financial claims.[3] They show that dividing a security into an "informationally sensitive" component and an "informationally insensitive" component makes informed trading more profitable. In other words, by isolating the informationally sensitive component as a separately traded security, informed traders have more to gain by collecting and revealing private information.

Recent theoretical work by Subrahmanyam (1991) and Gorton and Pennacchi (1993) suggests that the creation of security "baskets" improves uninformed investor welfare and, therefore, does not represent a redundant security. Subrahmanyam (1991) shows that liquidity traders will prefer baskets of securities over individual securities because the security-specific components of adverse selection tend to be diversified away. He goes on to show that (p. 19), "the tendency for the basket to serve as the lowest transaction cost market for discretionary liquidity traders is strong even when there exists traders who possess private information about systematic factors".

Gorton and Pennacchi (1993) agree that the trading of security baskets is beneficial to uninformed traders. They argue that because the expected profits of informed traders are an increasing function of security return variance (see Kyle's (1985) model), the creation of a composite security with a lower return variance reduces the incentives for informed trading. If informed and uninformed traders prefer different trading vehicles, then the design of the security (e.g. futures contract, option, composite security) may underlie the previously-documented lead-lag results discussed below.

Spot-Futures Relationships

Numerous studies have examined the lead-lag relationship between the (spot) stock index and its related futures. Ng (1987), for example, finds that futures returns generally lead spot returns for a variety of futures contracts including stock indices, commodities, and currencies. Herbst, McCormack and West (1987) use tick-by-tick data to demonstrate a lead-lag relationship between the Standard and Poor (S&P) 500 Index futures, the Kansas City Board of Trade Value Line Index (VL) futures, and their respective spot indices. Kawaller et al. (1987) show that, although S&P 500 Index cash returns sometimes lead futures for approximately one minute, the S&P 500 futures returns lead spot market returns by 20 to 45 minutes. Stoll and Whaley (1990) report similar results for the Major Market Index (MMI) and the S&P 500 Index, although their futures lead is only about five minutes and they find stronger evidence of spot returns leading futures.[4]

Wahab and Lashgari (1993) use a cointegration approach to examine the lead-lag relationship between the S&P 500 and the Financial Times Stock Exchange (FTSE) 100 futures and cash markets. Similar to the findings of Chan, Chan and Karolyi (1991), Wahab and Lashgari find a significant information role for both the spot and futures indices. Additional (international) evidence of futures leading spot is provided by Puttonen (1993), using a vector error correction (VEC) model, as well as by Grunbichler et al. (1994) and Abhyankar (1995).[5]

Spot-Options Relationships

Manaster and Rendleman (1982) conducted one of the first studies testing for the lead-lag relationship between stocks and options. The authors find that closing options prices contain superior information to that of closing stock prices and that it takes up to one day of trading for the stock prices to adjust. Bhattacharya (1987), Anthony (1988) and Finucane (1991) present evidence consistent with Manaster and Rendleman's findings.

Stephan and Whaley (1990) provide the first study positing that stock price changes lead option price changes. After correcting for various shortcomings in the previous studies, they find that stock price and volume changes lead option price and volume changes. Chan, Chung and Johnson (CCJ) (1993), however, have challenged these conclusions (see Diltz & Kim (1996) for recent evidence that reconciles these opposing views). After confirming Stephan and Whaley's results, CCJ show that the stock lead disappears when bid-ask midpoints are used instead of transaction prices.[6]

PDR BACKGROUND AND DATA

Portfolio Depository Receipts (PDRs)

PDRs began trading in Canada and the U.S. shortly after the demise of index participations (IPs). IPs are zero-net-supply positions, similar to a short sell transaction, which grant long holders the right to cash out their positions at an amount equal to the value of the underlying stock portfolio. Unlike short sells, however, buyers of IPs have the option to take delivery of the underlying portfolio only at specific times in the future. These cash index alternatives, pioneered by the Philadelphia Stock Exchange (PHLX) and AMEX, began trading on May 12, 1989 under the jurisdiction of the Securities and Exchange Commission (SEC).[7] However, on August 18, 1989, the U.S. Court of Appeals of the Seventh Circuit ruled that IPs belonged under the jurisdiction of the Commodity Futures Trading Commission (CFTC) instead of the SEC and investors were forced to close out their positions.

PDRs were created in an effort to overcome some of the legal and other design shortcomings of IPs. PDRs are best thought of as warehouse receipts for a pre-specified basket of stocks left on deposit with a unit investment trust. Investors can create or redeem the receipts by tendering or taking delivery of the underlying stock portfolio, and the receipts are traded on organized exchanges in the same manner as individual stocks. PDRs can be bought on margin and sold short while dividends on the underlying portfolio are accumulated and distributed quarterly by the trust. PDRs have certain advantages over stock index funds and stock index futures. Unlike index funds, the exchange-traded feature of PDRs improves liquidity and the timeliness of pricing information. And unlike futures, PDRs track the total returns (i.e. dividends included) on the underlying portfolio, do not have to be rolled-over at maturity, do not have daily price limits, and can be purchased in small increments.

We posit that PDRs serve the purpose of securing or separating the benefits of basket trading for uninformed traders, while futures contracts serve the same purpose for informed traders. PDRs, therefore, are seen as the informationally-insensitive derivative while futures are the informationally-sensitive derivative. Informed traders will benefit most from the derivative design that maximizes leverage and minimizes transaction costs (i.e, futures).[8] PDRs may play a significant allocation role due to discretionary liquidity trading demand, but are not expected to play a dominant information role. Subsequent empirical results, reported below, confirm this conceptual framework.

Description of Data

This study uses daily data on the PDR with the longest trading history, the Toronto Stock Exchange's (TSE's) TIP instrument, along with the TSE 35 Index and futures contract.[9] Each TIP unit represents an interest in a trust that holds the TSE 35 Index portfolio of stocks. The TSE 35 Index is composed of a (value-weighted) cross-section of Canada's largest and most liquid corporations. The value of each unit is approximately one-tenth the Index level, and a round lot consists of 100 units. Investors holding a predetermined number of units (i.e. approximately equal to one basket of the Index or 51,244 units as of

1992), or any multiples thereof, can redeem their TIPs for the underlying basket of stocks at any time. Investors holding this same predetermined number of TIPs can vote the shares by proxy, while smaller holdings are voted by the Discretions Committee.[10] Dividends paid on underlying shares are collected by a trust (Montreal Trust Company of Canada) and distributed to TIPs holders on a quarterly basis (April, July, October, and December). TIPs, as well as the underlying stocks making up the Index, are traded from 9:30 a.m. to 4:00 p.m. (Eastern), and the value of the Index is updated every 15 seconds.

The Toronto 35 Index futures contract is valued at \$500 times the futures price and is traded on the Toronto Futures Exchange from 9:15 a.m. to 4:15 p.m. (Eastern).[11] The contract months include the two nearest months and the next two quarterly months from the March, June, September, December cycle. Minimum price fluctuations are \$10 per contract and all contracts are cash settled. The last trading day is the Thursday before the third Friday of the contract month, while settlement occurs on the third Friday at the official opening level of the Index.

METHOD OF ANALYSIS AND EMPIRICAL RESULTS

Cointegration and Common Stochastic Trend

Since the TSE 35 Index, futures, and PDR instruments are likely to be driven by some common economic/market factors, the three series are expected to move together in the long run. Cointegration, referring to the existence of a stationary relationship among nonstationary series, is used to analyze the information transmission between these three closely related instruments. See Hasbrouck (1995) for detailed analysis of the relationship between the common stochastic trend, VECM, and price discovery.

Johansen (1988, 1991) develops the maximum likelihood estimator for a cointegrated system.[12] Let X_t be an $n \times 1$ vector of the indices; $n =$ the number of variables in the system (i.e. three in this case). If X_t is cointegrated, then it can be represented by a vector error correction model (VECM):

$$\Delta X_t = \mu + \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \varepsilon_t \quad (1)$$

where μ is a 3×1 vector of drifts, Π and Γ's are 3×3 matrices of parameters, and ε_t is a 3×1 white noise vector. The long-run relationship matrix, Π, has reduced rank of $r < 3$ and can be decomposed as $\Pi = \alpha\beta'$, where α and β are $n \times r$ matrices. The β matrix consists of the cointegrating vectors and α is the error correction (or equilibrium adjustment) matrix. The Johansen trace test statistic of the null hypothesis that there are at most r cointegrating vectors $0 \leq r \leq n$, and thus $m = n - r$ common stochastic trends between series is

$$trace = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i), \tag{2}$$

where the λ_i's are the $n - r$ smallest squared canonical correlations of X_{t-1} with respect to ΔX_t corrected for lagged differences and T is the sample size actually used for estimation. The Johansen maximum eigenvalue (λ_{max}) test statistic of the null hypothesis that there are r cointegrating vectors against $r + 1$ is

$$\lambda_{max} = -T \ln(1 - \hat{\lambda}_{r+1}), \tag{3}$$

where λ_{r+1} is the $(r + 1)$th greatest squared canonical correlation.

X_t can be represented by the following common factor model

$$X_t = \theta f_t + \tilde{X}_t \tag{4}$$

where f_t is an $m \times 1$ vector of I(1) common stochastic trends, θ is a $n \times m$ loading matrix, and X_t is an $n \times 1$ vector of I(0) transitory components. The common factor f_t is given by

$$f_t = \alpha_{\perp'} X_t \tag{5}$$

$$\theta = \beta_{\perp}(\alpha_{\perp'} \beta_{\perp})^{-1} \tag{6}$$

where $\alpha_{\perp}$ is a $n \times m$ matrix of full rank and orthogonal to α and β, respectively. Based on the framework of Johansen (1991), and particularly on the duality between the stationary relations and the common stochastic trends, Gonzalo and Granger (1995) develop the maximum likelihood estimator of $\alpha_{\perp}$. The likelihood ratio test of the null hypothesis H_o: $\alpha_{\perp} = A\varphi$, where A is an arbitrary $n \times q$ restriction matrix and φ is the $q \times m$ matrix of the reduced parameter space, is

$$-2 \ln Q = -T \sum_{i=r+1}^{n} \ln\{(1 - \hat{\lambda}^*_{i+(q-n)})/(1 - \hat{\lambda}_i)\}, \tag{7}$$

where λ^*_i is the ith largest eigenvalue from the model under the null hypothesis and is distributed as $\chi^2(m \times (n - q))$. See Gonzalo and Granger (1995) for detailed description. The test statistic (6), under the null that x_i is the common factor, is distributed as $\chi^2(2)$ if $n = 3$ and $m = 1$, i.e. one common stochastic trend in the three-variable system (which is the case in the current study).

Variance Decomposition and Impulse Response Functions

While Gonzalo and Granger's method tests the hypothesis that the common factor is derived from a particular market, it is also important to know how movement in one market can be explained by innovations in the common factor and how rapidly the innovations are transmitted to different markets. To address these two issues, the current paper employs the KPSW (1991) approach of variance decomposition (VDC) and impulse response functions (IRFs) to explore the single common stochastic trend in the system. KPSW identify the shocks with permanent effects on each series independently from those with only transitory effects. It is worth pointing out that the shock is an innovation to the common factor, instead of each individual series as is commonly done in the conventional VAR literature put forth by Sims (1980). The KPSW model has been used by Fisher, Fackler and Orden (1995) and Tse, Lee and Booth (1996), among others.

The VECM (1) is transformed into a vector moving average (VMA) model

$$\Delta X_t = \mu + C(B)\varepsilon_t, \tag{8}$$

where $C(B)$ is a 3×3 matrix polynomial in B. Since there is only one common factor in X_t, $C(1)$ is of rank 1 and there exists a 3×1 vector G such that $C(1) = JG'$, where J is $(1, 1, 1)'$. The VMA (7) is a reduced-form relation and is converted to a structural model

$$\Delta X_t = \mu + \Psi(B)\xi_t, \tag{9}$$

where $\Psi(B) = C(B)\Psi_o$ and $\xi_t \equiv (\xi_{1t}\ \xi_{2t}\ \xi_{3t})' = \Psi_o^{-1}\varepsilon_t$ is a 3×1 vector of serially uncorrected structural disturbances. To identify the common factor, KPSW restrict the matrix of long-run multipliers $\Psi(1)$ as

$$\Psi(1) = [J|0]. \tag{10}$$

According to (9), ξ_{1t} is the persistent shock (or shock to the common factor) with the long-run multiplier J, while ξ_{2t} and ξ_{3t} are transitory shocks. The impulse responses associated with ξ_{1t} are given by the first column of $\Psi(B)$.

Empirical Results

Table 1 summaries the descriptive statistics for spot, futures, and PDRs over the period from January 1991 to December 1994 (1008 observations). It shows that the three instruments have very similar variance, skewness, and kurtosis, and that they are highly correlated with one other. The ADF (Dickey & Fuller, 1979, 1981) and Phillips-Perron (Phillips (1987) and Phillips and Perron (1988)) unit root tests reported in Table 2 indicate that the three price series can be characterized as I(1) processes.

As mentioned above, since the three instruments are close substitutes and are driven by some common economic/market factor(s), they should move together

Table 1. Sample Statistics of Log Returns.

	ΔSpot	ΔFutures	ΔPDR
Mean (e-4)	2.12	2.14	2.13
	(0.74)	(0.35)	(0.36)
Median	0.00	0.00	0.00
Variance (e-5)	4.16	5.37	5.39
Skewness	–0.16	–0.10	–0.26
	(0.04)	(0.18)	(0.001)
Excess Kurtosis	1.44	0.96	1.26
	(<0.001)	(<0.001)	(<0.001)
Correlation Coefficients			
ΔSpot		0.91	0.88
ΔFutures			0.91

Asymptotic *p*-values are contained in parentheses.

in the long run. The Johansen tests presented in Table 3 demonstrate that the three instruments are cointegrated with two cointegrating vectors, i.e. $r=2$, suggesting that there is one common stochastic trend.[13] Results are robust to the number of lags k used in the VECM. The normalized coefficient vector of the common factor, $\alpha_{\perp}$, is (1.0 3.08–1.62)′ as shown in Table 4. Thus, this common factor is mostly derived from the futures market. The Gonzalo and Granger (1995) model is used to determine which market (instrument) is dominant in the cointegration system. The null hypothesis that the futures market is the

Table 2. Augmented Dickey-Fuller and Phillips-Perron Unit Root Tests.

	Spot	Futures	PDR
ADF (no trend)	–1.54	–1.53	–1.60
ADF (with trend)	–2.39	–2.15	–2.41
Phillips-Perron (no trend)	–1.48	–1.47	–1.54
Phillips-Perron (with trend)	–2.34	–2.36	–2.38

Note: The statistics are computed with 5 lags for the ADF tests and 5 non-zero autocovariances in Newey-West (1987) correction for the Phillips-Perron tests. Similar results are given for higher lags up to 20. The critical values for both statistics, which are asymptotically equivalent, are available in Fuller (1976, p. 373). The 5% critical values are − 2.86 (*no trend*) *and* –3.41 (*with trend*).

Table 3. Johansen Cointegration Tests.

	Trace	Critical Values at the 1% Level	λ_{max}	Critical Values at the 1% Level
$r=2$	2.36	11.65	2.36	11.65
$r=1$	48.97	23.52	46.60	19.19
$r=0$	139.90	37.22	90.95	25.75

Cointegration Vectors, β

	Spot	Futures	PDR
β_1'	1.000	–1.198	0.187
β_2'	1.000	0.331	–1.351

Results reported for $k=6$ (i.e. 5 lags of ΔX_t in the VECM) are qualitatively the same for $k=11$. The critical values are obtained from Osterwald-Lenum (1992).

common factor is not rejected (with p-value = 0.286), while the spot and PDRs markets are rejected at the five% significance level. Hence, the futures market can be considered the common factor driving the other two markets. These results are consistent with the VECM results in Table 5. It shows that both of the error correction terms, $z_{1,t-1}$ and $z_{2,t-1}$, are insignificant in the $\Delta\text{Futures}_t$ equation, but one of them is significant in the ΔSpot_t and ΔPDR_t equations. Therefore, the futures prices do not respond to deviations from the spot and PDR prices, but the spot and PDR prices do respond to deviations from the futures prices, suggesting that the futures market contributes more than the other two markets in the price discovery process. Harris et al. (1995) use this approach to examine price discovery between the NYSE and regional exchanges.[14] To further explore the above results, the IRF and VDC approach

Table 4. Common Stochastic Trend, $\alpha_\perp$.

	Spot	Futures	PDR
$\alpha_\perp'$	1.000	3.082	–1.615
H_o: x_i is the common factor[a]	9.034	2.505	13.443
p-value	0.011	0.286	0.001

[a] The test statistic is derived by Gonzalo and Granger (1995) and distributed as $\chi^2(2)$.

of KPSW are employed. Table 6 demonstrates that at the end of a 30-day horizon, 86% of the forecast error variance in ΔSpot, 93% in ΔFutures, and 68% in ΔPDR can be attributed to innovations to the common stochastic trend, ξ_{1t}. The impulse responses of X_t to an innovation to the common stochastic trend are reported in Table 7. In response to a shock generated at day 0, all markets respond significantly (greater than 90%) within day 1, indicating that the information transmission between markets is very quick and takes less than one day. However, at day 0, the spot market responds 76%, the futures market 91%, and the PDR 76%. This implies that the movement of the futures market is almost the same as the common factor. Note that the long-run (or 30 days) multiplier of the permanent shock is unity for each market.

Therefore, both the Gonzalo and Granger (1995) and KPSW (1991) models imply that the common factor is mainly derived from the futures market and

Table 5. Results of VECM.

$$\Delta X_t = a + b_1 z_{1,t-1} + b_2 z_{2,t-2} + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_5 \Delta X_{t-5} + \varepsilon_\tau$$
$$z_{i,t-1} = \beta_i' X_{t-1}, \quad i = 1 \textit{ or } 2$$

	Spot		Futures		PDR	
constant	–0.354*	(–2.29)	–0.215	(–1.21)	–0.631**	(–3.64)
$z_{1,t-1}$	–0.246**	(–2.62)	0.073	(0.08)	–0.014	(–0.13)
$z_{2,t-1}$	0.078	(1.59)	0.084	(1.48)	0.209**	(3.77)
$\Delta x_{1,t-1}$	0.025	(0.21)	0.275	(1.95)	–0.046*	(–2.13)
$\Delta x_{2,t-1}$	–0.015	(–0.11)	–0.183	(–1.22)	0.235	(1.60)
$\Delta x_{3,t-1}$	0.141	(1.48)	0.025	(0.22)	–0.368**	(–3.42)
$\Delta x_{1,t-2}$	0.014	(0.12)	0.151	(1.11)	0.139	(0.05)
$\Delta x_{2,t-2}$	–0.215	(–1.67)	–0.278	(–1.87)	–0.045	(–0.31)
$\Delta x_{3,t-2}$	0.222*	(2.20)	0.125	(1.07)	–0.091	(–0.80)
$\Delta x_{1,t-3}$	–0.028	(–0.25)	0.043	(0.33)	0.053	(0.42)
$\Delta x_{2,t-3}$	–0.073	(–0.59)	–0.081	(–0.57)	–0.016	(–0.12)
$\Delta x_{3,t-3}$	0.115	(1.14)	0.028	(0.24)	–0.028	(–0.25)
$\Delta x_{1,t-4}$	–0.155	(–1.49)	–0.061	(–0.50)	–0.064	(–0.55)
$\Delta x_{2,t-4}$	0.018	(0.16)	–0.019	(–0.14)	–0.005	(–0.04)
$\Delta x_{3,t-4}$	0.109	(1.15)	0.053	(0.48)	0.037	(0.34)
$\Delta x_{1,t-5}$	0.004	(0.04)	0.046	(0.45)	0.064	(0.65)
$\Delta x_{2,t-5}$	–0.048	(–0.51)	–0.027	(–0.24)	–0.093	(–0.88)
$\Delta x_{3,t-5}$	0.035	(0.43)	–0.033	(–0.35)	0.027	(0.29)

The VECM is estimated by the maximum likelihood approach of Johansen (1988, 1991).
t-statistics are in parentheses.
* significant at the 5% level.
** significant at the 1% level.

Table 6. Forecast Error Variance Decomposition.

Horizon	Spot	Future	PDR
1	0.8750	0.9496	0.7014
2	0.8755	0.9372	0.6832
3	0.8749	0.9350	0.6828
4	0.8741	0.9346	0.6829
5	0.8702	0.9339	0.6817
10	0.8656	0.9314	0.6780
30	0.8646	0.9306	0.6766

Entries are the fractions of forecast error variance to forecast ΔX_t that are due to the shocks to the common stochastic trend.

that the futures market drives the information transmission mechanism between markets, assuming that the common factor impounds all the long-run information. The IRFs also indicate the rapid (less than one day) information transmission between markets. These results are consistent with the view that informed traders prefer the security design of the futures instrument over the PDR instrument. Leverage and transaction costs are likely reasons for these preferences. The PDRs do not display a dominant price discovery role over the underlying stock index. This result is consistent with the view that PDRs are designed primarily for discretionary traders who desire the diversification benefits of basket trading but also want to minimize the probability of transacting with informed traders.

Table 7. Responses to Innovation to the Common Factor.

l	Spot	Future	PDR
0	0.7579	0.9129	0.7632
1	0.9215	0.9616	0.9279
2	0.9450	0.9525	0.9199
3	0.9802	0.9724	0.9409
4	0.9819	0.9710	0.9373
5	0.9645	0.9615	0.9294
10	0.9830	0.9704	0.9534
30	0.9995	0.9871	0.9804

Entries are the cumulative impulse responses of x_{it} to the shocks to the common stochastic trend, $\partial x_i/\partial \xi_{1,t-1}$, where ξ_{1t} is the shock to the stochastic trend.

CONCLUSION

This study examines the information role of a relatively new derivative security known as portfolio depository receipts (PDRs). PDRs are similar to warehouse receipts issued on an underlying basket of stocks and are currently traded on the Toronto Stock Exchange (TSE) and American Stock Exchange (AMEX). While considerable research has focused on the information roles of more traditional derivatives (e.g. futures and options), little is known of the price discovery role of PDRs. This paper investigates the lead-lag relationships among PDRs, the stock index, and stock index futures, as well as analyzes the underlying causes for such relationships. The results are useful to hedgers and arbitrageurs, and have implications for price discovery and security design.

It is posited that PDRs are designed primarily to serve the purpose of separating the benefits of basket trading for discretionary liquidity traders and, therefore, are not expected to display a price leadership role. PDRs are informationally-insensitive derivatives, while futures fulfil the role of an informationally-sensitive derivative. Empirical results, based on the common factor models of Gonzalo and Granger (1995) and KPSW (1991), add empirical support to this hypothesis. The results reported herein are consistent with Admati and Pfleiderer (1988), Bhushan (1991), and Chowdry and Nanda (1991). These studies show that discretionary traders will tend to concentrate their trading in specific time periods, specific assets, or on specific markets, respectively, in order to reduce the probability of transacting with informed traders. Our findings suggests that uninformed investors will also tend to concentrate trading in particular designs of derivative securities (i.e. PDRs).

NOTES

1. PDRs are exchange-traded securities representing a claim to an underlying basket of stocks held by an investment trust. The first PDR, known as Toronto 35 Index Participation units (TIPs), began trading on the Toronto Stock Exchange (TSE) on March 9, 1990. In 1993, the American Stock Exchange (AMEX) introduced the Standard & Poor's Depositary Receipt (SPDR). Both instruments became one of the most actively traded issues on its respective exchange shortly after its introduction. Due to this early success, the TSE and AMEX have recently launched additional PDRs known as Toronto 100 Index Participation units (HIPs) and Standard & Poor's MidCap 400 Depositary Receipts (MidCap SPDRs), respectively.

2. Boot and Thakor (1993) use the terms "informationally sensitive" and "informationally insensitive" to distinguish securities that actively impound information into prices from those that are passive observers of the pricing process.

3. The security design literature traces its roots to Modigliani and Miller's (1958) claim that capital structure, and security design in general, is irrelevant to firm value

under the assumptions of market completeness and no short sale restrictions. Between 1958 and 1988, that is, (Allen & Gale, 1988), little work had focused on optimal design characteristics since most researchers had taken the types of securities issued by firms as exogenously determined. Allen and Gale assume incomplete markets and short sale restrictions and construct optimal securities in which the earnings in each state of nature are totally allocated to the investor group that values them most. Additional design-related research can be found in Madan and Soubra (1991), Harris and Raviv (1989,1995), Duffie and Jackson (1989), Zender (1991), Cuny (1993), and Nachman and Noe (1994).

4. Additional corroborating evidence is offered by Kutner and Sweeney (1991) and Chan (1992). Kutner and Sweeney use minute-by-minute returns on S&P 500 Index cash and futures markets and confirm the lead of futures over spot, with futures leading spot by eight to 20 minutes and spot leading futures by one to five minutes. Chan shows that S&P 500 and the MMI lead-lag results are fairly robust to market conditions. Laatsch and Schwarz (1988), however, claim that the lead-lag relationship for the MMI between 1985–1988 is not stable across time periods (e.g. intra-day periods).

5. While the discussion above is related to lead-lag relationships in returns, several studies have also investigated lead-lag relationships in volatility. See, for example, Kawaller et al. (1990), Cheung and Ng (1991), Chan et al. (1991), and Abhyankar (1995).

6. CCJ (1993) hypothesize that Stephan and Whaley's results can be explained by the infrequent trading of the options. If this is the case, then small stock price changes may not be immediately reflected in the option due to its tick size. In other words, information releases may induce stock price changes which barely clear the one-eighth tick price barrier. The related (theoretical) option price change may not be large enough to clear the same barrier, thus making it unprofitable to trade. Under this scenario, the option will trade only after the stock has made more that one price change in the same direction.

7. The PHLX and AMEX products are referred to as cash index participations (CIPs) and equity index participations (EIPs), respectively. Similar instruments were also introduced on the Chicago Board Options Exchange (CBOE) and the New York Stock Exchange (NYSE). In general, the CBOE's market basket security (MBS) and NYSE's exchange stock portfolio (ESP) allow investors to trade portfolios of securities (e.g. the S&P 500 Index) in a manner similar to individual stocks.

8. In Canada, as in the U.S., futures market transaction costs are generally lower than stock market transaction costs for equivalent dollar amounts of underlying assets. In addition, while PDRs can be purchased on margin, the leverage capacity of futures contracts is considerably higher.

9. For additional information on the TSE 35 futures contracts and TIPs securities see Park and Switzer (1995).

10. According to TSE publications, the Discretions Committee is composed of senior TSE staff, non-member public governors, and independent non-member, non-staff individuals.

11. The different closing times for the Index and TIPs (4:00 p.m.) versus the futures contract (4:15 p.m.) creates a potential problem of nonsimultaneous prices. It should be noted, however, that the TIPs and Index results are not subject to nonsynchronous closing times, and that the leading role of futures is consistent with previous research.

12. Johansen tests are robust to various non-normal distributions as shown by Cheung and Lai (1993) and Lee and Tse (1996).

13. As shown by the last two rows of Table 3, the two cointegrating vectors are very close to the spread/differential between spot and futures prices and between spot and PDRs. Restricting the cointegrating vectors (as these spreads) does not qualitatively change the results.

14. Many papers have also used the error correction terms to describe the equilibrium dynamics between markets, e.g. Arshanapalli and Doukas (1993) and Arshanapalli, Doukas and Lang (1995).

REFERENCES

Abhyankar, A. H. "Return and Volatility Dynamics in the FT-SE 100 Stock Index Futures Markets". *Journal of Futures Markets*, 15 (1995), 457–488.

Admati, A. R., and P. Pfleiderer. "A Theory of Intraday Patterns: Volume and Price Variability". *Review of Financial Studies*, 1 (1988), 3–40.

Allen, F., and D. Gale. "Optimal Security Design". *Review of Financial Studies*, 1 (1988), 229–264.

Anthony, J. (1988). The Interrelationship of Stock and Option Trading Volume Data. *Journal of Finance*, *43*, 949–964.

Arshanapalli, B., & Doukas, J. (1993) International Stock Market Linkages: Evidence from the Pre- and Post-October 1987 Period. *Journal of Banking and Finance*, *17*, 193–208.

Arshanapalli, B., Doukas, J., & Lang, L. H. P. (1995). Pre- and Post-October 1987 Stock Market Linkages between U.S. and Asian Markets. *Pacific-Basin Finance Journal*, *3*, 57–73.

Bhattacharya, M. (1987). Price Changes of Related Securities: The Case of Call Options and Stocks. *Journal of Financial and Quantitative Analysis*, *22*, 1–15.

Bhushan, R. (1991). Trading Costs, Liquidity, and Asset Holdings. *Review of Financial Studies*, *4*, 343–360.

Boot, A. W. A., & Thakor, A. V. (1993). Security Design. *Journal of Finance*, *48*, 1349–1378.

Chan, K. (1992). A Further Analysis of the Lead-Lag Relationship between the Cash Market and the Index Futures Market. *Review of Financial Studies*, *5*, 123–152.

Chan, K., Chan, K. C., & Karolyi, G. A. (1991). Intraday Volatility in the Stock Index and Stock Index Futures Markets. *Review of Financial Studies*, *4*, 657–684.

Chan, K., Chung, Y. P., & Johnson, H. (1993). Why Option Prices Lag Stock Prices: A Trading-Based Explanation. *Journal of Finance*, *48*, 1957–1967.

Cheung, Y. W., & Lai, K. S. (1993). Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration. *Oxford Bulletin of Economics and Statistics*, *55*, 313–328.

Cheung, Y. W., & Ng, L. K. (1991). The Dynamics of S&P 500 Index and S&P Futures Intraday Price Volatilities. *Review of Futures Markets*, *9*, 121–154.

Chowdry, K. J., & Nanda, V. (1991). Multimarket Trading and Market Liquidity. *Review of Financial Studies*, *4*, 483–511.

Cuny, C. J. (1993). The Role of Liquidity in Futures Market Innovations. *Review of Financial Studies*, *6*, 57–78.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, *74*, 427–431.

Dickey, D. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, *49*, 1057–1072.

Diltz, D. J., & Kim, S. (1996). The Relationship between Stock and Option Price Changes. *Financial Review*, *31*, 499–519.

Duffie, D., & Jackson, M. O. (1989). Optimal Innovation of Futures Contracts. *Review of Financial Studies*, *3*, 275–296.

Finucane, T. (1991). Put-Call Parity and Expected Returns. *Journal of Financial and Quantitative Analysis*, *26*, 445–457.

Fisher, L. A., Fackler, P. L., & Orden, D. (1995). Long-Run Identifying Restrictions for an Error-Correction Model of New Zealand Money, Prices and Output. *Journal of International Money and Finance*, *14*, 127–147.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York: Wiley.

Gonzalo, J., & Granger, C. W. J. (1995). Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics*, *13*, 27–36.

Gorton, G. B., & Pennacchi, G. G. (1993). Security Baskets and Index-Linked Securities. *Journal of Business*, *66*, 1–27.

Grunbichler, A., Longstaff, F. A., & Schwartz, E. S. (1994). Electronic Screen Trading and the Transmission of Information: An Empirical Examination. *Journal of Financial Intermediation*, *3*, 166–187.

Harris, F. H. deB., McInish, T. H., Shoesmith, G. L., & Wood, R. A. (1995). Cointegration, Error Correction and Price Discovery on Informationally linked security markets. *Journal of Financial Quantitative Analysis*, *30*, 563–580.

Harris, M., & Raviv, A. (1989). The Design of Securites. *Journal of Financial Economics*, *24*, 255–287.

Harris, M., & Raviv, A. (1995). The Role of Games in Security Design. *Review of Financial Studies*, *8*, 327–367.

Hasbrouck, J. (1995). One Security, Many Markets: Determining the Contributions to Price Discovery. *Journal of Finance*, *50*, 1175–1199.

Herbst, A., McCormack, J., & West, E. (1987). Investigation of a Lead-Lag Relationship between Spot Indices and their Futures Contracts. *Journal of Futures Markets*, *7*, 373–382.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, *59*, 1551–1580.

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index. *Journal of Finance*, *42*, 1309–1329.

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1990). Intraday Relationships between Volatility in S&P 500 Futures Prices and the S&P 500 Index. *Journal of Banking and Finance*, *14*, 373–397.

King, R. G., Plosser, C. I., Stock, J. H., & Watson, M. W. (1991). Stochastic Trends and Economic Fluctuations. *American Economic Review*, *81*, 819–840.

Kutner, G. W., & Sweeney, R. J. (1991). Causality Tests Between the S&P 500 Cash and Future Markets. *Quarterly Journal of Business and Economics*, *30*, 51–74.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, *56*, 1315–1335.

Laatsch, F. E., & Schwarz, T. V. (1988). Price Discovery and Risk Transfer in Stock Index Cash and Futures Markets. *Review of Futures Markets*, *7*, 273–289.

Lee, T.-H., & Tse, Y. (1996). Cointegration Tests with Conditional Heteroskedasticity. *Journal of Econometrics*, *73*, 401–410.

Madan, D., & Soubra, B. (1991). Design and Marketing of Financial Products. *Review of Financial Studies*, *4*, 361–384.

Manaster, S., & Rendleman, R. (1982). Option Prices as Predictors of Equilibrium Stock Prices. *Journal of Finance*, *37*, 1035–1048.

Modigliani, F., & Miller, M. H. (1958). The Cost of Capital, Corporate Finance, and the Theory of Investment. *American Economic Review*, *48*, 261–297.

Nachman, D. C., & Noe, T. H. (1994). Optimal Design of Securities under Asymmetric Information. *Review of Financial Studies*, *7*, 1–44.

Newey, W. K., & West, K. D. (1987). A Simple Positive Semi-definite, Heteroskedasticity and Autorcorrection Consistent Covariance Matrix. *Econometrica*, *55*, 703–708.

Ng, N. (1987). Detecting Spot Forecasts in Futures Prices Using Causality Tests. *Review of Futures Markets*, *6*, 250–267.

Osterwald-Lenum, M. (1992). A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics*, *54*, 461–472.

Park, T. H., & Switzer, L. N. (1995). Index Participation Units and the Performance of Index Futures Markets: Evidence from the Toronto 35 Index Participation Units Market. *Journal of Futures Markets*, 187–200.

Phillips, P. C. B. (1987). Time Series Regression with a Unit Root in Time Series Regression. *Econometrica*, *55*, 277–301.

Phillips, P. C. B., & Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, *75*, 335–346.

Puttonen, V. (1993). Short Sales Restrictions and the Temporal Relationship between Stock Index Cash and Derivatives Markets. *Journal of Futures Markets*, *13*, 645–664.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*, 1–48.

Stephan, J., & Whaley, R. (1990). Intraday Price Change and Trading Volume Relations in the Stock and Stock Option Markets. *Journal of Finance*, *55*, 191–220.

Stoll, H. R., & Whaley, R. (1990). The Dynamics of Stock Index and Stock Index Futures Returns. *Journal of Financial and Quantitative Analysis*, *25*, 441–168.

Subrahmanyam, A. (1991). A Theory of Trading in Stock Index Futures. *Review of Financial Studies*, *4*, 17–51.

Tse, Y., Lee, T.-H., & Booth, G. G. (1996). The International Transmission of Information in Eurodollar Futures Markets: A Continuously Trading Market Hypothesis. *Journal of International Money and Finance*, *15*, 447–465.

Wahab, M., & Lashgari, M. (1993). Price Dynamics and Error Correction in Stock Index and Stock Index Futures Markets: A Cointegration Approach. *Journal of Futures Markets*, *13*, 711–742.

Zender, J. (1991). Optimal Financial Instruments. *Journal of Finance*, *46*, 1645–1663.

A DOUBLE SHARPE RATIO

Matthew R. Morey and H. D. Vinod

ABSTRACT

Sharpe's (1966) portfolio performance ratio, the ratio of the portfolio's expected return to its standard deviation, is a very well known tool for comparing portfolios. However, due to the presence of random denominators in the definition of the ratio, the sampling distribution of the Sharpe ratio is difficult to determine. Hence, measurement of estimation risk of the Sharpe ratio has often been ignored in practice. This paper uses the bootstrap methodology to suggest a new "Double" Sharpe ratio that allows an investor to make a tradeoff between risk-adjusted performance and estimation risk using the same weighting for estimation risk as the original Sharpe ratio uses for standard deviation. The use of this Double Sharpe ratio along with the original ratio, provides investors with a relatively simple method for measuring risk- and estimation risk-adjusted performance.

INTRODUCTION

The Sharpe (1966) portfolio performance ratio, the ratio of a portfolio's expected return to its standard deviation, is widely used and cited in the literature and pedagogy of finance. Indeed, in a recent finance literature search, over 30 papers published between 1995–1998 cited the Sharpe ratio.[1]

Despite its popularity, the Sharpe ratio suffers from a methodological limitation: because of the presence of random denominators in its definition and the difficulty in determining the sample size needed to achieve asymptotic

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 57–65.

ISBN: 0-7623-0798-6

normality, the Sharpe measure does not easily allow for evaluation of its own sampling distribution. As a result, assembling any notion of the estimation risk in the Sharpe index point estimate has been difficult.

We argue in this paper that the inability to easily construct a measure of estimation risk is a weakness of the Sharpe measure, for such information is valuable to investors. To illustrate, consider an investor who uses the Sharpe measure to evaluate portfolios. Such an investor will identify two different portfolios as having very similar performance if they have similar Sharpe point estimates. However, such an appraisal does not at all consider the range of uncertainty behind these point estimates; one portfolio may have little estimation risk on its Sharpe measure while the other may have a great deal of estimation risk. Such knowledge may sway an investor to prefer the portfolio with less estimation risk. Indeed not knowing the estimation risk of the Sharpe ratio is somewhat akin to not knowing the level of significance of the well-known Jensen or single index alpha model. Investors are trained in business schools to associate a positive and significant alpha with more reliable performance than a positive and non-significant alpha, but rarely is the same case made for the Sharpe ratio because of the above-described limitation.

This limitation of the Sharpe ratio has not gone unnoticed in the literature. In fact, several previous attempts have been made to construct methods to measure the estimation risk of the Sharpe ratio. The most well known is by Jobson and Korkie (1981) who use Taylor series (delta) approximations to derive confidence intervals on the Sharpe ratio. More recently work by Vinod and Morey (2000) has documented that a bootstrap methodology can achieve narrower, more reliable confidence intervals on the Sharpe ratio than the Jobson and Korkie method.

However, both of these methodologies still possess a shortcoming in that they provide no standard procedure of trading off performance for estimation risk. For example, consider two funds, one with a high Sharpe ratio and a very wide confidence interval, and another with a lower Sharpe ratio and yet a narrower confidence interval. Which should be preferred? What weighting should the estimation risk be given?

To deal with this issue we propose a modified version of the Sharpe ratio, which we call the Double Sharpe ratio, that, rather than calculating the confidence interval width, *directly* takes into account estimation risk in much the same fashion that the original Sharpe ratio takes into account the portfolio risk. In this way, estimation risk is implicitly in the performance measure with the same weighting as standard deviation receives in the original Sharpe ratio. As such, the Double Sharpe ratio allows investors to examine the risk- and estimation risk-adjusted performance in a relatively simple fashion.

The rest of the paper is organized as follows. Section 1 discusses the Sharpe measure and its methodological problems. Section 2 explains the Double Sharpe ratio. Section 3 explains how we empirically calculate the Double Sharpe Ratio. Section 4 provides a simple empirical example to help understand the use and value of the methodology. Section 5 concludes the paper with some final remarks and more on the practical use of the methodology.

1. THE SHARPE RATIO AND ESTIMATION RISK

Consider the following scenario in which the relative performance of n portfolios is to be evaluated.[2] In this scenario, $r_{i,t}$ represents the excess return from the i-th portfolio in period t, where $i = 1, 2, \ldots, n$. A random sample of T excess returns on the n portfolios is then illustrated by $r_t' = [r_{1t}, r_{2t}, \ldots r_{nt}]$, where $t = 1, 2, \ldots, T$ and where r_t is assumed to be multivariate normal random variable, with mean $\mu = \{\mu_i\}$, $i = 1, 2, \ldots, n$ and a covariance matrix $\sum = (\sigma_{ij})$ where $i, j = 1, 2, \ldots n$. It is well-known that the unbiased estimators of the $(n \times 1)$ mean vector and the $(n \times n)$ covariance matrix elements are,

$$\bar{r}_{1t} = \frac{1}{T}\sum_{t=1}^{T} r_{it}, \quad \text{and} \quad \mathrm{S} = \{\mathrm{s_{ij}}\} = \frac{1}{\mathrm{T}-1}\sum_{t=1}^{T} (r_{it} - \bar{r})(r_{jt} - \bar{r})'. \tag{1}$$

These two estimators are then used to form the estimators of the traditional Sharpe performance measure.

The population value of the Sharpe (1966) performance measure for portfolio i is defined as $Sh_i = \dfrac{\mu_i}{\sigma_i}$, $i = 1, 2, \ldots, n$. It is simply the mean excess return over the standard deviation of the excess returns for the portfolio. The conventional sample-based point estimates of the Sharpe performance measure used in (1) are then

$$\hat{S}h_i = \frac{\bar{r}_i}{s_i} \text{ for } i = 1, 2, \ldots, n. \tag{2}$$

Because of the presence of the random denominator s_i in the definition of (2), the Sharpe ratio does not permit an easy method for evaluating the estimation risk in the point estimate itself. This is because the small-sample distribution of the Sharpe measure is highly non-normal and hence the usual method for computing standard errors can be biased and unreliable.

2. THE DOUBLE SHARPE RATIO

The Sharpe ratio is defined in Eq. (2) as the ratio of the mean excess return to its standard deviation. We define the Double Sharpe ratio by

$$DSh_i = \frac{\hat{S}h_i}{s_i^{Sh}}, \tag{3}$$

where s_i^{sh} is the standard deviation of the Sharpe ratio estimate, or the estimation risk. As is clear in (3), the Double Sharpe penalizes a portfolio for higher estimation risk. One may question the weighting of the estimation risk; however, the implicit weighting of the estimation risk can be modified if desired without affecting our main proposal. For simplicity, our weighting is completely analogous to that implicit in the original Sharpe ratio.

For clarification, a couple of issues should be noted about the Double Sharpe ratio here before moving on. First, we, of course, assume the statistical paradigm is that the observed data are one realization from a random process. Hence, the observed means and variances are a random realization of the true mean and variance. Second, the Double Sharpe ratio deals only with estimation risk and not measurement risk. Estimation risk arises from the sampling variation in the estimates of mean and variance. This is distinctly different from measurement errors which arise from erroneous reports of prices, etc.

3. CALCULATING THE DOUBLE SHARPE RATIO

The question now is how to calculate the estimation risk, s_i^{sh}. To do this we utilize the well-known bootstrap methodology. To understand the basic ideas behind the bootstrap consider the following. Assume that a statistic, $\hat{\beta}$, is based on a sample of size, T. In the bootstrap methodology, instead of assuming the shape of the sampling distribution of $\hat{\beta}$, one empirically approximates the entire sampling distribution of $\hat{\beta}$ by investigating the variation of $\hat{\beta}$ over a large number of pseudo samples obtained by resampling. For the resampling, a Monte-Carlo type procedure is used on the available sample values. This is conducted by randomly drawing, with replacement, a large number of resamples of size T from the original sample. Each resample has T elements, however any given resample could have some of the original data points represented more than once and some not at all. Note that each element of the original sample has the same probability, $(1/T)$, of being in a sample.

The initial idea behind bootstrapping was that a relative frequency distribution of $\hat{\beta}$'s calculated from the resamples can be a good approximation to its sampling distribution. Hence, the underlying distribution of the statistic can be found when there is not enough data or when the statistic does not behave in the traditional manner as in the case of the Sharpe ratio.

In this paper, the resampling for the Sharpe measure is done "with replacement" of the original excess returns themselves for $j = 1, 2, \ldots, J$ or 999 times. Thus, we calculate 999 Sharpe measures from the original excess return series. The choice of the odd number 999 is convenient, since the rank-ordered 25-th and 975-th values of estimated Sharpe ratios arranged from the smallest to the largest, yield a useful 95% confidence interval. It is from these 999 Sharpe measures that we calculate s_i^{sh}.

The choice of 999 is rather standard in the bootstrap literature,[3] however the question of whether the results are robust to different resample sizes is obviously important. The answer to this question comes from research by Davison and MacKinnon (2000). They conducted elaborate experiments and concluded that the size of the resamples should be at least 399, but recommended a size of 1000 if computational costs were not high. They found that once the resampling size was above 399, the results were robust across larger samples. We found this to be true with our study as well.

4. AN EMPIRICAL EXAMPLE

As an illustration, we have calculated the Sharpe and Double Sharpe ratios for the 30 largest growth mutual funds (as of January 1998 in terms of overall assets managed).[4] Table 1 reports the results for each fund over the time period 1987–1997 with following column headings: (i) the excess mean monthly return;[5] (ii) the standard deviation of the excess monthly returns; (iii) the Sharpe ratio; (iv and v) the mean and standard deviation of the bootstrapped Sharpe ratios; (vi and vii) the lower [0.025] and upper [0.975] confidence values of the bootstrapped Sharpe value; (viii) the 95% confidence interval width; and (ix) the Double Sharpe ratio.

The results yield several interesting findings. First, the Sharpe ratios vary widely among the 30 growth funds. They range from 0.1346 (Fidelity Trend) to almost twice that level, 0.2683 (Fidelity Contrafund).

Second, the bootstrap distribution, which incorporates random variation in estimates of both the numerator and denominator, is generally a good approximation to the true sampling distribution of the estimate. In our case, the sampling distribution representing the estimation risk is non-normal with positive skewness. This is why the means of the bootstrapped Sharpe ratios (col. iv) are always slightly higher than the point estimates of Sharpe ratios (col. iii), which ignore the estimation risk altogether. Also, the standard deviation of the bootstrapped Sharpe ratios are quite variable, suggesting that

Table 1. Characteristics of the 30 Largest Growth Funds.

Fund Name	Excess Mean Monthly Return (%)	Standard Dev. of excess returns	Sharpe Ratio	Mean of Boot-strapped Sharpe Ratios	Standard Dev. of Boot-strapped Sharpe Ratios	Lower Confidence Value (0.025) for Sharpe Ratio	Upper Confidence Value (0.975) for Sharpe Ratio	95% Confidence interval width	Double Sharpe Ratio
AIM Value A	1.1150	4.529	0.2461	0.2569	0.1037	0.0672	0.4727	0.4055	2.372
AIM Weingarten A	0.9252	4.894	0.1890	0.1959	0.0948	0.0174	0.3818	0.3644	1.995
Amcap	0.8209	4.407	0.1863	0.1904	0.0967	0.0107	0.3867	0.3759	1.927
Amer Cent-Growth	0.9158	5.849	0.1566	0.1629	0.0942	–0.0149	0.3513	0.3662	1.662
Amer Cent-Select	0.7191	4.642	0.1549	0.1607	0.0949	–0.0247	0.3529	0.3777	1.632
Brandywine	1.1890	6.244	0.1904	0.1973	0.0964	0.0179	0.3912	0.3733	1.975
Davis NY Venture	1.0990	4.313	0.2547	0.2624	0.0986	0.0779	0.4715	0.3936	2.583
Fidelity Contrafund	1.2360	4.609	0.2683	0.2919	0.1225	0.0513	0.5319	0.4806	2.190
Fidelity Destiny I	1.1490	4.746	0.2421	0.2568	0.1041	0.0640	0.4694	0.4054	2.326
Fidelity Destiny II	1.2000	4.954	0.2423	0.2524	0.1076	0.0465	0.4788	0.4323	2.253
Fidelity Growth	1.0560	5.306	0.1991	0.2074	0.1026	0.0225	0.4231	0.4007	1.940
Fidelity Magellan	0.9955	4.644	0.2144	0.2234	0.1037	0.0402	0.4479	0.4077	2.068
Fidelity OTC	0.9595	5.069	0.1893	0.2026	0.1064	0.0125	0.4313	0.4189	1.780
Fidelity Ret. Growth	0.8455	4.761	0.1776	0.1907	0.1029	–0.0006	0.4053	0.4059	1.726
Fidelity Trend	0.7113	5.286	0.1346	0.1459	0.0989	–0.0309	0.3443	0.3751	1.361
Fidelity Value	0.8130	4.209	0.1932	0.2118	0.1136	0.0098	0.4486	0.4388	1.700
IDS Growth A	1.0230	5.371	0.1905	0.2017	0.0945	0.0214	0.3818	0.3605	2.017
IDS N.Dimensions	1.0680	4.399	0.2428	0.2447	0.0944	0.0652	0.4355	0.3703	2.572
Janus	0.9459	3.822	0.2475	0.2510	0.0964	0.0803	0.4649	0.3847	2.568
Janus Twenty	1.0580	5.113	0.2069	0.2092	0.0980	0.0245	0.4149	0.3904	2.112
Legg Mas. Val. Prim	0.9320	4.629	0.2013	0.2153	0.1049	0.0207	0.4288	0.4081	1.919
Neuberg&Ber Part	0.8707	3.786	0.2300	0.2385	0.1017	0.0464	0.4505	0.4041	2.261
New Economy	0.9237	4.450	0.2076	0.2194	0.0997	0.0409	0.4199	0.3790	2.081
Nicholas	0.8691	3.821	0.2274	0.2339	0.1035	0.0435	0.4512	0.4078	2.197
PBHG Growth	1.2530	7.079	0.1770	0.1813	0.0864	0.0064	0.3502	0.3438	2.047
Prudential Equity B	0.8303	4.224	0.1966	0.2087	0.1106	0.0140	0.4246	0.4106	1.777
T. Rowe Growth	0.7829	4.324	0.1811	0.1911	0.1038	0.0013	0.4032	0.4020	1.745
Van Kampen Pace	0.7598	4.519	0.1681	0.1885	0.1003	0.0004	0.4048	0.4044	1.627
Vanguard U.S. Growth	0.8838	4.459	0.1982	0.2017	0.0987	0.0125	0.4107	0.3982	2.009
Vanguard/Prime	0.9986	5.156	0.1937	0.2089	0.0967	0.0282	0.3956	0.3674	2.003

the estimation risk is variable. The lowest is PHBG with a value of 0.0864, the highest, more than 40% higher than PBHG, is the Fidelity Contrafund with a value of 0.1225.

Third, four of the 30 funds have lower bound confidence values that are negative. This indicates that these funds have Sharpe ratios which are not significantly different from 0 (at the 5% level) when estimation risk is considered. Moreover, the results on the confidence interval width show that it is not entirely intuitive which funds will have narrower confidence intervals; the fund with the highest standard deviation of returns actually has the narrowest confidence interval (PBHG Growth).

Fourth, the Double Sharpe ratios provide a different set of ranking of the funds. Indeed, the Spearman rho rank correlation test of Sharpe ratio rankings to the Double Sharpe rankings is 0.197 and the test does not reject the null hypothesis that the correlation of the two sets of rankings is 0. *Hence, it is not the case that funds with high Sharpe ratios will necessarily have low estimation risk*. Such results indicate that an investor who uses the Sharpe ratio should use the Double Sharpe ratio also. Only the latter incorporates the estimation risk. If the fund scores well in its Sharpe ratio and Double Sharpe ratio rankings, the investor can be assured that the fund has performed relatively well and is not subject to high estimation risk. Again, the weighting of the estimation risk in the Double Sharpe could be questioned. Since the investor is using the implicit weighting in the Sharpe ratio anyway, we start by weighting the estimation risk in a similar fashion. More sophisticated investors could of course adjust the weighting to reflect their individual preferences for representing the importance of estimation risk.

5. CONCLUSIONS

It is often the case that financial researchers assume that investors know the true mean, variance, covariance, etc. of stock returns. But in reality, investors rarely have this information. For example, to apply the elegant framework of modern portfolio theory, investors must estimate expected stock returns using whatever information is currently available. However, since there is often much random noise in financial markets, the observational statistics used by investors need not coincide with true parameters. This is increasingly being referred to as *estimation risk*.

The notion of estimation risk is becoming more and more of an issue in finance. In a recent paper, Lewellen and Shanken (2000) document that not accounting for estimation risk can lead to different empirical findings in regards to stock market return predictability and market efficiency. In this paper, we showcase how one of the most widely taught and used performance metrics, the Sharpe ratio, does not easily account for this estimation risk. This limitation of the Sharpe ratio is often ignored in spite of the fact that the Sharpe ratio is regularly taught in investments classes right along side with the Jensen and 4-index alpha models, which of course examine the underlying estimation error by also investigating the significance of the alpha.

To deal with this issue we propose a modified version of the Sharpe ratio, which we call the Double Sharpe ratio. Rather than calculating the confidence interval width, it *directly* takes into account estimation risk in much the same fashion that the original Sharpe ratio takes into account the portfolio risk. In this way, estimation risk is implicitly in the performance measure with the same weighting as standard deviation receives in the original Sharpe ratio. As such, the Double Sharpe ratio allows investors to examine the risk- and estimation risk-adjusted performance in a relatively simple fashion.

Since computational abilities have improved dramatically with the advent of more powerful personal computers, calculating the Double Sharpe ratio is not at all out the realm of most fund managers. The calculation process involves using a simple bootstrap methodology that can be easily programmed in most statistical packages. Indeed, there are many free downloadable bootstrap codes potentially useful for calculating the Double Sharpe ratios. Software packages including S-Plus, SAS, Eviews, Limdep, TSP and GAUSS have public websites, which can be searched for bootstrap options. See Vinod (2000) for references to software web sites and comparisons of numerical accuracy. By using the Double Sharpe ratio in conjunction with the original Sharpe ratio, the fund manager/investor can understand the performance of the portfolio well, in terms of a performance metric that most investors know, as well as getting a grasp of the estimation risk of the portfolio.

We do not claim that the Double Sharpe ratio is always better than others, or that it cannot be further improved. However, we do claim to provide more information regarding the estimation risk to those who still use the Sharpe ratio. Admittedly, the Double Sharpe ratio, since it is based on the Sharpe ratio suffers from the known limitations of the Sharpe ratio, arising from its arbitrary weighting of risk and ignoring negative returns, skewness and kurtosis.[6] However, Sharpe ratio remains one of the most-widely used and taught performance metrics in the finance literature. No one suggests exclusive reliance on the Sharpe ratio. In the same spirit, Double Sharpe ratio is intended to provide a relatively simple and straightforward method of dealing with the estimation risk, to be used in conjunction with other measures.

NOTES

1. According to the EconLit Database, April 1999.
2. The notation in this section is the same as Jobson and Korkie (1981).
3. See Davison and MacKinnon (2000) and Davison and Hinkley (1997).
4. These funds are classified as "Growth" by Morningstar Inc.
5. We subtract the monthly one-month U.S. T-Bill rates from the monthly fund returns to calculate the excess returns.
6. Note that our method also suffers from the same problems as the Sharpe ratio in that there may be some error in calculating the estimation risk due to non-synchronous trading. Kryzanowski and Sim (1990) address ways to correct for such problems.

REFERENCES

Davison, R., & MacKinnon, J. (2000). Bootstrap Tests: How Many Bootstraps. *Econometric Reviews*, *19*, forthcoming.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications*. New York: Cambridge University Press.

Jobson, J. D., & Korkie, B. M. (1981). Performance Hypothesis Testing with the Sharpe and Treynor Measures. *Journal of Finance*, *36*(4), 889–908.

Kryzanowski, L., & Sim, A. B. (1990). Hypothesis Testing with the Sharpe and Treynor Portfolio: Performance Measures Given Non-Synchronous Trading. *Economics Letters*, *32*, 345–352.

Lewellen, J., & Shanken, J. (2000). Estimation Risk, Market Efficiency and the Predictability of Returns. National Bureau of Economic Research, Working Paper No. 7699, May.

Sharpe, W. F. (1966). Mutual Fund Performance. *Journal of Business*, *39*(1), Part 2, January, 119–138.

Vinod, H. D., & Morey, M. R. (2000). Confidence Intervals and Hypothesis Testing for the Sharpe and Treynor Performance Measures: A Bootstrap Approach. *Computational Finance 99*. MIT Press, Chapter 3, 25–39.

Vinod, H. D. (2000). Review of GAUSS for Windows, Including its Numerical Accuracy. *Journal of Applied Econometrics*, *15*, 211–220.

This Page Intentionally Left Blank

INSTITUTIONAL OWNERSHIP, ANALYST FOLLOWING, AND MARKET LIQUIDITY

Shantaram P. Hegde and Susan M. Mangiero

ABSTRACT

We investigate the difference in the stock market liquidity of a sample of matched firms with high ("institutional favorites") *and low* ("neglected firms") *institutional equity ownership and analyst coverage. We report three findings that help to better understand the complex relationship between firm visibility and liquidity. First, stocks of highly visible companies trade in more liquid markets and with lower adverse selection costs, relative to those of less visible firms. Results reflect controls for firm size, industry classification and exchange listing. Second, the effects of a change in institutional ownership and analyst following on stock market liquidity seem to vary between* institutional favorites *and* neglected firms. *Specifically, the improvement in liquidity from an increase in institutional ownership and analyst coverage is typically minimal for* neglected firms. *In contrast, we find evidence of greater liquidity with an increase in institutional holdings and analyst coverage for* institutional favorites. *Third, market liquidity is likely to be lower at the opening of the market, especially for* neglected firm *stocks.*

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 67–94.

ISBN: 0-7623-0798-6

INTRODUCTION

Do financial institutions and analysts increase or decrease market liquidity? This question assumes growing importance because of the sharp increase in institutional stockholdings and analyst coverage in recent years.[1] An increase in institutional ownership and analyst coverage will impose adverse selection risk on the rest of the market and reduce market liquidity (Admati & Pfleiderer, 1988). This assumes that institutional investors and analysts engage primarily in the production of private information about the value of a firm. Furthermore, these parties are thought to trade on this information as a way of generating profit in a non-competitive market environment. However, if institutions and analysts instead disseminate available information about a firm, monitor managerial performance and attract more investors, visibility helps to promote active trading in the stock, increasing its market liquidity (Easley, O'Hara & Paperman, 1998; Seppi, 1990; Merton, 1987). This means that the overall liquidity effects depend upon whether financial institutions and analysts enhance information-motivated trading or liquidity-motivated trading in a stock.

Although several studies have examined the role of institutional owners and analysts in influencing a stock's market liquidity, the available evidence is mixed. For example, Jennings, Schnatterly and Seguin (1997) report that high institutional ownership leads to both lower bid-ask spreads and a smaller adverse selection component of the spread. In contrast, Kothare and Laux (1995) argue that institutional investors are informed traders and conclude that growth in institutional ownership is in part responsible for the widening of bid-ask spreads on NASDAQ issues. Similarly, Sarin, Shastri and Shastri (1996) find that higher institutional holdings are associated with wider spreads. However, they find no evidence to suggest that institutional owners trade on superior information.

The available evidence on the liquidity effects of analyst coverage is equally mixed. Brennan and Subrahmanyam (1995) use the number of analysts following a firm as a simple proxy for the number of individuals producing information about the value of a firm. They report that an increase in analyst coverage leads to a lower adverse selection cost and a deeper market due to enhanced competition among informed traders. Easley, O'Hara and Paperman (1998) report that analyst coverage tends to increase uninformed trading, thus leading to a lower adverse selection cost for high analyst stocks. They conclude that other proxies besides analyst coverage better capture the presence of information-based trading. On the other hand, Chung, McInish, Wood and

Wyhowski (1995) find a significant positive simultaneous relationship between the number of analysts following a stock and its bid-ask spread.

It is clear from this brief review of conflicting evidence that we do not have a solid understanding about the influence of institutional ownership and analyst following on the liquidity of underlying stocks. Our objectives in this study are twofold. First, we examine the relationship between liquidity and the durability of private information, including tests for time-of-day patterns. Second, we reexamine the liquidity effects of institutional equity holdings and analyst coverage by using an alternative research methodology, correcting for limitations with past studies.[2]

Existing literature on information-motivated trading suggests that liquidity effects may be conditional on both the number of informed traders as well as the durability of private information about the value of a security. To illustrate, consider the models of Holden and Subrahmanyam (1992) and Subrahmanyam (1991) that imply a variable relationship between information (liquidity) and visibility measures such as institutional holdings and analyst following. At low levels, the relationship is ambiguous because of informed traders' risk aversion and durability of private information. At high levels, liquidity improves with an increase in institutional and analyst interest in a stock because informed traders compete more aggressively and attract additional uninformed traders to the market (Admati & Pfleiderer, 1988; Easley, O'Hara & Paperman, 1998; Merton, 1987).

Regarding methodology, prior studies have relied on the application of linear regression to a large data set in order to evaluate the simultaneous relationship between liquidity and institutional ownership and/or number of analysts following a stock, while controlling for other explanatory variables. We employ a "matched pair" methodology that is quite effective in controlling for nuisance factors as well as in handling the nonlinear relationship between liquidity and the number of informed traders.[3]

Under this methodology, we construct two samples of stocks. *Institutional favorites* (*neglected firms*) are stocks with high (low) institutional ownership and analyst following. Direct comparison of market liquidity between these two samples is complicated because the *institutional favorites* are substantially larger in market capitalization than the *neglected firms*. Therefore, we construct two control groups of stocks that are matched with the *institutional favorites* and *neglected firms*, respectively, in terms of firm size, industry classification, and exchange listing.[4] Stocks included in the control groups belong to the middle level of institutional ownership and analyst following.

Our empirical analysis is based on transactional data for March 1995, using 29 matched pairs of stocks.[5] Based on the model of Lin, Sanger and Booth

(1995), we estimate the adverse selection component of quoted spreads as a measure of liquidity. Our results suggest that *neglected firm* stocks are more vulnerable to greater adverse selection risk (less liquidity) than stocks issued by *institutional favorites*. Interestingly, the liquidity benefits from increasing visibility with institutions and analysts seem to be minimal for the *neglected firms* in our sample. Only at the high level of institutional holdings and analyst following do we observe a significant improvement in liquidity. Finally, our findings suggest that the adverse selection cost is higher at the open of the trading day, particularly for *neglected firm* stocks.

The rest of this paper is organized as follows. Section 1 develops our propositions concerning the liquidity and information effects of institutional ownership and analyst coverage. Sample selection and transaction data used in the study are described in Section 2. We present the empirical results and discuss their implications in Section 3. Section 4 concludes the paper.

1. ALTERNATIVE PERSPECTIVES

Trading by informed investors makes market prices more efficient by bringing them closer to the true values of securities. However, informed trading produces a negative side effect for the less-informed traders, including market makers. They face the risk of buying at a higher price or selling at a lower price than the unknown true value. In other words, informed trading imposes "adverse selection risk", increases the cost of trading, and decreases market liquidity. In contrast, trading by uninformed investors (also called liquidity or noise traders) improves market depth by reducing the severity of adverse selection risk faced by other uninformed traders, including security dealers. Uninformed traders do not influence market prices, ignoring any new information that would otherwise be a factor for informed traders. For example, pension funds, mutual funds, hedge funds and other types of institutions often passively trade at market prices when required to re-balance their portfolio.

There is no consensus as to how to identify an informed trade though trade size and timing are frequently used measures. Seppi (1990) argues that trade size is an ambiguous measure of informed trading because informed institutions may either trade in blocks or submit a sequence of anonymous market orders at less favorable prices or both. Institutional equity ownership and analyst following are the other frequently used proxies. In this study, we use both institutional equity ownership and analyst coverage as simple proxies for informed trading, similar to Brennan and Subrahmanyam (1995).

A. Information-based Models

We begin with the pure information-based models. Both financial institutions and analysts are assumed to specialize in the production of private information about the value of a firm and trade on that information. Studies in information economics show that the relationship between the number of informed traders and the amount of adverse selection risk faced by uninformed traders and market makers is nonlinear and complex. The flow of private information to the market tends to increase directly with the number of informed traders, implying that the adverse selection cost likewise increases directly with the number of informed traders. This view is illustrated in Fig. 1.

To keep the graph simple, line segments are used, but the underlying theory does not necessarily imply that the relationship is linear. The horizontal axis represents the low and high levels of institutional holdings and analyst coverage, *n*, as proxies for informed trading. In our analysis, these two levels correspond to the *neglected firms* and *institutional favorites*, respectively. The

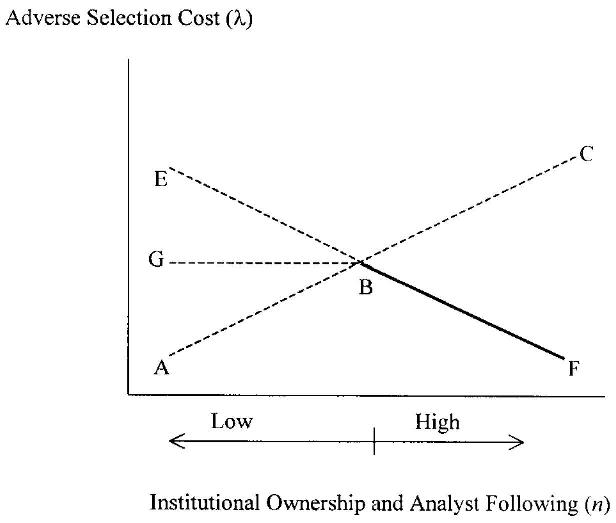


Fig. 1. Alternative Perspectives on the Relation Between the Adverse Selection Cost and Institutional Ownership and Analyst Coverage.

Line segments ABC and ABF denote the pure information-based models. Line segment EBF represents the pure liquidity-based model. The bold line segment denotes the convergence of the two models.

vertical axis denotes the adverse selection cost of trading, λ. Line segment ABC reflects the view that the flow of private information to the market and the adverse selection cost each increase monotonically with the level of institutional participation and analyst following.[6]

The foregoing discussion ignores the liquidity effect of competition among informed traders. Admati and Pfleiderer (1988) show that λ decreases with *n* when private information is short-lived and informed traders are risk neutral. This monotonic decline in λ is due to intense competition among the informed and is reflected by the line segment EBF in Fig. 1. Subrahmanyam (1991) relaxes the assumptions underlying the Admati and Pfleiderer model to allow for risk aversion among informed traders. He finds that λ increases with *n* for small *n* (line segment AB) because risk aversion dampens the intensity of trading by informed traders. He concludes that the relation between λ and *n* will be negative (line segment BF) only when *n* is sufficiently large.

When private information is durable, the predictions are even more complex. With long-lived information, Holden and Subrahmanyam (1992) point out that dealers and uninformed investors face more adverse selection risk in the initial rounds of trading, when their information disadvantage is greatest. But adverse selection risk decreases in the subsequent rounds of trading as market makers and noise traders gain information from the order flow and the trading process.

Overall, the pure information-based models predict a negative relationship between the adverse selection cost and the level of institutional ownership and analyst coverage when *n* is large. The relationship is ambiguous for small *n* because of risk aversion and durability of information. The dashed line segments AB and EB reflect these uncertainties. The dark line segment BF in Fig. 1 represents the unambiguous negative relationship.

B. Liquidity-based Models

Although many earlier theoretical studies have focused on the informational role of institutions and analysts, more recent work highlights their role in mitigating the degree of informational asymmetry between informed and uninformed investors. This line of work, dubbed here as the liquidity-based view, is skeptical of the ability of most institutions and analysts to produce valuable private information. Instead, this research contends that institutions and analysts engage primarily in the distribution of available information about a firm to a larger group of investors and in the monitoring of managerial performance.

Under this view, an increase in institutional ownership and analyst following accelerates the rate of flow of public information, reduces the information advantage of informed traders, and thus attracts more uninformed traders to the market (Easley, O'Hara & Paperman, 1998; Schleifer & Vishny, 1986).[7] As the investor base in a stock increases with institutional holdings and analyst coverage, dealers face less adverse selection risk and the firm faces a lower cost of equity capital (Merton, 1987). This view is consistent with the widespread evidence that most active mutual fund managers, investment advisors, and newsletters fail to outperform the market portfolio (Grinblatt, Titman & Wermers, 1995; Clements, 1998).

For instance, according to Lipper Analytical Services, 86% of diversified U.S. stock funds have under-performed the Standard and Poor's 500-stock index over the past 10 years. Furthermore, the average performance of stock picks by investment newsletters (before adjusting for trading costs) is very close to that of the market portfolio (Clements, 1998). Critics argue that analysts are under increasing pressure to deliver good news and promote their clients' stocks, enhancing the chances of their employers winning good investment banking deals from corporations. Consequently, analysts' research reports may be biased, failing to provide an independent assessment of a firm's prospects. A recent survey by First Call Corporation indicates that analysts currently have "strong buy" or "buy" recommendations on 66% of the nearly 6,000 stocks, but "sell" recommendations on just 1%, the rest being classified as "holds" (Siconolfi, 1998).

The message is that most institutions and analysts are unlikely to possess superior information. Instead, they may simply bring added publicity and visibility to the firms they follow, helping to expand the investor base. The resulting increase in the number of uninformed traders makes it easier for market makers to recover their losses from transacting with informed traders. Therefore, under the liquidity-based models, adverse selection risk decreases and market liquidity improves with an increase in institutional ownership and analyst following. This view conforms to the negatively sloped line segment EBF in Fig. 1.

Nowhere is there evidence that institutional ownership and analyst coverage jointly *cause* market liquidity. Rather, the relationship is simultaneously determined. According to the information-based models, institutions and analysts gravitate towards firms with greater information asymmetry in order to maximize expected profits from information gathering. Consequently, market makers face greater adverse selection and a wider bid-ask spread as both institutional holdings and number of analysts following a particular stock increase. In contrast, liquidity-based models postulate that analysts and institutions prefer liquid stocks in order to minimize trading costs. This means that market makers face a lower probability of information-motivated trades in stocks with greater institutional participation and analyst following.

C. Testable Implications

We know that the effect on market liquidity may be non-monotonic. It will depend upon the level of institutional ownership and analyst following, our proxies for informed trading. At low levels of institutional ownership and analyst coverage, there are likely to be fewer informed traders and less private information. However, there may be less intense competition among the (risk adverse) informed traders. Both of these aggravate the adverse selection problem at the margin. Moreover, private information is likely to be more durable for these *neglected firms*, due to less frequent firm-specific public announcements. As a result, when institutional holdings and analyst coverage are low (i.e. for *neglected firms*), the net effect on market liquidity seems to be ambiguous.

When the levels of institutional ownership and analyst following are high, there may be many more informed traders but they compete much more aggressively and attract far more liquidity trading. These factors will hasten the perishability of private information. Consequently, market liquidity may increase with an increase in institutional ownership and analyst coverage at the higher level. These arguments suggest that market liquidity may increase with an increase in institutional ownership and analyst following for *institutional favorites*, resulting from enhanced competition among informed traders as well as higher liquidity-motivated trading. In contrast, the competitive and visibility effects may be ambiguous for *neglected firms*.

Another issue of interest is the intra-day variation in the market liquidity of *neglected firms* and *institutional favorites*. When private information is durable, Holden and Subrahmanyam (1992) predict that uninformed investors and market makers face greater adverse selection risk in the initial rounds of trading. But this risk decreases in the subsequent rounds of trading as they gain information from the order flow. Brennan and Subrahmanyam (1998) observe that a large part of private information is likely to be short-lived, since most of it is nothing more than the advance knowledge of public information. Foster and Viswanathan (1993) note that some liquidity traders can afford to be patient and choose the best time to trade. Such discretionary traders would prefer to avoid trading at either the market open or close in order to minimize adverse selection risk. In addition, there may be a flurry of information-motivated trades just before the trading day ends because of limited trading opportunities thereafter.

Information-based trades may be more concentrated at the market *open* and *close* than at other times of the day for all stocks. Moreover, if *neglected firm* stocks are associated with more durable private information (perhaps due to less frequent firm-specific public announcements), then we would expect the trading process to convey more information to the uninformed throughout the day. This implies a greater decline in the adverse selection cost from the market *open* to the *midday* for *neglected firm* stocks. Thus, the temporal behavior of informed and liquidity traders suggests that market liquidity may be lower at both the market *open* (and *close*), particularly for *neglected firm* stocks with more durable private information.

In the real world, it is likely that both the pure information and pure liquidity effects are simultaneously associated with institutional ownership and analyst coverage. The empirical challenge is to evaluate not only their joint effects but also to disentangle the two sources of liquidity.

2. METHODOLOGY AND DATA

We noted that the liquidity effects might be ambiguous when only a few institutions and analysts are associated with a stock, but liquidity will likely increase with a rise in institutional holdings and analyst coverage. Moreover, the relationship between liquidity and institutional and analyst interest is simultaneous. To account for the potential non-monotonic relationship between liquidity and proxies for informed trading, we construct a matched pair sample – matched first on the basis of exchange listing (New York Stock Exchange, NYSE), next on industry classification (four-digit SIC code), and finally on firm size (market capitalization of equity).[8,9] We also require that each stock have an average price in excess of $3, trade at least ten times per day, and not have been split or reverse split during 1995. Applying these selection criteria to nearly 2,000 stocks traded on the NYSE in March 1995 yields a sample of 58 stocks or 29 matched pairs. The sample consists of four groups of stocks (number of firms in parentheses): *institutional favorites* (13), control group for *institutional favorites* (13), *neglected firms* (16), and control group for *neglected firms* (16). Because of high positive correlation between institutional and analyst interest and firm size, *institutional favorites* are substantially larger in market capitalization than the *neglected firms* in our sample. We use two control groups so that they can be better matched with the *neglected firms* and *institutional favorites*. The two control groups each represent the medium level of institutional holdings and analyst following but differ in firm size.

We use the number of institutional investors, percent of equity holdings of institutions, and the number of analysts providing earnings estimates as measures of institutional ownership and analyst coverage (collected, respectively, from "13f" filings with the Securities and Exchange Commission and Institutional Brokers Estimate System, I/B/E/S).[10,11] Most previous studies tend to focus on either institutional investors or the number of analysts following a stock, but rarely both. Because these variables are imperfectly correlated, we believe it is important to explicitly account for the joint effects of both institutional ownership and analyst coverage on market liquidity.[12]

We base our study on Trade and Quote Data (TAQ) for 23 trading days during March 1995 provided by the New York Stock Exchange. We choose transactions data because of the perishability of most private information. This means that a large part of information content is short-lived, unlikely to last beyond a trading day or a week. Holden and Subrahmanyam (1992) observe that private information tends to be revealed very quickly, even with as few as two informed agents. Admittedly, one month is a short time window, relative to several previous studies that use year-long transactions data (i.e. Foster & Viswanathan, 1993; Lin, Sanger & Booth, 1995). Our rationale for the choice of a short time interval is, besides its convenience, the fact that the database covers numerous trades and quotes. For example, the average number of trade observations available for each firm ranges from 799 for *neglected firms* to 2533 for *institutional favorites*, minimizing the potential small sample bias. Therefore, we believe that the use of trade and quote data over 23 days on a matched sample of stocks is sufficient to test for the non-monotonicity of information and liquidity effects attributable to institutional participation and analyst coverage.

Table 1 presents descriptive information on our sample. As shown in Panel A, *institutional favorites* are firms with: (i) institutional equity holdings of 75% or more; (ii) more than 300 institutional investors; and (iii) more than 12 analysts providing earnings estimates. In contrast, *neglected firms* are firms with: (i) institutional equity holdings of 25%; (ii) under 101 institutional investors; and (iii) less than 5 analysts following their stocks. Control groups have institutional participation and analyst coverage in between these extremes.

Summary statistics are shown in Panel B. *Institutional favorites* (control group firms for *institutional favorites*) have a median institutional ownership of 86% (51%), 341 (200) institutional investors, 15 (10) analysts, and market

Table 1. Descriptive Statistics.

The sample consists of 29 pairs of New York Stock Exchange-listed stocks that are matched on firm size and the four-digit SIC code. It covers 23 trading days in March 1995. Institutional holdings denote percentage of outstanding common stocks owned by '13f' financial institutions. Firm size is defined as the number of stocks outstanding multiplied by price per share. Panel B reports cross-sectional statistics.

Panel A: Sample Construction

	Institutional Holdings (%)	Number of Institutions	Number of Analysts
Institutional Favorites	75 to 100	> 300	> 12
Control Groups	40 to 60	101 to 300	5 to 12
Neglected Firms	0 to 25	0 to 100	0 to 4

Panel B: Summary Statistics

	Institutional Holdings (%)	Number of Institutions	Number of Analysts	Firm Size ($ billions)
Institutional Favorites (13 firms)				
Mean	86	362	17	1.71
Median	86	341	15	1.54
Standard Deviation	4	60	4	0.68
Favorites Control Group (13 firms)				
Mean	52	207	10	1.68
Median	51	200	10	1.45
Standard Deviation	5	64	2	0.82
Neglected Firms (16 firms)				
Mean	12	45	3	0.50
Median	13	44	3	0.60
Standard Deviation	6	30	1	1.10
Neglected Control Group (16 firms)				
Mean	52	181	9	0.51
Median	52	170	9	0.64
Standard Deviation	4	54	2	0.95

value of equity of $1.54 billion ($1.45 billion). In contrast, the *neglected firms* (control group firms for *neglected firms*) have a median institutional holding of 13% (52%), 44 (170) institutional investors, 3 (9) analysts, and market capitalization of $0.60 billion ($0.64 billion). Notice that the *institutional favorites* and *neglected firms* are vastly different in firm size, necessitating the need for the two separate control groups to better match the subject firms on size. The standard deviation estimates indicate that *institutional favorites* are a more homogeneous group than the *neglected firm* stocks. Thus, our matching procedure seems to be quite effective in maximizing the differences in institutional holdings and analyst coverage between *institutional favorites* and their control group as well as the *neglected firms* and their matched counterparts. Appendix A summarizes additional information about the sample firms.

Since we are interested in the information versus liquidity effects of institutional ownership and analyst coverage, we use the adverse selection component of the effective relative spread as derived by Lin, Sanger and Booth (1995) as our primary measure of market liquidity. In these models, the quoted relative bid-ask spread (defined as the difference between the ask and bid quotes, divided by the quote midpoint) is determined not only by the adverse selection cost but also by fixed order processing costs and order persistence. Assuming the quote midpoint as a proxy for the unknown true value of a stock, these models measure the effective relative spread as twice the absolute value of the percent difference between the transaction price and quote midpoint. Since a dealer faces the risk of trading with speculators who possess superior private information, she adjusts downward (upward) both the bid and ask quotes following a public sell (buy) order. Therefore, the dealer's risk of trading with informed individuals is reflected by the percentage change in her bid-ask quotes in response to trades.

We exclude the first trade of each day since specialists intervene in setting the opening market price.[13] We match trades and quotes by applying the procedure suggested by Lee and Ready (1991).[14] Following Lin, Sanger and Booth (1995), we estimate the adverse selection cost by regressing percentage changes in quote midpoints (quote revisions) on one-half the signed effective relative spreads for each stock over the sample period:

$$\text{Quote revision}_{t+1} = (\lambda \times {}^1\!/_2 \text{ the signed effective spread}_t) + \text{Error}_{t+1}, \qquad (1)$$

where λ is a measure of the adverse selection cost component of the signed effective spread.[15]

3. EMPIRICAL EVIDENCE

A. Results

Table 2 presents our estimates and test results with respect to the competitive and visibility effects of institutional holdings and analyst coverage. In addition to the adverse selection cost component, we present estimates of the quoted relative spread as well as the effective relative spread which are widely used as broader measures of trading costs. The quoted relative spread is defined as the difference between the bid and the ask quotes, divided by the quote midpoint, where the quote midpoint is located midway between the bid and ask quotes. The effective relative spread is given as twice the percentage difference between transaction price and quote midpoint.

Recall that the average number of trades for *institutional favorites* is about 2500 or more than thrice the trade number for *neglected firms*. The relatively

Table 2. Competitive and Visibility Effects.

The *Ask*, *Bid* and *Trade Prices* at time *t* are represented by A_t, B_t, and P_t, respectively. The quoted relative spread is $(A_t - B_t)/Q_t$, where $Q_t = (A_t + B_t)/2$. The quoted relative spread is computed by first averaging across all but the opening trades for each firm and then across firms within each sample. The effective relative spread is $2|\log(P_t/Q_t)|$ and is computed in a manner similar to the quoted relative spread. The adverse selection cost, λ, is estimated for each firm by the time series regression $\mathbf{\Delta Q_{t+1}} = \lambda \mathbf{z}_t + \mathbf{e}_{t+1}$ (Eq. 1) where $\Delta Q_{t+1} = Q_{t+1} - Q_t$, Q_t is the log of the quote midpoint, $z_t = P_t - Q_t$ and P_t is the log of trade price. R^2 denotes the squared correlation coefficient of the regression in Eq. 1. The superscript 'a' denotes that the estimate is significantly different from its matched counterpart at the 5% level, based on the Wilcoxon Signed Rank Test.

		Quoted Relative Spread		**Effective Relative Spread**		**Adverse Selection Cost**		
Sample	Average number of trades per firm	Mean (%)	Median (%)	Mean (%)	Median (%)	Mean (%)	Median (%)	R^2 (%)
Institutional Favorites	2533	1.39	1.31	0.55	0.54[a]	33	30[a]	16
Favorites Control Group	1591	1.53	1.57	0.69	0.64	40	41	21
Neglected Firms	798	2.87	2.50	1.19	0.94	40	37	25
Neglected Control Group	1458	2.17	1.70	1.02	0.72	40	43	21

large number of trades associated with *institutional favorites* is consistent with their greater market liquidity. Median values of the quoted relative spread increase sharply from 1.31% for *institutional favorites* to 2.50% for *neglected firms*. Reflecting the fact that many trades actually take place inside the bid-ask quotes, median values of the effective relative spreads are substantially smaller at 0.54% for *institutional favorites* and 0.94% for *neglected firms*.

The adverse selection cost data, generated from the firm-specific regressions, exhibits scant evidence of normality. Accordingly, we use the Wilcoxon Signed Rank Test to determine if the distribution of spread estimates for *institutional favorites* is significantly lower than (i.e. located to the left of) that for the *neglected firms*.[16] These tests (not reported in Table 2) indicate that both the quoted and effective spreads are significantly smaller, at the 0.1% level, for *institutional favorites* as compared to neglected stocks.

Although these results are in agreement with the findings of some earlier studies, they are suspect because, as noted in panel B of Table 1, the median firm in the *institutional favorites* group is more than twice as large, in terms of market capitalization, as the median neglected firm. As big firms tend to have lower spreads, it is quite likely that the smaller observed spreads for *institutional favorites* are attributable to their larger firm size, rather than to their higher institutional ownership and analyst coverage. To correct for this firm size bias, we compare the *institutional favorites* and *neglected firms* with their respective matched (on firm size, industry classification, and exchange listing) control groups. These additional tests show no significant difference (at the 5% level) in quoted and effective spreads between the *neglected firms* and their matched counterparts. *Institutional favorites*, on the other hand, do have significantly (at the 5% level) lower effective relative spreads than that of the control group. This is not the case for quoted spreads.

For our purposes, the adverse selection component of the effective spread is a superior gauge of market liquidity because it specifically captures the informational versus liquidity effects of institutional holdings and analyst coverage. From the last three columns of Table 2, we notice that the median adverse selection cost for *institutional favorites* is 30% of the effective spread, compared to 37% for *neglected firms*.[17] More importantly, our first testable implication involves a lower-tail test between the *institutional favorites* and their control group, but a two-sided test between the *neglected firms* and their control. These additional tests show that the *institutional favorites*' median adverse selection cost of 30% is significantly lower (at a 5% level) than the 40% for the control group. However, the difference in the distributions of adverse selection costs between *neglected firms* and their matched counterparts is not statistically significant.[18]

In our sample, the median effective spread per share for *institutional favorites*, their control group, neglected stocks, and their control group, respectively, are \$0.125, \$0.136, \$0.144, and \$0.130. Multiplying these dollar effective spreads by the corresponding median estimates of λ reported in Table 2, we get the following median estimates of λ per share: *institutional favorites*, \$0.038; *institutional favorites* control group, \$0.056; *neglected firms*, \$0.053; and *neglected firms* control group, \$0.056. Thus, even in economic terms the adverse selection cost is much lower for the *institutional favorites*, companies with greater institutional ownership and analyst following.

Overall, the liquidity cost estimates presented in Table 2 highlight two results. First, the adverse selection cost is significantly lower (and market liquidity higher) for the *institutional favorites* relative to the *neglected firms*. Second, and more importantly, an increase in institutional holdings and analyst following, after controlling for firm size, produces a significant improvement in market liquidity only for institutional favorites, and not for neglected stocks. In other words, the relationship between the adverse selection cost, λ, and the number of informed traders, n, (as proxied by institutional holdings and analyst coverage) seems to be non-monotonic.

The observed difference in liquidity in our sample is insignificant when the number of informed traders increases from the low to the medium level (i.e. *neglected firms*). When n increases from the medium to the high level (i.e. *institutional favorites*), the liquidity gains are significant. These results are consistent with the models of Admati and Pfleiderer (1988), Subrahmanyam (1991), and Holden and Subrahmanyam (1992). They comport with the line segment GBF in Fig. 1. The flat line segment GB reflects the insignificant change in λ for neglected stocks. The negatively sloped line segment BF indicates the decrease in λ for *institutional favorites*.

Table 3 presents test results on our conjectures regarding the timing effects, i.e. that market liquidity will be lower at the market open and close. This is particularly true for the *neglected firm* stocks because of the durability of private information. In these tests, we first attempt to control for the level of institutional holdings and analyst following in each group of stocks and then test for variation in market liquidity measures across the time of day. We divide the trading day into three parts. The market *open* covers all eligible trades (except the opening trade) between 9:30 and 10:30 a.m. The market *close* includes all trades that occur between 3:00 and 4:00 p.m. and *midday* covers trades during 10:30 a.m. and 3:00 p.m. The adverse selection component of the effective spread is re-estimated separately for each of these three intra-day time intervals.

The average number of hourly trades per firm in Table 3 shows the number of observations underlying the estimate of adverse selection costs at the market *open*, *midday* and *close* as well as the frequency of trading. Consistent with the existing evidence, the average trading frequencies are higher at the market *open* and *close* for all the four groups (Foster & Viswanathan, 1993). For the *institutional favorites*, the median adverse selection cost at the open is 34% of the effective spread as compared with 30% midday. According to the Wilcoxon Signed Rank Test, the distribution of adverse selection cost at the market open is significantly (at a 2% level) higher than (i.e. to the right of) that at the midday. In the same vein, the market *open* cost distributions for the *institutional favorites* control group, *neglected firms* and their control group are all significantly higher (under 10%) than those at the midday. These results support our conjecture that the adverse selection cost is higher and market liquidity lower for all stocks at the beginning of daily trading.

The median drop in the adverse selection cost of *neglected firm* stocks from market *open* to *midday* is 4%, equal to that for *institutional favorites*. The corresponding decreases for the two control groups are 0% and 9%. Therefore,

Table 3. Timing Effects.

The trading day is divided into three time windows: *Open* (9:30 a.m. to 10:30 a.m.), *Midday* (10:30 a.m. to 3:00 p.m.) and *Close* (3:00 p.m. to 4:00 p.m.). The **adverse selection cost**, λ, is estimated separately for each time window by running the time series regression for each firm $\Delta \mathbf{Q}_{t+1} = \lambda \mathbf{z}_t + \mathbf{e}_{t+1}$ (Eq. 1) where (a) $\Delta Q_{t+1} = Q_{t+1} - Q_t$; (b) Q_t is the log of the quote midpoint, $z_t = P_t - Q_t$; and (c) P_t is the log of trade price. The superscript 'a' ('b') denotes that the estimate is significantly different from its *Midday* counterpart at the 5% (10%) level, based on the Wilcoxon Signed Rank Test.

				Adverse		**Selection**		**Cost**	
				Open		**Midday**		**Close**	
Sample	Number of Hourly Trades per Firm (Open	Midday	Close)	Mean (%)	Median (%)	Mean (%)	Median (%)	Mean (%)	Median (%)
Institutional Favorites	507	347	423	37	34[a]	33	30	31	32
Favorites Control Group	322	221	278	46	41[b]	38	41	36	35
Neglected Firms	111	85	110	41	39[b]	36	35	40	36
Neglected Control Group	215	173	209	46	46[a]	37	37	37	39

we find only weak evidence of any difference in the decline in adverse selection cost from market *open* to *midday* across all stocks. Finally, these test results show little evidence of an increase in the adverse selection cost at the market *close*. The absence of an increase in adverse selection cost at the market *close* may imply that most private information is accumulated overnight, is short-lived and hence traded on early in the day.[19,20]

B. Discussion

Our analysis uncovers three important results. First, in our sample we notice no significant change in the adverse selection cost component of the effective bid-ask spread for *neglected firm* stocks. That is, there is little change in adverse selection cost as the median institutional holdings and number of analysts each increase from 13% and 3, respectively, for *neglected firm* stocks to 52% and 9 for the related control group. Second, we find a significant drop in adverse selection cost as the median institutional ownership and analyst coverage goes up from 51% and 10 for the *institutional favorites* control group to 86% and 15 for *institutional favorites*. Finally, we document a significant increase in adverse selection cost at the market *open* (relative to *midday*), particularly for stocks with lower institutional holdings and analyst following. The first two findings suggest that institutions and analysts make little difference to the liquidity of stocks up to the median level of their equity ownership and coverage. It is only at the high level of institutional holdings and analyst following that we find a significant improvement in liquidity.

Before we examine the implications of these findings, we need to first establish that our estimates of adverse selection cost, derived from a small matched sample of 29 pairs of stocks with transactions data over 23 trading days, are reliable and representative. The matching procedure enhances control over noise variables but inevitably takes its toll on the sample size.[21] Since we use the same estimation procedure as used by Lin, Sanger and Booth (LSB, 1995), it is informative to compare our sample and related estimates with those of LSB. The LSB study is based on transactional data for the entire year of 1988 and a sample of 150 NYSE stocks – 50 each selected from the low, medium and high trading volume categories. Their average firm sizes are 0.21, 1.12 and 5.66 billion U.S. dollars, respectively, for the three volume groups. From Panel B of Table 1, our *institutional favorites* are considerably smaller ($1.71 billion) than their high trading volume group, but our *neglected firms* are much larger in size ($0.60 billion) than their low volume sample. Average quoted relative spreads for the LSB sample are 2.10%, 1.03% and 0.60%, respectively, for the low, medium and high volume stocks. From Table 2, our average quoted relative spreads are 1.39% for *institutional favorites* and 2.87% for *neglected firms*. Average number of trades per firm (projected over a 23-day period), as reported by LSB, are 552, 1,380 and 3,680, respectively, for the three volume groups. Our samples of 798 trades for *neglected firms* and 2533 trades for *institutional favorites* are somewhat similar to the LSB sample size. Finally, LSB's average λ estimates are 0.42, 0.42 and 0.33, respectively, for the low, medium and high volume groups.[22] These estimates compare favorably with our average estimates of 0.40 for neglected stocks and the two control groups and 0.33 for *institutional favorites*. Thus, even though our sample is small and the study period short, our estimates of the adverse selection cost seem reasonable and representative of those based on a larger sample used by LSB. Moreover, the quote revision model used in our study complements the alternative procedures employed in other studies of institutional holdings and analyst coverage.

Unfortunately, our results are not directly comparable with the findings of three recent studies that focus on the informational role of institutions and analysts (Brennan & Subrahmanyam, 1995; Jennings et al., 1997; Sarin et al., 1996) because of differences in the procedures used to estimate the adverse selection cost of trading. We use quote revisions in estimating λ whereas others focus on transaction price changes, order flow, and other variables.[23] Unlike our work based on matched pairs, these studies report regressions results based on much larger samples. It is worth noting, however, that Sarin et al. (1996) find no relationship between adverse selection cost and institutional ownership. This is broadly consistent with our first result regarding *neglected firm* stocks. Brennan and Subrahmanyam and Jennings et al. conclude that high analyst coverage and institutional holdings lead to a smaller adverse selection cost. This comports with our second result concerning *institutional favorites*. We believe our matched pair methodology is better suited to account for the non-monotonic relation between liquidity and the number of informed traders and thus able to reconcile the conflicting evidence reported by these studies.

Our first finding that the overall liquidity effects are minimal up to the medium level of institutional ownership and analyst coverage is also consistent with Easley, Keifer, O'Hara and Paperman (1996). They report no difference in the risk of information-based trading between the medium and low volume stocks. Our *neglected firm* stocks and the two control groups are respectively comparable with the low and medium volume stocks used in previous studies.[24] This finding implies that either the pure information effect washes out the pure liquidity effect at the low and medium level of institutional holdings and analyst following or that liquidity of trading is hardly affected at these levels.

The first implication suggests that the added liquidity-motivated trades, induced by the incremental institutional and analyst interest, leave the adverse selection cost unchanged, even though the amount of private information increases as we go from the low to the medium level. We think, however, that such an offsetting pattern in the pure information and liquidity effects is less plausible. What is perhaps more realistic is the second implication that, at the low and middle level of institutional ownership and analyst coverage, the liquidity effects are inconsequential. Thus, our first result suggests that up to the medium level, most institutions and analysts are neither informed nor capable of attracting liquidity-motivated traders. More specifically, the overall liquidity influence of institutions and analysts is not typically distinguishable from those of a few small individual investors at the low and medium levels of visibility.

Our second finding shows that the adverse selection cost decreases for *institutional favorites*. These results suggest that an increase in institutional ownership and analyst following from the medium to the high level leads not only to greater production of private information, but also to intense competition among informed traders and attracts more uninformed traders. These findings imply that less than half the number of institutions and analysts produce and trade on private information and are consistent with the available evidence that only a small fraction of institutional investors and advisors produce abnormal returns attributable to superior information and analytical ability. From the firm's point of view, these results imply that the liquidity effects emerge only after a stock attracts a critical mass of interest from institutions and analysts. Firms seeking to improve the market liquidity of their stocks should augment their network of institutional investors and analysts as much as possible. Our third and final result indicates that the adverse selection cost is significantly higher at the market *open* for all stocks. We find only weak evidence to suggest that stocks with fewer institutional holdings and analyst coverage experience better improvement in liquidity from the market *open* to *midday.*

C. Vicious Circle of Illiquidity

In our sample, firms with low institutional ownership and analyst following are typically small in size (median market capitalization of $600 million), while those with the high level are large ($1.5 billion). Small firms have a less liquid market (median quoted spread of 2.50%) whereas the large stocks trade in a more liquid market (median quoted spread of 1.31%). An alternative (but not mutually exclusive) explanation, consistent with our results, incorporates the fact that institutions and analysts prefer to trade in big lots for both information and liquidity reasons (Seppi, 1990). Information production involves large fixed costs and thus induces informed parties to trade large blocks of equity in order to recover their costs. Even the trades that are triggered by institutional portfolio re-balancing tend to be fairly large in size. Informed institutions and analysts tend to shy away from small stocks, presumably because of their risk aversion, the difficulty in establishing or liquidating positions in thin markets, and the limited potential for lucrative investment banking deals offered by these firms. Consequently, other informed traders such as officers of the firm, venture capitalists (insiders) and other non-institutional investors, face less competition. Outside investors face a higher adverse selection cost in small stocks. Large firms, on the other hand, attract more institutions and analysts because of their already deep markets and their ability to offer promising investment banking opportunities. Competition among the informed attracts more uninformed trading and lowers the adverse selection cost. This implies that liquidity gains from the growing institutionalization of the stock market are not uniformly distributed across all stocks. The small, illiquid stocks (*neglected firm* stocks) continue to remain illiquid. The large, liquid stocks (*institutional favorites*) tend to become more liquid with an increase in institutional participation and analyst coverage.

Finally, it is important to point out two limitations of this study. First, our investigation of liquidity effects examines only two variables – institutional ownership (number and percentage of stock owned) and analyst following. Admittedly, these are broader, aggregate indicators in comparison to the finer (less noisy) proxies for information production and trading (i.e. institutional trades or analyst recommendations). For instance, a proprietary trade by an investment bank may convey more information than a trade by a pension fund. So it is quite possible that our inability to uncover any change in the liquidity effects up to the medium level of institutional ownership and analyst coverage is simply because these variables are not quite robust in detecting information-motivated trading. Therefore, it is important to bear in mind that our findings are limited to the ability of institutional ownership and analyst following to proxy for information and liquidity effects. We do not, however, make an attempt to fully address the broader question of whether institutions and analysts are informed traders. For a more complete understanding of their informational role, it is important to examine additional measures of information trading, such as, institutional trades and analyst recommendations (Beneish, 1991; Chan & Lakonishok, 1995). Second, our sample is admittedly small, both in time (one month) and in the cross section of stocks (58). It is clearly important to extend the sample period and increase the number of stocks to assess the robustness of the non-monotonic relationship between market liquidity and institutional holdings and analyst coverage documented in this study.

4. CONCLUSION

Institutional investors and financial analysts promote both information-motivated and liquidity-motivated trading in securities. The widely documented evidence that only a few institutional investors and advisors outperform the market suggests that a small fraction of them produce and trade on valuable private information. This leaves the remainder to promote added visibility and stimulate liquid trading in the securities they follow. However, there exists mixed evidence on the overall effects of institutional ownership and analyst coverage on market liquidity. The extant literature points out that the relationship between market liquidity and the number of informed traders may be non-monotonic.

Drawing from this literature, we examine whether the liquidity effects vary as a function of the level of institutional holdings and analyst following. Specifically, we investigate the following testable implications. First, we test whether the effects of a change in institutional holdings and analyst following on market liquidity are ambiguous for *neglected firm* stocks, stocks with low institutional holdings and analyst coverage. Second, we test whether liquidity increases with an increase in institutional ownership and analyst following for *institutional favorites*, driven by intense competition among informed traders and added liquidity-motivated trading. Third, we test whether market liquidity is lower at the market *open* (*close*), particularly for *neglected firm* stocks with more durable private information.

We test these conjectures by constructing matched samples of firms with varying levels of institutional ownership and analyst coverage. Empirical tests based on March 1995 transactions data show that *institutiona*l *favorites* experience a significantly lower adverse selection cost relative to their matched counterparts. However, when institutional holdings and analyst coverage increase from the low to the medium level, we find no significant reduction in the adverse selection cost for *neglected firms*. Moreover, we find evidence of higher adverse selection cost at the market *open*, particularly for stocks with lower institutional holdings and analyst coverage.

Thus, our results provide support for a non-monotonic relation between the number of informed traders and market liquidity. They suggest that institutional investors and financial analysts promote market liquidity, but only at the high level of institutional equity ownership and analyst coverage. At the low and medium levels, these variables have, on average, an insignificant effect on market liquidity.

NOTES

1. Several studies show that institutions now own approximately 50 cents of every dollar of common stock outstanding.
2. As in many other studies, we use institutional equity ownership and analyst following as proxies for informed-trading. Easely and O'Hara (1987) suggest block trades as an alternative proxy for informed trading. In contrast, Seppi (1990) observes that an informed institution may choose either a block order at a good price or a sequence of anonymous market orders at less favorable prices or both. Empirical evidence in Holthausen, Leftwich and Mayers (1987) suggests that the information effect of block trades is small. Therefore, it is not clear that block trades are a better proxy than institutional ownership and analyst following for informed trading.
3. Our choice of the matched pair methodology is similar in spirit to that of Easley (1998).
4. See Affleck-Graves, Hegde and Miller (1994) for evidence that the adverse selection component of spread varies across exchanges.
5. The tradeoff between sample size and controlling for extraneous information indicators is considerable. The most onerous sacrifice in sample size comes from matching firms, ranked by institutional interest *and* analyst following, by market capitalization and SIC industry code. Note that all but two of the *institutional favorite* stocks are included in the S&P 500 Index. Only two of the *institutional favorite* control group firms are included in this broad-based market index. None of the *neglected firms* are members of this broad-based market index.
6. In Brennan and Hughes (1991), brokerage commissions decrease with the price level of a stock and neglected firms split their stocks to attract more analysts. This imposes more adverse selection risk on their stockholders.
7. Coffee (1991) offers an alternative view. He argues that institutional investors collect information primarily to comply with fiduciary rules, rather than to trade on private information. Moreover, institutions are bound by law to invest in relatively liquid securities.
8. The difference in market value of equity is under 25% between the matched pairs.
9. Easley, O'Hara and Paperman (1998) justify the matched pair methodology on the grounds that it does not require the specification of the functional form between adverse selection costs and the number of informed traders, with respect to trading volume. This means that the matched pair technique is useful for non-liner and simultaneous models.
10. Our measures of institutional participation are approximate because Rule 13f-1 of the Securities Exchange Act of 1934 requires only those managers with investment discretion over $100 million in equities to report their holdings.
11. We are grateful to I/B/E/S International, Inc. for providing data on the number of analysts.
12. For example, none of the visibility measures are highly correlated for *institutional favorites*. In stark contrast, analyst following and the number of

institutional owners are highly positively correlated for the *neglected firm* control group while correlation for this same pair of variables for *neglected firms* is negative.

13. In addition, we weed out trades that are settled outside the normal procedure such as those that deviate more than 25% in price from adjacent trades. We also exclude trades and quotes with obvious reporting errors such as zero or negative prices (quantities).

14. Specifically, we match each trade with a quote that precedes it by at least five seconds.

15. We use logarithms of the transaction price and quote midpoint data to adjust for the discrete nature of quotations.

16. Formal tests for non-normality were run on adverse selection cost data for each firm. The evidence was overwhelming, in favor of a non-normal distribution.

17. As in Lin, Sanger and Booth (1995), we set any negative estimate of λ to zero. Estimates of λ in excess of 1 are set to unity.

18. In addition, we also conduct tests for differences in trade size but find little significant variation across the three levels of institutional and analyst interest.

19. Foster and Viswanathan (1993) find an intraday U-shaped pattern in λ though the increase in λ at the market close is insignificant. Lin, Sanger and Booth (1995) report a significant drop in λ from the *open* to *midday*, as well as the *midday* to *close*.

20. We also investigated whether there are higher adverse selection costs and lower trading volumes on Mondays vis-à-vis other days of the week (Foster & Viswanathan, 1990, 1993) and found little interday differences in our sample. We suspect the insignificant result is due to the fact that our sample includes only four Mondays.

21. Furthermore, trading activity in March is not biased by the year-end effect. NYSE statistics suggest that the March trading volume is representative of the remaining months in 1995.

22. These estimates are taken from the 50 to 75 percentile of trade size. See Table 3 in the Lin, Sanger and Booth paper.

23. Following Glosten and Harris (1988), Brennan and Subrahmanyam (1995) estimate λ by regressing transactions price changes on the signed order flow. Jenning et al. (1997) and Sarin et al. (1996) employ the George, Kaul and Nimalendran (1991) procedure which employs a cross-sectional regression of estimated spreads on the quoted spreads to estimate λ. (This method relies on first estimating the serial covariance of the difference in trade-to-trade returns and the subsequent bid-to-bid returns.) In addition, some of these researchers obtain other estimates of adverse selection cost but none make use of the Lin, Sanger and Booth procedure.

24. For example, Lin, Sanger and Booth report average daily trading volume, deflated by firm size of 0.15%, 0.18% and 0.19%, respectively, for the low, medium and high trading volume samples. In our sample, the corresponding values are 0.06%, 0.12% and 0.28% for *neglected firms*, the two control groups and the *institutional favorites*, respectively.

ACKNOWLEDGMENTS

We thank two anonymous referees, Thomas O'Brien, Gim Seow, seminar participants at the University of Connecticut, Sacred Heart University, 1997 meeting of the Southern Finance Association, and 1999 meeting of the Eastern Finance Association, respectively, for valuable comments and suggestions.

REFERENCES

Admati, A. R., & Pfleiderer, P. (1988). A Theory of Intraday Patterns: Volume and Price Variability. *Review of Financial Studies*, *1*, 3–40.

Affleck-Graves, J., Hegde, S. P., & Miller, R. E. (1994). Trading Mechanisms and the Components of the Bid-Ask Spread. *The Journal of Finance*, *49*, 1471–1488.

Beneish, M. D. (1991). Stock Prices and the Dissemination of Analysts' Recommendations. *Journal of Business*, *64*, 393–416.

Brennan, M. J., & Hughes, P. A. (1991). Stock Prices and the Supply of Information. *Journal of Finance*, *46*, 1665–1691.

Brennan, M. J., & Subrahmanyam, A. (1995). Investment Analysis and Price Formation in Securities Markets. *Journal of Financial Economics*, *38*, 361–381.

Brennan, M. J., & Subrahmanyam, A. (1998). The Determinants of Average Trade Size. *The Journal of Business*, *71*, 1–25.

Chan, L. K. C., & Lakonishok, J. (1995). The Behavior of Stock Prices Around Institutional Trades. *Journal of Finance*, *50*, 1147–1174.

Chung, K. H., McInish, T. H., Wood, R. A., & Wyhowski, D. J. (1995). Production of Information, Information Asymmetry, and the Bid-Ask Spread: Empirical Evidence from Analysts' Forecasts. *Journal of Banking and Finance*, *19*, 1025–1046.

Clements, J. (1998). The Truth Investors Do Not Want to Hear on Index Funds and Market Soothsayers. *The Wall Street Journal*, May 12, C1.

Coffee, J. C. Jr. (1991). Liquidity versus Control: The Institutional Investor as Corporate Monitor. *Columbia Law Review*, *91*, 1277–1368.

Easley, D., & O'Hara, M. (1987). Prices, Trade Size and Information in Security Markets. *Journal of Financial Economics*, *19*, 69–90.

Easley, D., Kiefer, N. M., O'Hara, M., & Paperman, J. (1996). Liquidity, Information, and Infrequently Traded Stocks. *Journal of Finance*, *51*, 1405–1437.

Easley, D., O'Hara, M., & Paperman, J. (1998). Are Financial Analysts Informed Traders? *Journal of Financial Markets* (forthcoming).

Foster, F. D., & Viswanathan, S. (1993). Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models. *Journal of Finance*, *48*, 187–211.

George, T. J., Kaul, G., & Nimalendran, M. (1991). Estimation of the Bid-Ask Spread and Its Components: A New Approach. *Review of Financial Studies*, *4*, 623–656.

Glosten, L. R., & Harris, L. E. (1988) Estimating the Components of the bid/ask spread. *Journal of Financial Economics*, *21*, 123–142.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum Strategies, Portfolio performance and Herding: A Study of Mutual Fund Behavior. *American Economic Review*, *85*, 1088–1105.

Holden, C. W., & Subrahmanyam, A. (1992). Long-Lived Private Information and Imperfect Competition. *Journal of Finance*, *47*, 247–270.

Holthausen, R., Leftwich, R., & Mayers, D. (1987). The Effect of Large Block Transactions on Security Prices. *Journal of Financial Economics*, *19*, 237–267.

Jennings, W. W., Schnatterly, K., & Seguin, P. J. (1997). Institutional Ownership, Information and Liquidity. Working paper, University of Michigan.

Kothare, M., & Laux, P. (1995). Trading Costs and the Trading System for Nasdaq Stocks. *Financial Analysts Journal*, *51*, 42–53.

Lee, C. M. C., & Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance*, *46*, 733–746.

Lin, J. C., Sanger, G. C., & Booth, G. G. (1995). Trade Size and Components of the Bid-Ask Spread. *Review of Financial Studies*, *8*, 1153–1183.

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance*, *42*, 483–510.

New York Stock Exchange (1996). *Fact Book: 1995 Data* (New York Stock Exchange, Inc.).

Sarin, A., Shastri, K. A., & Shastri, K. (1996). Ownership Structure and Stock Market Liquidity. Working Paper, University of Pittsburgh.

Seppi, D. J. (1990). Equilibrium Block Trading and Asymmetric Informatuion. *Journal of Finance*, *45*, 73–94.

Shleifer, A., & Vishny, R. W. (1986). Large Shareholders and Corporate Control. *Journal of Political Economy*, *94*, 461–488.

Siconolfi, M. (1995). Out of Sync: This Securities Analyst Calls It as He Sees It: That's the Problem. *The Wall Street Journal*, May 18, A1.

Subrahmanyam, A. (1991). Risk Aversion, Market Liquidity, and Price Efficiency. *Review of Financial Studies*, *3*, 417–441.

APPENDIX A: SAMPLE DESCRIPTION OF INSTITUTIONAL FAVORITES AND THEIR CONTROL GROUP FIRMS

This appendix identifies each firm included in our sample, its SIC code, percent of equity capital held by institutions, number of institutional investors and number of analysts.

Matched Pair	Firm Name	SIC Code	Institutional Holdings (%)	Number of Institutions	Number of Analysts
	Panel A:		*Institutional*	*Favourites*	
1	Becton Dickinson	3841	87	529	19
2	Illinois Central	4011	91	312	15
3	Ryder System	7513	93	403	14
4	Brinker International	5812	83	344	23
5	TJX Companies	5651	85	300	14
6	General Signal	3823	85	329	14
7	Black & Decker	3546	80	405	17
8	ASARCO	3331	84	341	14
9	Federal Paper Board	2631	81	328	12
10	Liz Claiborne	2339	90	384	18
11	Scientific-Atlanta	3663	84	367	14
12	Bethlehem Steel	3312	87	334	18
13	Louisiana Land	2911	86	332	28
	Panel B:	*Favorites'*	*Control*	*Group*	
1	Boston Scientific	3841	47	278	10
2	Frontier Corporation	4813	44	252	12
3	ADT Limited	7382	58	140	6
4	Morrison Restaurants	5812	46	172	9
5	Family Dollar Stores	5531	56	149	7
6	Sunbeam Corporation	3634	49	250	11
7	Case Corporation	3523	51	125	12
8	Owens-Illinois	3221	52	118	11
9	Potlatch Corporation	2631	51	267	12
10	Shaw Indus.	2273	57	267	12
11	Harsco Corporation	3489	56	200	8
12	Vulcan Materials	3281	58	174	9
13	Bemis Company	2671	49	295	9

APPENDIX B: SAMPLE DESCRIPTION OF NEGLECTED FIRMS AND THEIR CONTROL GROUP FIRMS

This appendix identifies each firm included in our sample, its SIC code, percent of equity capital held by institutions, number of institutional investors and number of analysts.

Matched Pair	Firm Name	SIC Code	Institutional Holdings (%)	Number of Institutions	Number of Analysts
	Panel A:		*Neglected*	*Firms*	
1	Spelling Entertainment	7812	13	78	2
2	Circle K Corporation	5411	12	50	4
3	Petrie Stores	5621	22	90	1
4	United Water Res.	4941	15	75	3
5	Laclede Gas	4923	15	55	1
6	SGS-Thomson N.V.	3674	10	61	2
7	Elsag Bailey Process	3823	0	0	4
8	Toy Biz	3944	11	32	3
9	Sola International	3851	17	36	1
10	Mikasa Inc.	3263	12	35	4
11	Falcon Building	3563	16	23	3
12	BIC Corporation	3591	0	0	2
13	International Specialty	2819	15	60	4
14	Manville Corporation	2631	15	84	2
15	Zapata Corporation	2077	12	38	0
16	Marvel Entertainment	2721	0	0	4

APPENDIX B continued

Matched Pair	Firm Name	SIC Code	Institutional Holdings (%)	Number of Institutions	Number of Analysts
		Panel B:	*Neglected*	*Firm Control*	*Group*
1	Comdisco Inc.	7377	52	170	9
2	Service Merchandise	5399	51	132	8
3	Morrison Restaurants	5812	46	172	9
4	ONEOK Inc.	4923	48	244	5
5	Rollins Environ Sv.	4959	57	191	10
6	Boston Scientific	3841	47	278	10
7	AMETEK Inc.	3823	58	151	5
8	CML Group	3949	54	153	7
9	Sunrise Medical	3842	51	175	10
10	Medusa Corporation	3241	54	112	6
11	Varco International	3533	53	123	6
12	Teleflex Inc.	3728	58	170	9
13	Valspar Corporation	2851	51	124	11
14	Potlatch Corporation	2631	51	267	12
15	Hartmarx Corporation	2311	50	167	8
16	Shaw Indus.	2273	57	267	12

EUROPEAN STOCK MARKETS: AN ERROR CORRECTION MODEL ANALYSIS

Asim Ghosh and Ronnie J. Clayton

ABSTRACT

The integration of European capital markets has increased over the last two decades. Government-imposed barriers to the flow of capital across European countries have been gradually reduced. Furthermore, the development and growth of derivative securities and the idea of having a unified common currency, the European Currency Unit, have stimulated financial integration. Applying the theory of cointegration in case of European stock markets, this study investigates whether the stock prices are predictable. The empirical evidence suggests that the indices are pair wise cointegrated and hence predictable during the period investigated. The appropriate error correction model is estimated and is then used to perform out-of-sample forecasting. The Root Mean Square Error (RMSE) from the error correction model is compared with that of a naive model. Using the error correction model reduces the RMSE, on average, across the European Stock Market Indices by almost 19%.

1. INTRODUCTION

Economies world-wide have become more interdependent. The country in which a firm is founded and headquartered has less importance as many of

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 95–107.

ISBN: 0-7623-0798-6

these firms have significant international operations and look to the international activity to generate much of their revenue. Operating internationally offers diversification of markets and revenue base that is not attainable in their domestic economy. Europe has seen particularly impacted as the move toward the European Union took root and grew. Government-imposed barriers to the flow of capital across European countries have been gradually reduced. The development and growth of derivative securities and the idea of having a unified common currency, the European Currency Unit, have stimulated financial integration. The natural result is that the integration of European capital markets has increased over the last two decades. Applying the theory of cointegration in case of European stock markets, this study investigates whether the stock prices in European markets are predictable.

Understanding the integration of financial markets in various locations in the global economy is important because of our continued drive to reduce risk for the level of return that we expect to receive or to increase expected return for the risk that we are willing to take. Capital Market Theory tells us that diversification is the primary way to accomplish the risk/return goal. There is debate, however, as to whether investors need to seek diversification in the international arena.

A. The Debate Over International Diversification

Burton Malkiel vs. The Senior Chairman of Vanguard Group, John Bogle, Malkiel points out the benefits of international diversification and supports including foreign stocks as part of any investment strategy.[1] Bogle argues that, while foreign investments may be acceptable, they are not necessary in light of the recent performance of the stock markets in the United States. If one takes a short-term approach, the United States markets have, indeed, provided significantly greater return than foreign markets. However, over a longer period of time stock markets in the United States have provided, essentially, the same average return as that provided in foreign markets.[2] General knowledge tells us that international diversification can be beneficial in terms of portfolio risk and/or return. It is in investors' interest to better understand the relationship between various equity markets. This study concentrates on the equity markets in Europe.

Economic theory posits that certain pairs of financial time series are expected to move together in the long-run. In the short-run they may deviate from each other, but investors' tastes and preferences, market forces and government intervention will bring them back to their equilibrium.

B. The Literature

Kasa (1992) investigates the international stock markets and finds a single common stochastic trend in the U.S., Japan, England, Germany and Canada. Corhay, Rad and Urbain (1995) examine the Pacific-Basin stock markets of Australia, Japan, Hong Kong, New Zealand and Singapore. They do not find evidence of a common stochastic trend among these markets. Kwan, Sim and Cotsomitis (1995) study the stock markets of Australia, Hong Kong, Japan, Singapore, South Korea, Taiwan, the U.K., the USA and Germany. Their evidence suggests that these markets are not efficient in weak form. They find the existence of significant lead-lag relationships between equity markets. Hung and Cheung (1995) investigate the stock markets of Hong Kong, Korea, Malayasia, Singapore and Taiwan and find no evidence of cointegration. Freimann (1998) examines the correlation of stock market indices throughout Europe and finds that these markets, in general, have shown increased correlation in more recent years.

C. Dominant Economies

If the European equity markets have become more integrated, the relationship may be manifested through one or more dominant economies and markets. Knowledge of the dominant markets will help portfolio managers to decide upon the proportion of their portfolio to allocate to a particular economy to enhance the risk/return characteristics. The present study investigates whether the stock markets of Ireland, Norway, Sweden, Denmark, France, the U.K. and Germany exhibit cointegration and are, thus, predictable. Evidence presented in this study supports this hypothesis implying that an effort to predict the stock index of one country based on the stock market index of an economically dominant country will prove useful. The premise of predictability is reinforced by means of out-of-sample forecasting.

The paper is organized as follows. Section 2 briefly examines the concept of cointegration and error correction. Section 3 presents the data, the model, a discussion of unit roots and the theory of cointegration. Section 4 examines the empirical evidence. This research shows that the stock market indices of Ireland (IR), Sweden (SW), Denmark (DN), Norway (NW), France (FR), Switzerland (SZ), United Kingdom (U.K.), and Germany (DM) are all integrated processes. The stock indices of IR, SW and DN are individually cointegrated with U.K. while the stock indices of NW, FR and SZ are cointegrated with DM. An appropriate error correction model is constructed and it is shown to be statistically significant and potentially useful for forecasting the stock market indices of IR, NW, SW, DN, FR, and SZ. Section 5 concludes the paper.

2. RELATIONSHIPS AMONG ECONOMIC TIME SERIES

Nonstationarity and Cointegration

Economic theory is based on equilibrium relationships among a set of variables. Statistical analyses are applied to financial time series to investigate such relationships. Classical statistical inference will be valid if the series are stationary. However, most financial time series are not stationary. Any regression model developed using nonstationary variables will be misspecified. Financial economists difference the variables in the model to achieve stationarity and use the differenced series in statistical analysis. Valid statistical inference is achieved, but at the cost of losing valuable long-run information.

The concept of cointegration introduced by Granger (1981) and further developed by Engle and Granger (1987), incorporates the presence of nonstationarity, long-term relationships and the short-run dynamics in the modeling process. The notion of cointegration is briefly described below while a more detailed description can be found in several sources, including Engle and Granger (1991), Davidson and MacKinnon (1993), Banerjee, Dolado, Galbraith and Hendry (1993) and Hamilton (1994). A financial time series is said to be integrated of order one i.e. I(1), if it becomes stationary after differencing once. If two series are integrated of order one, they may have a linear combination that is stationary without requiring differencing. If the linear combination of two such series is stationary, they are cointegrated according to Engle and Granger. When regression analysis is involved, ignoring non-stationarity can lead to spurious regression (see Granger and Newbold (1974)). In the following analysis, it is shown that nonstationarity, long-run relationships and short-run deviations are integrated within a dynamic specification framework.

3. DATA AND METHODOLOGY

Data

The data used in this study consist of monthly value-weighted stock market indices of IR, NW, SW, DN, FR, SZ, U.K., and DM and all are expressed in terms of U.S. Dollars. The data are obtained from DSC Data Services, Inc.

covering the time period January 1960 through June 1994. After eliminating six missing observations 408 observations are used for investigation. The natural logarithms of the indices are used for estimation purposes. We employ 375 observations for model estimation and the remaining 33 are used for out-of-sample forecasting.

Methodology

If a financial time-series becomes stationary after first differencing it is said to by integrated of order one, i.e. I(1). Let us consider two time series x_t and y_t, which are both I(1). Usually, any linear combination of x_t and y_t will be I(1). But if there exists a linear combination $z_t = y_t - \alpha - \beta\, x_t$ which is I(0), then x_t and y_t are cointegrated according to Engle and Granger with the cointegrating parameter β. Cointegration links the long-run relationship between integrated financial series, to a statistical model of those series.

To test whether the two indices are cointegrated, it is necessary to test that each index is I(1). Testing for unit roots is conducted by performing the augmented Dickey-Fuller (ADF) (1981) regression, which can be written as:

$$\Delta y_t = a_0 + a_1 y_{t-1} + \sum_{i=1}^{p} a_i \Delta y_{t-1} + \varepsilon_t \tag{1}$$

where p is selected large enough to ensure that the residuals ε_t are white noise. For sufficiently large values of p, the ADF test loses its power. An alternative test proposed by Phillips and Perron (PP) (1988), which allows weak dependence and heterogeneity in disturbances, is performed using the following regression:

$$y_t = b_0 + b_1 y_{t-1} + u_t \tag{2}$$

where u_t is serially correlated.

Testing for cointegration is undertaken once it is found that each series contains one unit root. Test statistics utilize residuals from the following cointegrating regression:

$$C_t = a + bF_t + e_t \tag{3}$$

where C_t and F_t are the regressand and the corresponding regressor. For example if IR is the regressand then U.K. is the regressor. If NW is the regressand then DM is the regressor and similarly.

If the two series are cointegrated, then e_t will be I(0). The ADF test is performed on the estimated residuals, e_t, from Eq. (3):

$$\Delta e_t = ae_{t-1} + \sum_{j=1}^{q} \phi_j \Delta e_{t-j} + v_t \tag{4}$$

where q is large enough to make v_t white noise. The estimated residuals are also subject to the following PP test:

$$e_t = \alpha + \beta e_{t-1} + \gamma_t \tag{5}$$

where γ_t is serially correlated.

Once it is established that the series are cointegrated, their dynamic structure can be exploited for further investigation. Engle and Granger's bivariate framework is used for several reasons.[3] From conceptual point of view it will allow us to understand the influence of a dominant regional market force (such as Germany and U.K.). Our presumption is the equity markets of IR, SW, and DN are primarily driven by the equity market of U.K. However, the equity market of DM may have non-negligible effect. Likewise the equity markets of NW, FR, and SZ are expected to be dominated by the equity market of DM. Engle and Granger show that cointegration implies and is implied by the existence of an error correction representation of the indices involved. Error correction model (ECM) abstracts the short- and long-run information in the modeling process. The appropriate ECM to be estimated is given

$$\Delta C_t = a_0 + ae_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta F_{t-i} + \sum_{j=1}^{m} \delta_j \Delta C_{t-j} + u_t \tag{6}$$

where n and m are large enough to make u_t white noise.

Engle and Granger propose a two-step estimation procedure for the estimation of the parameters of model (6). First, C_t is regressed on F_t and the residuals are collected from model (3) by using the ordinary least squares (OLS). The ECM with the appropriate specification of the dynamics is estimated by the OLS in the second stage. The appropriate values of n and m are chosen by the Akaike information criterion (AIC) (1974).

The existence of an error correction model implies some Granger causality between the series, which means that the error correction model can be used for forecasting. The error correction model is expected to provide better forecasts than the ones from a naïve model. The forecasting performance of the error correction model is compared to that of a benchmark naive model by means of root mean squared error (RMSE). For forecasting this naive model used the most recent information available concerning actual value. The forecasting equation is

$$F_{t+i} = x_t \tag{7}$$

where F_{t+i} = forecast for period $t + i$

t = present period

i = number of periods ahead being forecast

x = latest actual value (for period t)

4. EMPIRICAL ANALYSIS

All the stock indices are tested to ensure they are I(1). The results of the ADF and the PP tests are shown in Table 1. The level series demonstrate that they

Table 1. Augmented Dickey-Fuller (ADF) and Phillips Perron (PP) Tests for Unit Roots in Monthly Stock Indices[a].

	Stock Indices		Critical Value (10%)
	ADF	PP	
Levels:			
IR	–0.36	–0.58	–2.57
NW	–1.45	–1.47	–2.57
SW	–1.94	–1.91	–2.57
DN	–2.21	–1.94	–2.57
FR	–1.39	–1.39	–2.57
SZ	–1.43	–1.81	–2.57
U.K.	–1.96	–2.16	–2.57
DM	–0.02	–0.36	–2.57
First Difference :			
IR	–5.72	–17.01	–2.57
NW	–4.94	–18.60	–2.57
SW	–4.84	–20.32	–2.57
DN	–5.45	–18.99	–2.57
FR	–4.92	–19.03	–2.57
SZ	–5.54	–17.22	–2.57
U.K.	–5.02	–17.58	–2.57
DM	–4.44	–16.03	–2.57

[a] Tests are conducted using autoregressive representations of the monthly stock indices of IR, NW, SW, DN, FR, SZ, U.K., and DM covering the time period January 1960 through June 1994 (408 observations). Critical values are taken from MacKinnon (1991).

have a unit root in their autoregressive representations. This evidence indicates that the series are nonstationary. Now the difference series are checked for the presence of a unit root. The ADF and the PP tests clearly reject the null hypothesis of the presence of a unit root. This implies that the difference series are I(0). Therefore, the stock indices are I(1) for all indices.

Since it is established that each series is I(1), it is imperative to test whether there exists a linear combination of two corresponding indices that is I(0). If there exists a long-run relationship, they must be cointegrated. Results of the tests of cointegration are presented in Table 2. The ADF and PP tests reject the null hypothesis of no cointegration at the 10% level of significance except the ADF test for SW, and DN. This observation reinforces the notion that cointegration unites the long-run relationship between the relevant variables.

Cointegration implies that the series have an error correction representation and, conversely, an ECM implies that the series are cointegrated (Engle & Granger). The ECM (6) provides a better representation of the stochastic dynamic relationship between the series by enlarging the information set. For example, the last period's equilibrium error is incorporated through the error correction term. Short-run deviations in one period are adjusted through lagged variables in the next period.

Table 3 presents the estimates of the parameters of model (6) for IR, NW, SW, DN, FR, and SZ. The intercepts from model (6) are found to be statistically significant for most indices. This appears to imply the presence of a linear trend in the data generating process. The error correction term is negative and statistically significant. This evidence is consistent with the theory

Table 2. Cointegration Between European Stock Market Indices[a].

	Regressor				Critical Value(10%)*
Regressand:	U.K.		DM		
	ADF	PP	ADF	PP	
IR	–3.26	–3.82	–	–	–3.04
NW	–	–	–3.44	–3.58	–3.04
SW	–2.97	–3.43	–	–	–3.04
DN	–2.71	–3.07	–	–	–3.04
FR	–	–	–3.51	–3.56	–3.04
SZ	–	–	–3.49	–3.56	–3.04

[a] Augmented Dickey-Fuller (ADF) and Phillips Perron (PP) tests for cointegration between the logarithms of (IR, U.K.), (SW, U.K.), (DN, U.K.), (NW, DM), (FR, DM), and (SZ, DM) covering the period January 1960–June 1994. Critital values are taken from MacKinnon (1991).

i.e. the error correction coefficient is supposed to be negative. If the change in the corresponding regressand (such as ΔIR_t) is above its average value, the error correction term is positive. In this case ΔIR_t will move downward to follow the long-run equilibrium attractor making the coefficient negative. If ΔIR_t is below its average position, the error correction term is negative, however it will move upward to follow the long-run equilibrium attractor and the coefficient will be negative. These coefficients measure the speed with which the system moves towards its equilibrium relationship. Lagged variables of ΔIR, ΔU.K., ΔDM and ΔSZ are statistically significant. These findings show that the deviations in one period are adjusted in the following periods.

Table 3. Error Correction Model Estimation[a].

Regressor	Regressand	
	Coefficient	SE
	ΔIR	
Constant	0.0086	0.0031
e_{t-1}	–0.0664	0.0143
$\Delta U.K._{t-1}$	0.1498	0.0377
ΔIR_{t-1}	–0.1270	0.0497
	ΔNW	
Constant	0.0057	0.0036
e_{t-1}	–0.0326	0.0115
	ΔSW	
Constant	0.0080	0.0036
e_{t-1}	–0.0365	0.0142
	ΔDN	
Constant	0.0070	0.0027
e_{t-1}	–0.0264	0.0106
	ΔFR	
Constant	0.0048	0.0032
e_{t-1}	–0.0429	0.0149
	ΔSZ	
Constant	0.0036	0.0025
e_{t-1}	–0.0514	0.0151
ΔDM_{t-2}	0.1128	0.0525
ΔSZ_{t-2}	–0.1291	0.0544

[a] Estimates of the parameters from the error correction model (6) for IR, NW, SW, DN, FR and SZ covering the time period January 1960 through June 1994.

Estimation of an error correction model implies the existence of causality between the respective changes in the stock indices implying that the stock indices are predictable. The estimated error correction model is used to develop 33 one-step ahead forecasts for ΔIR, ΔNW, ΔSW, ΔDN, ΔFR, and ΔSZ. These forecasts are then compared with naive univariate forecasts. Table 4 presents the summary statistics for these forecasts. The forecast standard deviations of the naive model are higher than those for ECM (6) thus leading to a smaller root mean squared error.

Table 5 presents the ratio of the ECM RMSE to the naïve RMSE. The ECM reduces root mean squared error (RMSE) over the naïve model by 23%, 13%,

Table 4. One-Step-Ahead Forecasts[a].

Statistic	Naïve	ECM
	ΔIR	
Mean	0.0081	0.0151
SD	0.0500	0.0106
RMSE	0.0620	0.0480
	ΔNW	
Mean	0.0108	–0.0009
SD	0.0560	0.0038
RMSE	0.0680	0.0591
	ΔSW	
Mean	0.0092	0.0100
SD	0.0749	0.0044
RMSE	0.0990	0.0735
	ΔDN	
Mean	–0.0006	0.0089
SD	0.0473	0.0032
RMSE	0.0569	0.0462
	ΔFR	
Mean	0.0068	0.0002
SD	0.0378	0.0040
RMSE	0.0485	0.0413
	ΔSZ	
Mean	0.0127	–0.0020
SD	0.0391	0.0091
RMSE	0.0522	0.0433

[a] Summary statistics for 33 one-step-ahead forecasts for IR, NW, SW, DN, FR and SZ covering the time period January 1960 through June 1994.

Table 5. RMSE Ratio.

RATIO	ECM/Naïve
ΔIR	0.77
ΔNW	0.87
ΔSW	0.74
ΔDN	0.81
ΔFR	0.85
ΔSZ	0.83
Average	0.81

26%, 19%, 15% and 17% for ΔIR, ΔNW, ΔSW, ΔDN, ΔFR, and ΔSZ respectively. On average the ECM reduces RMSE almost 19% over the naïve model showing that it improves the out-of-sample forecasting accuracy. This evidence establishes the dominance of model (6) over model (7).

To summarize, the evidence presented in this paper demonstrates that the ECM (6) is better than the model (7) in predicting the change in indices. Out-of-sample forecasts reinforce the supremacy of model (6). This evidence appears to support the premise that the stock prices are predictable in European countries and may present attractive investment opportunities.

5. CONCLUSION

This study investigates whether the European stock market prices can be forecast using the theory of cointegration. Monthly stock index series are used in the analysis. Each series is tested for the presence of a unit root in its autoregressive representation. It is found that each series is integrated of order one. Each index is tested for the existence of a long-run equilibrium relationship in a bivariate framework. It is then found that each market index is cointegrated with one of the dominant market indices (either U.K. and DM).

Cointegration implies and is implied by the existence of an error correction model. As mentioned earlier, ECM integrates the short- and the long-run information in modeling the data and proves to be a superior modeling technique compared to the naive method. ECM would be more effective in forecasting the stock market indices. Investors can benefit in designing their trading strategies by using this framework.

The dominance of model (6) over model (7) is established by means of out-of-sample forecasts. ECM reduces the RMSE of the stock indices for all the countries investigated in this study by a considerable margin. Knowledge that cointegration exists and that there is a predictable relationship between a dominant index and other indices will aid portfolio managers as they seek to allocate investment funds in the various markets. Appropriate allocation should allow these managers to exploit diversifying risk/return relationships to enhance their portfolios' performance. This framework has the potential for applying in various other stock markets in the world.

NOTES

1. See the *Wall Street Journal* article written by Jonathan Clements entitled "Two Pros Clash on Merit of Foreign Stocks," May 18, 1999, page C1.
2. Ibbotson Associates, Chicago, IL, indicates that over the past five years the average annual return on United States equity securities has outpaced foreign equity security returns by a ratio of approximately three-to-one (18.1% to 5.9%). However, taking a longer-term perspective, since 1970, the average annual return on United States equity securities is equal to the equity return provided in foreign markets (13.7%).
3. The approach of Engle and Granger is used rather than that of Johansen (1988) to examine the relationship of the dominant economies with those that they dominate. In such a bivariate framework, the results of the two procedures are identical.

REFERENCES

Akaike, H. (1974). A new look at statistical model identification. *IEEE Trans. Auto. Control*, *19*, 716–723.

Banerjee, A., Dolado, J., Galbraith, J., & Hendry, D. (1993). *Co-Integration, Error Correction, and the Econometric Analysis of Non-stationary Data*. New York: Oxford University Press.

Corhay, A., Rad, A., & Urbain, J. (1995). Long run behavior of Pacific-Basin stock prices. *Applied Financial Economics*, *5*, 11–18.

Davidson, R., & MacKinnon, J. (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Dickey, D., & Fuller, W. (1981). The likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*, 1057–1072.

Engle, R., & Granger, C. (1987). Cointegration and error correction representation, estimation, and testing. *Econometrica*, *55*, 251–276.

Engle, R., & Granger, C. (1991). *Long-run Economic Relationships Readings in Cointegration*. New York: Oxford University Press.

Freimann, E. (1998). Economic Integration and Country Allocation in Europe. *Financial Analysts Journal*, (September/October).

Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, *16*, 121–130.

Granger, C. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, *48*, 213–228.

Granger, C., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 111–120.

Hamilton, J. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hung, B., & Cheung, Y. (1995). Interdependence of Asian Emerging Equity Markets. *Journal of Business Finance & Accounting*, *22*, 281–288.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, (June/September).

Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics*, *29*, 95–124.

Kwan,A., Sim, A., & Cotsomitis, J. (1995). The causal relationships between equity indices on world exchanges. *Applied Economics*, *27*, 33–37.

MacKinnon, J. (1991). Critical values for cointegration tests. In: R. Engle & C. Granger (Ed.), *Long-run Economic Relationships Readings in Cointegration*. New York: Oxford University Press.

Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*, 335–346.

This Page Intentionally Left Blank

ALTERNATIVE METHODS FOR ROBUST ANALYSIS IN EVENT STUDY APPLICATIONS

Lisa A. Kramer

ABSTRACT

A variety of test statistics have been employed in the finance and accounting literatures for the purpose of conducting hypothesis tests in event studies. This paper begins by formally deriving the result that these statistics do not follow their conventionally assumed asymptotic distribution even for large samples of firms. Test statistics exhibit a statistically significant bias to size in practice, a result that I document extensively. This bias arises partially due to commonly observed stock return traits which violate conditions underlying event study methods. In this paper, I develop two alternatives. The first involves a simple normalization of conventional test statistics and allows for the statistics to follow an asymptotic standard normal distribution. The second approach augments the simple normalization with bootstrap resampling. These alternatives demonstrate remarkable robustness to heteroskedasticity, autocorrelation, non-normality, and event-period model changes, even in small samples.

1. INTRODUCTION

This paper focuses on test statistics underlying financial event studies. In such studies, one analyzes the information content of corporate events, making use

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 109–132.

ISBN: 0-7623-0798-6

of stock returns for a collection of firms. The goal is to determine whether a particular financial event, such as an equity issue, debt offering, merger, or regulation change, had a significant effect on firms' returns. Evidence of a significant effect may shed light on various questions including the impact on shareholder welfare, market efficiency, or the effectiveness of government intervention. Techniques have evolved from the seminal event study of stock splits by Fama, Fisher, Jensen and Roll (1969) to the point that event study methods are a common source of "stylized facts" which influence policy decisions and the direction of research. Recent surveys of the state of the art of event study methods are provided by MacKinlay (1997) and Binder (1998).

Common characteristics of returns data can seriously compromise inference based on event study methods, a point that has not gone unnoticed in the literature. De Jong, Kemna and Kloek (1992, p. 29) report that "results obtained under the usual assumptions on the error process (homoskedastic, normal distribution) shows that ignoring the fat tails and the heteroskedasticity may lead to spurious results." Campbell and Wasley (1993, p. 74) find that with daily NASDAQ returns, conventional test statistics "depart from their theoretical unit normal distribution under the null hypothesis." Several studies have provided modifications to conventional techniques which successfully address some of the concerns that arise in practice.[1] This study study contributes to the advancement of event study methodology by providing a feasible means of effectively dealing with a large set of problems under a very wide range of conditions.

The main components of the paper are as follows. In Section 2, I provide a simple correction to conventional test statistics that ensures they follow their assumed asymptotic distribution. In Section 3, I demonstrate that in finite samples, the characteristics of returns data lead to bias in the conventional test statistics, compromising inference. The corrected test statistics show little or no bias. In Section 4, I present a two-step bootstrap procedure that also yields test statistics robust to the problems that plague conventional methods. Conclusions follow.

2. EVENT STUDY METHODS: CONVENTIONAL AND NEW

There are two broad goals in conducting financial event studies: testing for a significant *information effect* in stock returns at the time of the event announcement (examples include Schipper and Thompson (1983), Malatesta and Thompson (1985) and Eckbo (1992)) and identifying factors which *determine* the information effect (see, for example, work by Eckbo, Maksimovic and Williams (1990) and Prabhala (1997)). The first of these, testing for an information effect, is the focus of this study.

In testing for an information effect, one collects a series of consecutive stock returns for a sample of firms of interest along with the corresponding returns on a market index portfolio. A market model is then estimated for each of the firms, and tests are conducted to see if there is evidence of an impact on firms' stock returns at the time of the event. There are several possible approaches available for formulating test statistics to detect the information effect. The most flexible and most commonly adopted of these is based on a multivariate regression model including a dummy variable to pick up the event. The methodology is clearly laid out by Thompson (1985) and is briefly summarized as follows.[2]

Define N as the number of firms being considered. R_{it} is the return on firm i's share where $i = (1, \cdots, N)$, M_{it} is the return on the market index portfolio, and ε_{it} is an error term. The point in time at which the event announcement potentially impacts the firms' returns is denoted $t = +1$, hence a dummy variable is defined to equal 1 for $t = +1$ and zero otherwise.[3] The dummy variable picks up the unanticipated portion of the return, i.e. the event effect. For each of the N firms being considered, a market model is estimated over a time period of length T such as $t = (-130, \cdots, +10)$, *including* the date of the event and several days following the event:[4,5]

$$R_{it} = \beta_{i0} + \beta_{i1} M_{it} + \beta_{iD} D_{it} + \varepsilon_{it}, \qquad i = (1, \cdots, N). \tag{1}$$

Estimating the market model produces a t-statistic, t_i, for each of the N estimated dummy variable coefficients $\hat{\beta}_{iD}$. These are used for testing the null hypothesis of no abnormal event day returns:

$$Z = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}}. \tag{2}$$

Typically, the Z statistic is assumed to be distributed as standard normal. In fact, standard normality would follow only if the t_i were themselves independent and identically distributed as normal with mean zero and unit variance, which they are not in practice. At best, the t_i are Student t-distributed with a variance that differs from unity.[6] As a result, Z does not approach standard normal, even as the number of firms, N, is increased.

To achieve asymptotic standard normality, Z must be normalized by the theoretical standard deviation of the t_i. In practice, with returns displaying even a small degree of heteroskedasticity, non-normality, etc., the t_i have a standard deviation that differs from the theoretical standard deviation. Thus the sample standard deviation of the t_i must be used to normalize Z. The standard deviation of the t_i is defined as:

$$\hat{\sigma}_N = \sqrt{\frac{\sum_{i=1}^{N} (t_i - \bar{t}\,)^2}{N-1}}$$

where $\bar{t}$ is the mean of the t_i. Then the normalized test statistic is obtained by dividing Z by $\hat{\sigma}_N$:

$$\tilde{Z} = \frac{Z}{\hat{\sigma}_N}. \tag{3}$$

This normalized test statistic, $\tilde{Z}$, follows an asymptotical standard normal distribution (as $N \rightarrow \infty$). In practice, $\tilde{Z}$ is very robust relative to Z. This is detailed directly below.

3. IMPACT OF RETURNS CHARACTERISTICS ON TEST STATISTIC PROPERTIES

In this section, I quantify the extent to which many well-documented characteristics of the data used in event studies compromise the use of event study test statistics. As I show below, neglecting features of the data such as heteroskedasticity, autocorrelation, non-normality, and changes in event-period variance can lead to test statistics which do not follow their assumed distribution, even as N grows quite large. That is, the properties of returns data compromise finite sample inference. Experiments investigating the size of test statistics are reported immediately below, and the power experiments follow in Section 4.

Examining the validity of standard event study techniques under circumstances commonly encountered in practice requires the use of data that are generated to display properties that closely match those of actual returns data. The use of simulated data in establishing the properties of test statistics is conventional and profuse. In the event study literature, simulated data is employed by authors including Acharya (1993) and Prabhala (1997). Thus I first employ simulated data in this paper, which facilitates a clear analysis of the marginal impact of many features of the data. For comparative purposes, at the end of Section 4, I also provide results from experiments employing actual CRSP data as used by Brown and Warner (1980, 1985) and others.

In investigating size, i.e. the test statistic's behavior when there is no event present, I consider a sample of $N = 30, 50, 100$, or 200 firms estimated over a time period of $T = 141$ days. The experimental design is as follows.

1. Disturbances, market returns, and model parameters were generated for each firm, and then each of the N firms' returns were generated according to a basic market model:

   $$R_{it} = \beta_{i0} + \beta_{i1} M_{it} + \varepsilon_{it}$$

   where $t = (-130, \cdots, +10)$ and $i = (1, \ldots, N)$.

   The properties ascribed to the generated data closely matched those of actual financial returns data. Values were chosen to conservatively mimic what is observed in actual returns as documented in the appendix. The unanticipated return series, ε_{it}, was generated with a standard deviation of 0.77, skewness of 0.15, kurtosis of 6.2, an event period jump in variance of 100%. The coefficient on market returns was generated with an event period increase of 100% and firms' autocorrelation coefficient was generated to be 0.1.
2. OLS was used to estimate each firm's market model, yielding N event-day dummy variable t-statistics. Then Z was computed according to Eq. (2) and $\tilde{Z}$ was computed according to Eq. (3).

   $$Z = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}} \tag{2}$$

   $$\tilde{Z} = \frac{Z}{\hat{\sigma}_N} \tag{3}$$

Recall that $\hat{\sigma}_N$ is the standard deviation of the t_i statistics.

3. Steps 1 and 2 were repeated a total of 1000 times. Each of these 1000 replications can be thought of as a single event study, each generating the Z and $\tilde{Z}$ statistics. By simulating many event studies, strong statements can be made about the behavior of the event study tests.
4. The statistical size of the Z and $\tilde{Z}$ statistics was then evaluated. First, actual rejection rates were computed at various levels of significance, then each actual rejection rate was compared to the assumed "nominal" rejection rate (size) for the test. For example, for a hypothesis test conducted at the $\alpha = 1\%$ level of significance, under the null hypothesis of no abnormal return, the test statistic should indicate rejection 1% of the time if the statistical size of the test is correct. Thus, the actual rejection rate for the 1% test would be compared to its nominal size of 0.01. Standard confidence intervals were constructed around the nominal size to see if actual rejection rates differ significantly from those expected under the null hypothesis. With a large number of replications, strong conclusions can be reached because the

confidence intervals around the nominal size values for each statistic are quite small.

Panel A of Table 1 contains results of this experiment for various sample sizes. The first column lists the α level for conventional tests of significance, 1%, 5%, and 10%. The remaining columns contain results for samples of 30, 50, 100, and 200 firms respectively. The value listed in each cell is the actual rejection rate for the Z statistic of Eq. (2) at the particular significance level. Rejection rates which differ significantly from their nominal size are denoted with asterisks.

Table 1. Test Statistic Properties: Non-Normal Data with DGP Changes. Rejection Rates for Tests at Common Significance Levels. 1000 Replications, Various Numbers of Firms

Panel A: Z Statistic Using the Standard Normal Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.043***	0.040***	0.044***	0.054***
0.05	0.102***	0.098***	0.110***	0.119***
0.10	0.161***	0.158***	0.181***	0.191***
Panel B: $\tilde{Z}$ Statistic Using the Standard Normal Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.009	0.006	0.010	0.012
0.05	0.054	0.045	0.054	0.059*
0.10	0.104	0.096	0.107	0.113*
Panel C: $\tilde{Z}$ Statistic Using the Bootstrap Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.014	0.013	0.012	0.012
0.05	0.059*	0.048	0.057	0.059*
0.10	0.109	0.095	0.111	0.117**
Panel D: Z Statistic Using the Bootstrap Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.011	0.009	0.007	0.011
0.05	0.052	0.046	0.054	0.060*
0.10	0.109	0.095	0.110	0.113*

* Significant at the 10% level
** Significant at the 5% level
*** Significant at the 1% level

Two things are remarkable about these results. First, the conventional Z statistic dramatically over-rejects for any significance level, α, and for any number of firms considered. The rejection rates are anywhere from 60% to 530% higher than they are expected under the null which, in practice, could cause a researcher to conclude that an event effect is present when there is none. Second, the degree of over-rejection is not reduced by increasing the sample size, N. This is not surprising given the Z statistic requires a normalization to follow an asymptotic (in N) standard normal distribution, as discussed in Section 2.

Panel B of Table 1 contains details on the experiment for the case of $\tilde{Z}$. Examining this portion of the table reveals that implementing the normalization eliminates the bias relative to the traditional method. For example, when Z was compared to the assumed standard normal distribution for the case of 50 firms, a rejection rate of 9.8% was observed although a rejection rate of 5% was expected under the null hypothesis. Under similar conditions, the $\tilde{Z}$ statistic rejection rate is 4.5% (which is not statistically different from 5%). The difference in performance of the Z and $\tilde{Z}$ statistics arises because of the normalization which permits $\tilde{Z}$ to follow an asymptotic standard normal distribution.

All of the previously discussed results are for the case where data simultaneously display a variety of commonly observed characteristics. An advantage of using simulated data is the ability to examine the marginal impact of particular features of the data individually and at various levels of intensity (for example, with different degrees of skewness and/or kurtosis, with higher or lower levels of event-period variance changes, with characteristics that vary across firms, etc.). Kramer (1998) reports the outcome of such experiments. The results indicate that conventional test statistics are significantly biased even when the characteristics of the data are less extreme relative to the data considered above. Event-period variance changes, autocorrelation, non-normality, and other common features of the data are each individually capable of invalidating inference, even when present at very low levels relative to what is conventionally observed in the actual returns data used in event studies. Significantly biased test statistics are observed under a wide range of circumstances, even when very large samples of firms are employed.

It is important to emphasize that the severe problems plaguing the Z statistic are not solved by increasing the number of firms in the sample. It has been shown above that in order to achieve asymptotic standard normality, Z statistics require a normalization for non-unit variance. Without this normalization, test statistics display no improved performance as N, the number of firms, is increased.

4. THE TWO-STAGE BOOTSTRAP APPROACH

The aim of the above Monte Carlo analysis has been to demonstrate the practical repercussions of using a conventional event study test statistic. The significant bias of the conventional approach was eliminated by using a normalized version of the commonly employed test statistic. In this section, I present another alternative method which involves two stages.

The first stage involves normalizing the conventional Z statistics as shown in the previous section:

$$\tilde{Z} = \frac{Z}{\hat{\sigma}_N} \tag{3}$$

where $\hat{\sigma}_N$ is the standard deviation of the t_i statistics. The second stage involves using bootstrap resampling to generate the empirical distribution of the normalized test statistic. This alternative procedure will also be shown to demonstrate improved performance relative to traditional methods.

The Bootstrap

Use of the bootstrap involves repeatedly sampling from the actual data in order to empirically estimate the true distribution of a test statistic. This method was initially introduced by Efron (1979) as a robust procedure for estimating the distribution of independent and identically distributed data. Since its inception, the bootstrap's performance under a variety of conditions has been examined in depth in the statistics literature. Work by Liu (1988) establishes the suitability of adopting the bootstrap under conditions most applicable to this setting: that of independent but not necessarily identically distributed observations. Provided the random observations are drawn from distributions with similar means (but not necessarily identical variances) and provided the first two moments are bounded, use of the bootstrap is valid. Several books and articles on the subject provide a good overview of the use of the bootstrap for empirical work in many fields, including Hall (1992), LePage and Billard (1992), Efron and Tibshirani (1993), Hjorth (1994), Li and Maddala (1996) and Manly (1998).

Some of the many applications in finance include Malliaropulos' (1999) paper on mean reversion, a study of mutual fund performance by Cai, Chan and Yamada (1997), Kothari and Shanken's (1997) study of expected real stock market returns, an investigation of long-horizon security price performance by Kothari and Warner (1997), Malliaropulos' (1996) study of the predictability of long-horizon stock returns using the bootstrap, Liang, Myer and Webb's (1996) bootstrap estimation of the efficient frontier for mixed-asset portfolios, and an investigation of long-horizon predictability of exchange rates by Mark (1995). In the context of event studies, Marais (1984) uses bootstrapped p-values to conduct inference in conjunction with the standardized residual approach, and Larsen and Resnick (1999) use the bootstrap with cross-sectional stochastic dominance analysis. Given the abundant use of the bootstrap in finance, economics, and beyond, I consider the marginal impact of the bootstrap in the context of event studies.

Implementation

In Fig. 1, I provide a diagrammatic representation of the two-stage bootstrap approach for conducting inference in event studies. The procedure is also discussed in detail below. While the explanation makes use of the Z statistic which emerges from a regression framework employing a dummy variable, the steps can be modified in a straightforward manner to enable inference based on any common event study test statistic, as shown by Kramer (1998).

1a. Estimate the appropriate event study market model for each of the N firms in the sample. The simplest possibility is as follows:

$$R_{it} = \beta_{i0} + \beta_{i1} M_{it} + \beta_{iD} D_{it} + \varepsilon_{it}, \qquad i = (1, \cdots, N). \tag{1}$$

The estimation yields N t-statistics: one for each firm's estimated dummy variable coefficient. As shown in Fig. 1, this collection of t-statistics forms the pool of data upon which the conventional Z statistic is based.

$$Z = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}} \tag{2}$$

A researcher interested in conducting conventional inference would stop at this point and compare the value of the Z statistic to a critical value from the assumed standard normal distribution. As indicated earlier, that distribution does not apply, even asymptotically as N is increased.

1b. Normalize the Z statistic obtained in Step 1a to account for the fact that its variance differs from unity in practice. First, compute the sample standard deviation of the t_i:

$$\hat{\sigma}_N = \sqrt{\frac{\sum_{i=1}^{N} (t_i - \bar{t})^2}{N - 1}}$$

where $\bar{t}$ is the mean of the t_i statistics. Then, divide Z by $\hat{\sigma}_N$ to yield the normalized version of Z:

$$\tilde{Z} = \frac{Z}{\hat{\sigma}_N}. \tag{3}$$

In the next stage of this method, the empirical distribution of $\tilde{Z}$ will be constructed, facilitating reliable inference.

2a. Under the null hypothesis, the distribution of $\tilde{Z}$ is centered about zero, and hence the empirical distribution must be constructed such that it is also

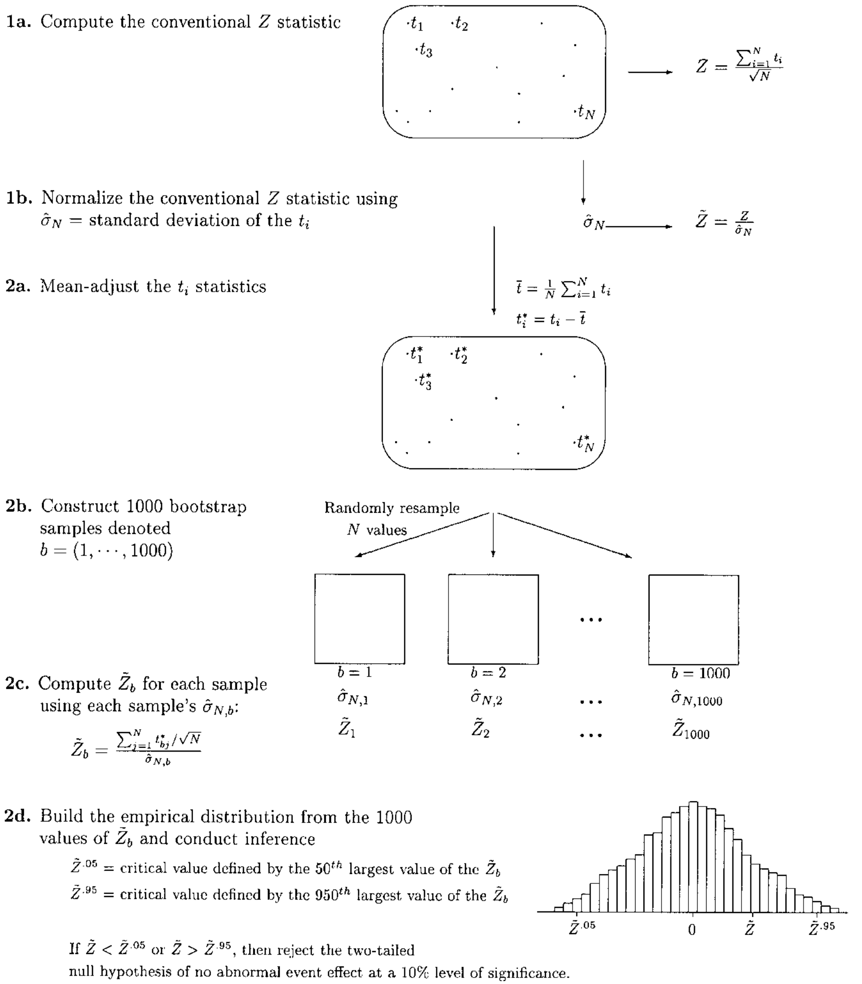


Fig. 1. The Two-Stage Bootstrap Approach.

centered about zero. Notice that the N actual t-statistics calculated in Step 1a have a mean, $\bar{t}$, which generally differs from zero. If these t-statistics were used directly to build the empirical distribution, the result would be a distribution centered about the actual value of $\tilde{Z}$. This would occur because in the absence of imposing the null distribution, the distribution of the sample would be replicated, with the sample mean exactly in the middle. Therefore, prior to constructing the empirical distribution, the t-statistics must be adjusted to impose the null hypothesis of no event day abnormal returns (i.e. zero mean). Accordingly, a collection of *mean-adjusted* t-statistics, denoted t_i^* is assembled by deducting $\bar{t}$ from each of the individual t-statistics: $t_i^* = t_i - \bar{t}$.

The N mean-adjusted t-statistics are, of course, mean zero, and they constitute the collection of statistics – the population – from which bootstrap samples are drawn in the next step. Having mean-adjusted the t-statistics, the empirical distribution will be centered about zero, allowing one to conduct inference under the null hypothesis of no event.

2b. The mean-adjusted data are used to construct an empirical distribution for $\tilde{Z}$ under the null. This involves drawing many random samples, called "bootstrap samples," from the population of t_i^* statistics. As shown in Fig. 1, a single bootstrap sample is constructed by randomly drawing with replacement N observations from the collection of t_i^* statistics. A total of 1000 such bootstrap samples, individually denoted $b = (1, \cdots, 1000)$, are constructed, with each bootstrap sample containing N observations.[7] The particular composition of each bootstrap sample varies randomly. For example, the first sample might contain duplicate occurrences of some of the t_i^* statistics and no occurrences of other t_i^* statistics, the second sample might contain duplicate occurrences of some different t_i^* statistics, etc. The make-up of each of the 1000 constructed bootstrap samples is determined purely by chance.

2c. The construction of the normalized test statistic for each bootstrap sample is parallel to that of the $\tilde{Z}$ statistic shown earlier. Compute Z_b for each of the 1000 bootstrap samples, where the $b = (1, \cdots, 1000)$ subscript is used to specify the particular bootstrap sample and the $j = (1, \cdots, N)$ subscript distinguishes between particular observations within a given bootstrap sample:

$$Z_b = \frac{\sum_{j=1}^{N} t_{b,j}^*}{\sqrt{N}}. \tag{4}$$

Define the mean of the $t_{b,j}^*$ statistics in a particular bootstrap sample as $\bar{t}_b^*$. The standard deviation of the $t_{b,j}^*$ for each bootstrap sample is:

$$\hat{\sigma}_{N,b} = \sqrt{\frac{\sum_{j=1}^{N} (t^*_{b,j} - \bar{t}^*_b)^2}{N-1}}.$$

Each of the 1000 Z_b statistics of Eq. (4) is then normalized by the corresponding $\hat{\sigma}_{N,b}$, yielding 1000 $\tilde{Z}_b$ statistics:

$$\tilde{Z}_b = \frac{Z_b}{\hat{\sigma}_{N,b}} \tag{5}$$

2d. Ordering the collection of 1000 $\tilde{Z}_b$ statistics from smallest to largest defines the empirical distribution. The histogram depicted at the bottom of Fig. 1 is an example of such an empirical distribution. Inference is conducted by comparing the $\tilde{Z}$ statistic from Step 1b to critical values from the empirical distribution. For example, with 1000 bootstrap samples, a 5% left-tail critical value, $\tilde{Z}^{0.05}$, is the 50th largest value of the $\tilde{Z}_b$ statistics and a 5% right-tail critical value, $\tilde{Z}_b^{0.95}$, is the 950th largest of the $\tilde{Z}_b$ statistics. If the value of the $\tilde{Z}$ statistic happens to be larger than 95% of the bootstrap $\tilde{Z}_b$ statistics (i.e. exceeding $\tilde{Z}_b^{0.95}$) or smaller than 5% of the bootstrap $\tilde{Z}_b$ statistics (i.e. falling below $\tilde{Z}^{0.05}$), one rejects at the 10% level of significance the two-sided null hypothesis of no abnormal returns.

To summarize, applying Steps 1a–2d of the two-stage bootstrap approach based on the conventional Z statistic basically involves computing Z using the actual t_i statistics and normalizing it with the variance of the t_i to impose unit variance. This yields $\tilde{Z}$. The t_i statistics are then mean-adjusted to form a population of statistics, the t_i^*, from which random re-samples are drawn. 1000 bootstrap samples are formed, each containing N observations, and $\tilde{Z}_b$ is computed for each bootstrap sample. The collection of all the $\tilde{Z}_b$ statistics defines the empirical distribution. Finally, event study inference is conducted by comparing $\tilde{Z}$ to critical values from the empirical distribution.[8]

Size

Using the experimental design of Section 3, I now explore the performance of the two-stage approach. As before, data were generated with skewness of 0.15, kurtosis of 6.2, an increase in event period variance of 100%, an increase in the event period market return coefficient of 100%, and an autocorrelation coefficient of 0.1. 1000 replications were conducted for computing the rejection rates for the test statistics, and 1000 bootstrap samples were drawn. Results indicate the two-stage approach eliminates the bias of conventional test statistics shown in Section 3, and the statistical power of the two-stage approach compares favorably with conventional methods.

Panel C of Table 1 contains the results for the $\tilde{Z}$ statistic using bootstrapped critical values. As with the normalized test statistic shown in Panel B, virtually no bias remains. Extensive sensitivity analysis (unreported) indicates that the performance of the technique is robust to perturbations in the properties ascribed to the generated data. The bootstrap approach maintains its unbiased size when applied to data displaying a very wide range of commonly observed properties.

Another alternative method one might consider is to simply take the conventional Z statistic and compare it a distribution constructed using the bootstrap procedure in the spirit of Marais (1984) (without first normalizing for non-unit variance). Results from these experiments, shown in Panel D of Table 1 indicate such an approach has properties similar to the other alternative methods discussed in this paper.

Power

The power of the event study test statistics is evaluated by employing data generated to have a positive abnormal return at the time of the event. Abnormal returns of 0.5%, 0.7%, and 0.9% and a sample size of 50 firms are considered here. According to convention, I document size-adjusted power.[9] Results are shown in Table 2.

Table 2. Power Comparisons.
Rejection Rates for Tests at Common Significance Levels.
1000 Replications, 50 Firms

Abnormal Return	Sig. Level (α)	(A) Z Using Std. Normal	(B) $\tilde{Z}$ Using Std. Normal	(C) $\tilde{Z}$ Using Bootstrap	(D) Z Using Bootstrap
0.009	0.01	1.000	0.999	0.999	1.000
	0.05	1.000	1.000	1.000	1.000
	0.1	1.000	1.000	1.000	1.000
0.007	0.01	0.988	0.986	0.962	0.990
	0.05	0.998	0.999	0.996	0.999
	0.1	1.000	0.999	0.999	0.999
0.005	0.01	0.778	0.810	0.737	0.822
	0.05	0.946	0.954	0.929	0.948
	0.1	0.986	0.987	0.976	0.988

Column A contains the rejection rates based on comparing the Z statistic to its commonly assumed standard normal distribution, Column B contains rejection rates based on comparing the $\tilde{Z}$ statistic to the standard normal distribution, and Column C contains rejection rates based on comparing the $\tilde{Z}$ statistic to its bootstrapped distribution. Column D pertains to the case of comparing the conventional Z statistic to a bootstrap distribution. The first case considered is that of adding abnormal returns of 0.9% on the event day. As shown in the top panel of Table 2, all of the test statistics reject almost 100% of the time in this case. When abnormal returns of 0.7% are added on the event day, shown in the middle panel of the table, results are qualitatively similar. Rejection rates for conventional inference and the inference based on the alternative approach remain very close to 100%. With abnormal returns of 0.5%, shown in the bottom panel, the conventional rejection rates and the rejection rates for the alternative approaches fall slightly, but the performance is qualitatively similar across methods. The overall conclusion to draw is that power under the alternative approaches is very comparable to that of conventional methods. Rejection rates are almost identical for all cases considered. Similar results are obtained under various conditions.

Other Nonparametric Methods

In Kramer (1998), I evaluated the properties of nonparametric methods including the rank test of Corrado (1989) and the sign test. While both the rank test and the sign test demonstrated less bias than conventional Z test statistics, significant bias remained, even for samples with large number of firms. A sample size of 1000 firms was required before the majority of the rejection rates were insignificantly different from what would be expected from an unbiased test statistic.

Thus, for inference in samples of fewer than 1000 firms (and in general), I advocate use of one of the alternative approaches documented in this paper.[10] The outstanding performance of these approach applied to data exhibiting a wide range of the characteristics of actual returns is testament to their robustness. Valid and powerful inference is facilitated, even in situations where conventional methods fail.

Use of Actual Returns Data in Monte Carlo Experiments

The experiments documented earlier in this paper make use of simulated data. Some past researchers have employed *actual* returns data for investigating the performance of event study test statistics. Thus, for comparative purposes, I replicated the above experiments using actual returns data in place of simulated data. I followed the approach laid out by Brown and Warner (1985) and Boehmer, Musumeci and Poulsen (1991), creating portfolios of firms by randomly resampling from actual CRSP data. The results from these experiments are qualitatively identical to those presented above.

For example, in exploring event-induced variance, I once again considered the impact of a 100% increase with various sample sizes. Results appear in Table 3. As before, Panel A applies to the conventional Z statistic using a

Table 3. Test Statistic Properties: Using Actual Data.
Rejection Rates for Tests at Common Significance Levels.
1000 Replications
100% Increase in Event Period Variance

Panel A: Z Statistic Using the Standard Normal Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.033***	0.036***	0.044***	0.048***
0.05	0.106***	0.098***	0.110***	0.117***
0.10	0.172***	0.173***	0.166***	0.172***
Panel B: $\tilde{Z}$ Statistic Using the Standard Normal Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.011	0.012	0.006	0.015*
0.05	0.051	0.044	0.054	0.048
0.10	0.107	0.096	0.103	0.104
Panel C: $\tilde{Z}$ Statistic Using the Bootstrap Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.010	0.012	0.006	0.014
0.05	0.054	0.044	0.054	0.051
0.10	0.112	0.096	0.103	0.100
Panel D: Z Statistic Using the Bootstrap Distribution				
Sig. Level (α)	30 Firms	50 Firms	100 Firms	200 Firms
0.01	0.012	0.015*	0.011	0.015*
0.05	0.056	0.052	0.059*	0.054
0.10	0.108	0.106	0.105	0.105

* Significant at the 10% level
** Significant at the 5% level
*** Significant at the 1% level

standard normal distribution, Panel B is for the $\tilde{Z}$ statistic compared to a standard normal, Panel C is for the $\tilde{Z}$ statistic using the bootstrap distribution, and Panel D is for the case of the Z statistic compared to a bootstrap. Samples of 30, 50, 100, and 200 firms were considered, and 1000 replications were conducted. Results shown in the table indicate that the conventional Z statistic is significantly biased while the three alternative methods are mostly unbiased. Unreported experiment results based on other commonly encountered levels of increased variance (50% and 200%) and considering other typical characteristics of the data were qualitatively similar.[11]

5. CONCLUSIONS

There are several different analytic techniques commonly used in event studies – both parametric and nonparametric. These approaches differ in market model specification and estimation or differ in the calculation of the statistics used for hypothesis testing. A common feature of all of the approaches, however, is that basic underlying conditions must be satisfied for test statistics to have their assumed distribution. These conditions are typically violated in practice, invalidating inference. Monte Carlo results presented above indicate the statistical size of commonly employed test statistics is significantly biased when data exhibit characteristics identical to those observed in actual stock returns.

Researchers attempting to conduct event studies with small samples typically recognize that conventionally assumed distributions may be inappropriate for conducting hypothesis tests, and hence they may attempt to collect data for a larger set of firms. (Event studies on samples of 15 or 20 firms are, nonetheless, fairly common.) In the past, the rationale for such an increase in sample size may have been based on appeals to asymptotic theory. In this paper, however, I have argued that even with large numbers of firms, the asymptotic distribution of conventional test statistics is not valid.

As a solution to such problems, I have presented alternative testing procedures based on normalizing conventional Z statistics and/or empirically estimating their distribution with the bootstrap. I presented evidence to establish the validity of using such approaches on data with properties like those of actual stock returns. The techniques are not prone to exhibit biased size in common situations which render bias in conventional techniques – with no sacrifice of power – hence I recommend their use for hypothesis testing in event studies. In situations where the incidence of non-normality or heteroskedasticity in the data is particularly extreme, it may be prudent to adopt the two-stage bootstrap approach, as in practice it will be most robust to the presence such features of the data.

NOTES

1. For example, Boehmer, Musumeci and Poulsen (1991) propose an alternative event study test statistic to account for event-induced variance, Brockett, Chen and Garven (1994) suggest that event study regression models should account for ARCH and stochastic parameters, and Corhay and Tourani Rad (1996) recommend accounting for GARCH effects. Nonparametric alternatives to conventional methods have also been proposed. For example Marais (1984) uses the bootstrap, Larsen and Resnick (1999) use the bootstrap with cross-sectional stochastic dominance analysis, and Dombrow, Rodriguez and Sirmans (2000) apply Theil's non-parametric regression in the estimation of abnormal returns.

2. Kramer (1998) considers additional event study methods such as the cross-sectional approach and the Boehmer, Musumeci and Poulsen (1991) standardized cross-sectional method. Experiment results indicated that except for the standardized cross-sectional method, all methods exhibit the high degree of problems documented in this paper.

3. I assume single-day event periods here for notational ease. To test the significance of *cumulative* effects over *multiple* event days one simply defines the dummy variable accordingly.

4. Though more complex alternatives are available, in this paper I consider the relatively simple market model used frequently in the literature. See for example, Brown and Warner (1985), MacKinlay (1997) and Binder (1998). This permits me to determine how robust various event study methods are to unintentionally neglected features of actual data. Even with more complex models, the problems I document in this paper can arise because the true data generating process underlying stock returns is unknown to the researcher. In practice, one should always use a model that attempts to avoid mis-specification. In conjunction with a well-specified model, the methods I propose below offer a high degree of robustness.

5. The notation T_i would be used to allow for different length estimation periods across firms. In most cases, T_i is constant across firms, in which case the use of T is unambiguous.

6. If the t_i were distributed as Student t, then the theoretical variance of the t_i would be $\dfrac{T-k}{T-k-2}$ (where k is the number of parameters estimated in Eq. (1); $k=3$ in this case). With Student t-distributed t_i, the distribution of Z would approach standard normal as the length of the time series, T, is increased (assuming changes in the market model parameters can be correctly specified which can prove difficult or impossible). In practice, characteristics of returns data are likely to lead to Z with an unknown distribution.

7. Sources such as Efron and Tibshirani (1993) indicate that 1000 bootstrap samples are sufficient for constructing confidence intervals. I verified this result through extensive experimentation. Increasing the number of bootstrap samples to 10,000 leads to no marked change in results.

8. It is worth emphasizing that use of the bootstrap in this setting requires that the t_i statistics be *independently* distributed. For applications where cross-firm correlation may be present (or for applications where time-series correlation may be present in the case of a multi-day event-period), use of the moving sample bootstrap may be advisable. See Liu and Singh (1992) for details.

9. See, for example, Davidson and MacKinnon (1993) for a description of size-adjusted test statistics.

10. For the two-stage bootstrap approach, the computations for a single event study require little additional CPU time relative to conventional methods, and they can be undertaken with any standard statistical package. Several statistical analysis packages, including Shazam, have the bootstrap method built right in.

11. Conventional test statistics generally showed highly significant bias while the alternative methods showed little or no bias.

12. I verified these researchers' findings myself using actual CRSP data. The reported values appear to be correct with the exception of minor typographical errors. I also verified the magnitude of Kon's reported moment values on other sample periods and found the values to be quite similar with the exception of the time around the crash of 1987.

13. Furthermore, variance estimates used in variously defined versions of the Z statistic embed different assumptions about the behavior of the disturbances. If the estimated market model does not properly account for the time-varying behavior of the disturbances, then different conclusions may be reached with the use of different test statistics, even when employing the same data.

ACKNOWLEDGEMENT

I am grateful to Brad Barber, Glen Donaldson, Espen Eckbo, Wayne Ferson, Burton Hollifield, Mark Kamstra, Maurice Levi, Felice Martinello, Brendan McCabe, Nagpurnanand Prabhala, Ken White, and seminar participants at Simon Fraser University, the University of British Columbia, the University of Toronto, and participants of the Western Finance Association meetings. I thank the editor and an anonymous referee for helpful comments. All remaining errors are my own. This work was financially supported by the Social Sciences and Humanities Research Council of Canada.

REFERENCES

Acharya, S. (1993). Value of latent information: Alternative event study methods. *Journal of Finance*, *48*, 363–385.

Binder, J. J. (1998). The Event Study Methodology since 1969. *Review of Quantitative Finance and Accounting*, *11*(2), September, 111–137.

Boehmer, E., Musumeci, J., & Poulsen, A. B. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, *30*, 253–273.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance. *Journal of Econometrics*, *52*, 5–59.

Boudoukh, J., Richardson, M. P., & Whitelaw, R. F. (1994). A tale of three schools: Insights on autocorrelations of short-horizon stock returns. *Review of Financial Studies*, *7*(3), 539–573.

Brockett, P. L, Chen, H. M., & Garven, J. R. (1999). A new stochastically flexible event methodology with application to proposition 103. *Insurance Mathematics and Economics*, *25*(2), 197–217.

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics*, *8*, 205–258.

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics*, *14*, 3–31.

Cai, J., Chan, K. C., & Yamada, T. (1997). The performance of Japanese mutual funds. *Review of Financial Studies*, *10*(2), 237–273.

Campbell, C. J., & Wasley, C. E. (1993). Measuring security price performance using daily NASDAQ returns. *Journal of Financial Economics*, *33*, 73–92.

Corhay, A., & Tourani Rad, A. (1996). Conditional heteroskedasticity adjusted market model and an event study. *The Quarterly Journal of Economics and Finance*, *36*(4), 529–538.

Corrado, C. J. (1989). A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics*, *23*, 385–395.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.

de Jong, F., Kemna, A., & Kloek, T. (1992). A contribution to event study methodology with an application to the Dutch stock market. *Journal of Banking and Finance*, *16*, 11–36.

Dombrow, J., Rodriguez, M., & Sirmans, C. F. (2000). A complete nonparametric event study approach. *Review of Quantitative Finance and Accounting*, *14*(4), June, 361–380.

Donaldson, R. G., & Hatheway, F. (1994). *An expanded approach to the empirical analysis of illegal insider trading and stock prices*. University of British Columbia manuscript.

Donaldson, R. G., & Kamstra, M. (1997). An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance*, *4*, 17–46.

Eckbo, B. E. (1992), Mergers and the value of antitrust deterrence. *Journal of Finance*, *47*, 1005–1029.

Eckbo, B. E., Maksimovic, V., & Williams, J. (1990). Consistent estimation of cross-sectional models in event studies. *The Review of Financial Studies*, *3*, 343–365.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, *7*, 1–26.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, *34*, 420–429.

Fama, E. F., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*, 1–21.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *48*, 1779–1801.

Gupta, A., LeCompte, R. L. B., & Misra, L. (1993). FSLIC assistance and the wealth effects of savings and loans acquisitions. *Journal of Monetary Economics*, *31*, 117–128.

Hall, P. (1992). *The bootstrap and Edgeworth expansion*. New York: Springer-Verlag.

Hjorth, J. S. U. (1994). *Computer intensive statistical methods: Validation, model selection and bootstrap*. New York: Chapman & Hall.

Kon, S. J. (1984). Models of stock returns – a comparison. *Journal of Finance*, *34*, 147–165.

Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: a time-series analysis. *Journal of Financial Economics*, *44*(2), May, 169–203.

Kothari, S. P., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics*, *43*, March, 301–340.

Kramer, L. A. (1998). Banking on event studies: statistical problems, a new bootstrap solution and an application to failed bank acquisitions. Ph.D. Dissertation, University of British Columbia.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: volume versus GARCH effects. *Journal of Finance*, *45*, 221–229.

Larsen, G. A., Jr., & Resnick, B. (1999). A performance comparison between cross-sectional stochastic dominance and traditional event study methodologies. *Review of Quantitative Finance and Accounting*, *12*, 103–112.

LePage, R., & Billard, L. (1992). *Exploring the limits of bootstrap*. New York: John Wiley & Sons, Inc.

Li, H., & Maddala, G. S. (1996). Bootstrapping time series models. *Econometric Reviews*, *15*, 115–158.

Liang, Y., Myer, F. C. N., & Webb, J. R. (1996). The bootstrap efficient frontier for mixed-asset portfolios. *Real Estate Economics*, *24*, 247–256.

Liu, R. Y. (1988). Bootstrap procedures under some non-iid models. *Annals of Statistics*, *16*, 1696–1708.

Liu, R. Y., & Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In: R. LePage & L. Billard (Eds), *Exploring the Limits of Bootstrap*. New York: John Wiley & Sons, Inc.

MacKinlay, C. A. (1997). Event studies in economics and finance. *Journal of Economic Literature*, *34*, March, 13–39.

Malatesta, P. H., & Thompson, R. (1985). Partially anticipated events: A model of stock price reactions with an application to corporate acquisitions. *Journal of Financial Economics*, *14*, 237–250.

Malliaropulos, D. (1999). Mean reversion in southeast Asian stock markets. *Journal of Empirical Finance*, *6*(4), 355–384.

Malliaropulos, D. (1996). Are long-horizon stock returns predictable? A bootstrap analysis. *Journal of Business Finance and Accounting*, *23*, 93–106.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, *36*, 394–419.

Manly, B. F. J. (1998). *Randomization, bootstrap and Monte Carlo methods in biology.* New York: Chapman & Hall.

Marais, M. L. (1984). An application of the bootstrap method to the analysis of squared standardized market model prediction errors. *Journal of Accounting Research*, *22*, Supplement, 34–54.

Mark, N. C. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review*, *85*(1), March, 201–218.

McCabe, B. (1989). Misspecification tests in econometrics based on ranks. *Journal of Econometrics*, *40*, 261–278.

Nelson, D. B. (1990). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*, 347–370.

Officer, R. R. (1967). The distribution of stock returns. *Journal of the American Statistical Association*, *67*, 807–812.

Prabhala, N. R. (1997). Conditional methods in event-studies and an equilibrium justification for standard event-study procedures. *The Review of Financial Studies*, Spring, 1–38.

Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J., & Mykytka, E. F. (1979). A probability distribution and its uses in fitting data. *Technometrics*, *21*, May, 78–82.

Ramberg, J. S., & Schmeiser, B. W. (1972). An approximate method for generating symmetric random variables. *Communications of the Association for Computing Machinery*, *15*, 987–990.

Ramberg, J. S., & Schmeiser, B. W. (1974). An approximate method for generating symmetric random variables. *Communications of the Association for Computing Machinery*, *17*, 78–82.

Schipper, K., & Thompson, R. (1983). The impact of merger-related regulations on the shareholders of acquiring firms. *Journal of Accounting Research*, *21*, 184–221.

Thompson, R. (1985). Conditioning the return-generating process on firm-specific events: A discussion of event study methods. *Journal of Financial and Quantitative Analysis*, *20*, 151–168.

APPENDIX
FURTHER DETAILS ON EXPERIMENT DESIGN

For the Monte Carlo approach outlined in Section 3 above, non-normality and changes in the Data Generating Process (DGP) were incorporated in the generated data for each firm. Several empirical studies and actual data were consulted in order to choose parameter values that would accurately reflect the properties of actual returns data. The choice of particular parameter values for generating data and the choice of the algorithms employed are motivated below.

- *Non-Normality and Heteroskedasticity*
  There is considerable evidence that returns data are not normally distributed. In particular, skewed and fat-tailed distributions have been documented extensively. For some early evidence of these characteristics in firms' return data, see Mandelbrot (1963), Fama (1965), and Officer (1967). An investigation by Kon (1984) reveals, almost without exception, significant skewness and excess kurtosis (fat tails) among daily returns for 30 individual stocks and 3 standard market indexes. Bollerslev, Chou and Kroner (1992, page 11) state it is "widely recognized that the unconditional price or return distributions tend to have fatter tails than the normal distribution." They also observe that even by accounting for Autoregressive Conditional Heteroskedasticity (ARCH), one may fail to capture all of the non-normality in returns data: "standardized residuals from the estimated models . . . often appear to be leptokurtic." ARCH is a well documented empirical regularity of stock returns data, as evidenced by its voluminous literature (much of

which is cited in the survey by Bollerslev, Chou and Kroner (1992)). Neglecting the time-varying variance returns data can result in non-normality of the disturbances in a market model and hence *Z* statistics which do not follow their assumed distributions, even for large samples.

Thus, in determining properties to assign the data generated for experiments, I consulted past studies such as those by Kon (1984) and Lamoureux and Lastrapes (1990) which document the statistical properties of actual stock returns and market returns data,[12] including mean, variance, skewness, and kurtosis, as well as conditional heteroskedasticity parameters. The intention was to simulate stock returns and market returns data with specific statistical properties that closely match actual data. Kon documents the first four moments over a particular sample period for 30 individual firms traded on the NYSE and for several market indices. His results for the moments are as follows. Positive skew for individual firms was observed in twenty-eight of the thirty cases considered. Of these positive cases, the skewness ranged between 0.0678 and 0.9080. The median of all thirty cases was about 0.32, and most values were between 0.30 and 0.40. Excess kurtosis was observed in all 30 stocks considered by Kon. The range in the kurtosis coefficient was 4.8022 to 13.9385, with a median of about 6.3, and with most values between 5 and 7. The standard deviation of returns exceeded 0.77 for all firms and was observed as high as 2.89. For the experiments conducted in this study, skewness up to 0.15, kurtosis up to 6.2, and a standard deviation of 0.77 were adopted in generating conservative data for firms' disturbances.

- *Autocorrelation*
  The evidence on autocorrelation in stock returns is voluminous. For example, Donaldson and Kamstra (1997) report autocorrelation coefficients for daily data in several international and domestic market indices ranging from 0.07 to 0.3. Likewise, Boudoukh, Richardson and Whitelaw (1994) report autocorrelation coefficients for returns on portolios of U.S. firms ranging from below 0.1 to above 0.5. For the experiments conducted in this study, a conservative autocorrelation coefficient of 0.1 is employed in generating market returns and firms' returns.
- *Changes in the DGP*
  There is considerable evidence that the data generating process for returns can change dramatically during the time of an event. For example, Boehmer, Musumeci and Poulsen (1991) report that most researchers who have investigated event-induced variance changes have found variances can increase anywhere from 44% to 1100% during event periods. Donaldson and Hatheway (1994) also find evidence of variance changes – both increases

and decreases – during periods of insider trading. Of their cases where a *rise* in variance is observed during the event period, the amount of the increase ranges from about 4% to 680%. Failure to model such changes can render the market model specification invalid leading to event study test statistics which do not follow their assumed distributions, even asymptotically.[13] Likewise, several researchers have established that market model coefficients can also undergo changes around the time of the event (or follow some distribution or some pattern over time). De Jong, Kemna and Kloek (1992) find evidence that the coefficient on the market return is not necessarily constant over the estimation period in event studies. They find it often follows a mean-reverting AR process. Donaldson and Hatheway (1994) demonstrate the importance of allowing for changes in the intercept and market return coefficient at the time of the event. They find that the market return coefficient can fall by as much as 106% or rise by as much as 4238% in the collection of firms they consider. Brockett, Chen and Garven (1999) demonstrate the importance of allowing a time-varying stochastic market return coefficient.

Thus, for the experiments conducted in this paper, The event period variance can increase by as much as 100% and the event period market return coefficient can rise by as much as 100%. The change in event period variance was incorporated in the data by re-scaling event period disturbances to have a variance up to 100% larger than that of non-event period disturbances.

- *Generating the Data*

There are many reasonable options for generating non-normal returns data, including the following. Returns can be modeled to incorporate excess kurtosis by using a Student *t*-distribution with low degrees of freedom. (Bollerslev and Wooldridge (1992), for example, use a Student t with 5 degrees of freedom to generate fat-tailed data for their Monte Carlo simulations.) Skewness can be incorporated by making use of asymmetric models of conditional variance, such as the EGARCH model of Nelson (1990) or the Modified GARCH model of Glosten, Jagannathan and Runkle (1993). Alternatively, both skewness and excess kurtosis can be simultaneously incorporated by making use of an algorithm described in Ramberg, Dudewicz, Tadikamalla and Mykytka (1979). (This algorithm is a generalization of Tukey's lambda distribution, and it was developed by Ramberg and Schmeiser (1974, 1975). For an application in the context of a simulation study, see McCabe (1989).) Basically, the Ramberg et al. algorithm allows one to select particular values for the first four moments of a distribution in generating random variates. For experiments reported in this paper, I adopted the Ramberg et al. algorithm to generate returns data with the first four moments matching those of actual data. Results of experiments based on use of the Student *t*-distribution to generate leptokurtic but symmetric data are qualitatively similar.

In this study, parameters for the Ramberg et al. algorithm were selected by consulting various studies (including Kon (1984)) and by examining actual returns data. Disturbance terms for all firms were conservatively generated with kurtosis of 6.2, skewness of 0.15, standard deviation of 0.77, and zero mean. The change in the market return coefficient during the event period was incorporated as follows. During the non-event period, $t = (-130, \cdots, -11)$, the coefficient of market returns was set to equal one, while during the event period, $t = (-10, \cdots, +10)$, the coefficient doubled. A market model intercept of zero was assumed for all firms.

Allowing firms to have different baseline values for the market return coefficient and the market model intercept (matching values observed in practice, for example) does not affect the experiment results. The explanation for this invariance with respect to choice of parameter value relies on the fact that OLS is unbiased. As explained in Section 3, the market returns, M_{it}, and the disturbances, ε_{it}, are generated with particular properties to match actual data. Then setting $\beta_{i0} = 0$ and $\beta_{i1} = 1$, firms' returns, R_{it}, are generated according to the basic market model $R_{it} = \beta_{i0} + \beta_{i1}M_{it} + \varepsilon_{it}$. When R_{it} is regressed on M_{it} and a constant, the OLS estimators for β_{i0} and $\beta_{i1} = 1$ are unbiased. That is the estimated intercept and market return coefficient are equal to the chosen values on average. Thus, restricting the chosen values to be zero and one for all firms is a harmless simplification. See Marais (1984, p. 42) for further details.

A TEST OF A NEW DYNAMIC CAPM

Robert Faff, Robert Brooks and Tan Pooi Fan

ABSTRACT

This paper develops a simple version of a dynamic CAPM by the inclusion of a lagged dependent variable in the market model framework. Our tests use the multivariate approach developed by Gibbons (1982) applied to Australian industry portfolio returns over the period 1974 to 1995. While, our results indicate that neither the static CAPM nor the dynamic CAPM can be rejected, some indirect evidence finds in favour of the dynamic model.

1. INTRODUCTION

The capital asset pricing model (CAPM) still represents a fundamental theoretical paradigm of the finance literature. Despite this, many recent attacks (see *inter alia* Fama & French (1992, 1993, 1995, 1996, 1997); Davis (1994); and He & Ng (1994)) suggest that the CAPM is fundamentally flawed. However, there exist many strong rebuttals of these attacks (see *inter alia* Black (1993); Kothari, Shanken & Sloan (1995); Pettengill, Sundaram & Mathur (1995); Flannery, Hameed & Harjes (1997); Loughran (1997); Kothari & Shanken (1997); and Clare, Priestley & Thomas (1998)). Thus, resolution of the debate about the CAPM remains fundamentally an empirical question.

As argued by Jagannathan, Kubota and Takehara (1998), the finance literature has developed in three directions in an attempt to discover alternative asset pricing models which may be superior to the static CAPM. Specifically, they identify: (a) consumption-based intertemporal asset pricing models; (b)

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 133–159.

ISBN: 0-7623-0798-6

dynamic versions of the CAPM; and (c) linear factor models. While the consumption based models have struggled to gain strong empirical support, the dynamic versions of the CAPM and linear factor models have had much greater success.[1]

The aim of the current paper is to continue this line of research endeavour by developing and testing a new version of a "dynamic" CAPM. The simplest method to generate a dynamic version of the CAPM is by the inclusion of a lagged dependent variable in the market model. This idea in and of itself is not new. However, while Cartwright and Lee (1987), for example, employ a lagged dependent variable in the context of the market model, they do not explore the implications of such a dynamic specification for the CAPM.

The inclusion of a lagged dependent variable may be justified on both (theoretical) economic and econometric grounds. Hence, our model begins with the idea that one response to markets in which thin trading (trading frictions) is a problem is to augment the standard market model with a lagged dependent variable. This produces what we term a dynamic market model which, when interpreted in a CAPM framework, produces a dynamic version of the CAPM. In this model, which defines both a short-run and a long-run beta, a theoretical restriction on the dynamic market model can be tested. Following the work of Gibbons (1982) and others, we choose to conduct a multivariate test of the non-linear constraints imposed by the null model using a sample of Australian industry portfolio data. Specifically, we employ monthly data over the sample period 1974 to 1995.

The Australian evidence regarding the CAPM is well represented by Ball, Brown and Officer (1976); Faff (1991); Wood (1991); Faff (1992); Brailsford and Faff (1997); and Faff and Lau (1997). Ball, Brown and Officer (1976) represents the first published test of the CAPM using Australian data. They employed the basic univariate testing methodology in vogue at the time and found evidence supporting the model. Faff (1991) applied the multivariate test proposed by Gibbons (1982) to the zero-beta CAPM over the period 1958 to 1987. His results showed only moderate support for the model. Similarly, Wood (1991) applying the Shanken (1985) multivariate cross-sectional regression approach, also found only weak support for the CAPM. While Faff (1992) was primarily concerned with conducting a multivariate test of the APT, he did run comparative tests against the CAPM. While he found that the APT outperformed the CAPM, neither model seemed to provide an adequate explanation of seasonal mispricing in Australian equities. Brailsford and Faff (1997) test a conditional CAPM in a GARCH-M framework, showing there is some sensitivity to the return measurement interval used. Finally, Faff and Lau

(1997) apply the generalised method of moments (GMM) test of MacKinlay and Richardson (1991) over the period 1974 to 1994 with mixed results.

We use Australian data for two main reasons. First, justification for our dynamic specification of the CAPM, in part, comes from the well known data problem of thin trading. However, it is important for our research design to strike a balance between this consideration and that of having a market that is of sufficient interest to an international audience, as well as having a reliable and well tried dataset. We believe an Australian dataset is ideal for this task. On the one hand it can be seen that a relatively few, very large companies dominate the Australian market. For example, around 40% of market capitalisation and trading value is produced by just 10 stocks, whereas a similar number of the largest U.S. stocks constitute only about 15% of the total U.S. market.[2] Moreover, there are typically prolonged periods in which many smaller Australian companies do not trade.[3] On the other hand, despite the above argument, the Australian equity market belongs to the dominant group of developed markets. For instance, as at the end of 1996 it ranked tenth among all markets in terms of market capitalisation at about $U.S.312,000 million. Interestingly, this is not greatly dis-similar from the size of the Canadian market which ranked sixth.[4]

Second, we join the growing chorus of researchers appealing to the arguments of Lo and MacKinlay (1990) regarding the concern about data snooping in finance research. For example, Jagannathan, Kubota and Takehara (1998, 321–322) argue for the use of Japanese data for their purposes because they wanted ". . . to provide new evidence that cannot be anticipated from the results that are reported in earlier empirical studies and to minimise the biases that arise due to *data snooping*".[5]

The paper is organised as follows: Section 2 outlines the data used; Section 3 provides a detailed explanation of the empirical framework and motivation for the dynamic CAPM; Section 4 details and discusses the results and we finish with a conclusion in Section 5.

2. DATA

In this paper, we analyse the continuously compounded monthly returns on Australian equities, calculated from the Price Relatives File of the Centre from Research in Finance (CRIF) at the Australian Graduate School of Management. Specifically, we report the results based on 24 value weighted industry based portfolios. The proxy chosen for the market portfolio, is the value-weighted Australian market index supplied by the CRIF database. The dataset comprises 264 observations from January 1974 to December 1995.

For the length of sample we consider it is reasonable to conjecture that because of the substantial change in the Australian economy taking place, it is advisable to allow for a potential structural break in our analyses. A number of key events merit investigation. First, the floating of the Australian dollar in December 1983 and second the stock market crash of October 1987. The impact of these events on the beta risk of Australian industries has previously been investigated for the banking sector (see *inter alia* Harper & Scheit (1992), Brooks & Faff (1995), Brooks, Faff & McKenzie (1997)) and for industry portfolios generally (see *inter alia* Brooks & Faff (1997), Brooks, Faff & McKenzie (1998)). The results of this research generally indicate an impact on beta risk, which may generalise to an impact on testing of the CAPM.

Accordingly, to examine the possibility of structural change and also the effect of deregulation in our analysis, the period from January 1974 to December 1995 is partitioned into 3 non-overlapping sub-periods, namely: (a) a pre-float period: January 1974 to November 1983; (b) a post-float, pre-crash period: December 1983 to September 1987; and (c) a post-crash period: November 1987 to December 1995.

3. EMPIRICAL FRAMEWORK

3.1. Market Model Augmented by a Lagged Dependent Variable

To set the stage for our analysis, we begin by examining a standard market model as did Gibbons (1982):[6]

$$R_{it} = a_i + b_i\, R_{mt} + \varepsilon_{it} \tag{1}$$

where R_{it} is the return on the i^{th} asset in period t; a_i is the intercept term; b_i is the estimate of the systematic risk of asset i; R_{mt} is the return on market index in period t and ε_{it} is a mean zero random error term.

It is now standard econometric practice to apply diagnostic tests to estimated regressions and a standard diagnostic test is the Durbin-Watson test. When initially proposed by Durbin and Watson (1950, 1951), the test was designed to detect AR(1) disturbances. However, as documented by King and Evans (1988) and Kariya (1988), the Durbin-Watson test can also be interpreted as a test for MA(1) disturbances, ARMA (1,1) disturbances and the presence of a lagged dependent variable. Thus, it is not clear how a researcher should respond to a significant Durbin-Watson statistic.

A possible response to finding a significant Durbin-Watson statistic is to estimate the standard market model with a lagged dependent variable:

$$R_{it} = a_i + b_i\, R_{mt} + \theta_i R_{it-1} + \varepsilon_{it} \tag{2}$$

This is the specification examined by Cartwright and Lee (1987) using a sample of forty-nine U.S. stocks. An interpretation of the inclusion of a lagged dependent variable relates to the literature on thin trading. For instance, consider a version of the Dimson estimator (Dimson, 1979) which could be based on an infinite lead and an infinite lag of the market return,

$$R_{it} = a_i + \sum_{j=-\infty}^{+\infty} b_{ij} R_{mt+j} + e_{it} \tag{3}$$

If one conceives of the b_{ij} following a geometric progression then solving produces the dynamic market model given in equation [2].[7] Hence, as argued by Cartwright and Lee (1987, p. 136), adopting the lagged dependent variable version of Eq. [2] as a simple means of specifying a dynamic market model, is attractive since it allows us to avoid the issue of choosing the appropriate lag length. A feature of this model is that incorporates a short-run beta and a long-run beta $\left(\frac{b_{i0}}{1-\theta_i}\right)$. Further, as argued by Cartwright and Lee (1987, p. 137):

> . . . in some cases there is friction in the trading process and . . . the long-run coefficient . . . might be a better systematic risk measure than the traditional [static] risk measure . . .

3.2. Stability Tests

Given our earlier arguments regarding the potential for structural breaks over our full sample period, we further extend our analysis of the dynamic market model to incorporate appropriately defined dummy variables to accommodate sub-period analysis:

$$\begin{aligned} R_{it} = \alpha_i &+ b_{1i} D_1 {*}\, R_{mt} + b_{2i} D_2 {*}\, R_{mt} + b_{3i} D_3 {*}\, R_{mt} + b_{4i} D_{CRASH} {*}\, R_{mt} \\ &+ \theta_{1i} D_1 {*} R_{i,t-1} + \theta_{2i} D_2 {*} R_{i,t-1} + \theta_{3i} D_3 {*} R_{i,t-1} + \varepsilon_{it} \end{aligned} \tag{4}$$

where D_1 takes a value of unity for the period from January 1974 to November 1983, and zero otherwise; D_2 takes a value of unity for the period from December 1983 to September 1987, and zero otherwise; D_3 takes a value of unity for the period from November 1987 to December 1995, and zero otherwise; and finally, D_{CRASH} takes a value of unity for the October 1987, and zero otherwise. While somewhat arbitrary, the choice of the breakpoints delineating these subperiods does have some justification. In November 1983 the Australian dollar was floated on foreign exchange markets – commonly seen as the 'signature' event for the deregulation of Australia's financial markets. Further, in October 1987 we have the international stock market crash. Given that our analysis incorporates a number of subperiods, several possible hypotheses arise as to the stability of the slope coefficients over time.

First, consider the set of hypotheses which relate to the short-run beta. We can test for: (a) the equality of the short-run beta between the pre-float and pre-crash periods (H_{01}: $b_1 = b_2$); (b) for the equality between the pre-float and post-crash periods (H_{02}: $b_1 = b_3$); (c) for the equality between the pre-float and crash periods (H_{03}: $b_1 = b_4$); (d) for the equality between the pre-crash and post-crash periods (H_{04}: $b_2 = b_3$); (e) for the equality between the pre-crash and crash periods (H_{05}: $b_2 = b_4$); (f) for the equality between the post-crash and crash periods (H_{06}: $b_3 = b_4$); (g) for the equality across the pre-float, pre-crash and post-crash periods (H_{07}: $b_1 = b_2 = b_3$); (h) for the equality across the pre-float, pre-crash and crash periods (H_{08}: $b_1 = b_2 = b_4$); (i) for the equality across the pre-float, post-crash and crash periods (H_{09}: $b_1 = b_3 = b_4$); (j) for the equality across the pre-crash, post-crash and crash periods (H_{010}: $b_2 = b_3 = b_4$); and (k) for the equality across the pre-float, pre-crash, post-crash and crash periods (H_{011}: $b_1 = b_2 = b_3 = b_4$).

Second, there exists a set of hypotheses which relate to the equality and significance of the lagged dependent variable across the subperiods. Specifically, we can test (l) for the equality of the lagged response across the pre-float and the pre-crash periods (H_{012}: $\theta_1 = \theta_2$); (m) for the equality of the lagged response across the pre-float and the post-crash periods (H_{013}: $\theta_1 = \theta_3$); (n) for the equality of the lagged response across the pre-crash and post-crash periods (H_{014}: $\theta_2 = \theta_3$); (o) for the equality of the lagged response across the pre-float, pre-crash and post-crash periods (H_{015}: $\theta_1 = \theta_2 = \theta_3$); and (p) for the significance of the lagged response across the pre-float, pre-crash and post-crash periods (H_{016}: $\theta_1 = \theta_2 = \theta_3 = 0$).

Third, there exists a set of hypotheses which relate to the long-run beta. Specifically, we can test (q) for the equality of the long-run beta across the pre-float and pre-crash periods (H_{017}: $b_1/(1 - \theta_1) = b_2/(1 - \theta_2)$); (r) for the equality across the pre-float and post-crash periods (H_{018}: $b_1/(1 - \theta_1) = b_3/(1 - \theta_3)$); (s) for the equality across the pre-crash and post-crash periods (H_{019}: $b_2/(1 - \theta_2) = b_3/(1 - \theta_3)$); and (t) for equality across the pre-float, pre-crash and post-crash periods ((H_{020}: $b_1/(1 - \theta_1) = b_2/(1 - \theta_2) = b_3/(1 - \theta_3)$).

3.3. Multivariate Tests of the Static CAPM

While Cartwright and Lee (1987) leave their analysis at the level of the dynamic market model, it is natural to extend it to the case of the dynamic CAPM, analogous to the well researched link between the static versions of the market model and the CAPM. Accordingly, once the preliminary analysis outlined in previous sections has been completed, we apply the multivariate methodology developed by Gibbons (1982) to test the CAPM in both a static and dynamic context. First, we test if the static CAPM can be rejected. The static CAPM can be expressed as:

$$\mathrm{E}(\mathrm{R_i}) = \mathrm{R_f} + \beta_\mathrm{i}[\mathrm{E}(\mathrm{R_m}) - \mathrm{R_f}] \tag{5}$$

and upon rewriting [5] we obtain:

$$\mathrm{E}(\mathrm{R_i}) = \mathrm{R_f}(1 - \beta_\mathrm{i}) + \beta_\mathrm{i}\mathrm{E}(\mathrm{R_m}) \tag{6}$$

If the risk free asset ($\mathrm{R_f}$) does not exist (or is unobservable), we have a zero-beta version of CAPM in which the risk-free rate ($\mathrm{R_f}$), is replaced with the expected return on the zero-beta portfolio (γ). Equation [6] becomes:

$$\mathrm{E}(\mathrm{R_i}) = \gamma(1 - \beta_\mathrm{i}) + \beta_\mathrm{i}\mathrm{E}(\mathrm{R_m}) \tag{7}$$

If we take the expectations of the standard market model (unrestricted model):

$$\bar{R}_i = a_i + b_i \bar{R}_m \tag{8}$$

and compare it to the zero-beta CAPM (Eq. [7]), the well-known nonlinear restriction on the market model coefficients across equations occurs viz:

$$\mathrm{H_0} : \alpha_\mathrm{i} = \gamma(1 - \mathrm{b_i}), \tag{9}$$

where i run across the N assets being studied. This represents the null hypothesis that the CAPM is consistent with the data.

Whilst there are many multivariate tests to choose from, we have decided to use the likelihood ratio test (LRT) statistic (following Gibbons (1982)) adjusted for small sample bias (following Gibbons, Ross and Shanken (1989) and MacKinlay and Richardson (1991)). Specifically, the LRT is defined as follows:

$$\mathrm{LRT} = (\mathrm{T} - \mathrm{N} - 1)^* \ln(\mathrm{DRC_r}/\mathrm{DRC_u}) \tag{10}$$

where T is the number of monthly observations, N is the number of equations and $\mathrm{DRC_r}$ ($\mathrm{DRC_u}$) is the determinant of the contemporaneous variance-covariance matrix estimated from the residuals of the restricted (unrestricted) system. LRT is distributed asymptotically as a chi-square test with (N − 1) degrees of freedom.

3.4. Multivariate Tests of the Dynamic CAPM

Finally, we need complete the framework necessary to test the validity of the dynamic CAPM specification. Taking the expectation of the “dynamic” market model specification of Eq. [2]:

$$\bar{R}_i = a_i + b_i \bar{R}_m + \theta_i \bar{R}_i \tag{11}$$

Equation [11] can be re-written as :

$$(1 - \theta_i)\bar{R}_i = a_i + b_i \bar{R}_m$$

and upon dividing both sides by $(1 - \theta_i)$ we have:

$$\bar{R}_i = \left[\frac{a_i}{1 - \theta_i}\right] + \left[\frac{b_i}{1 - \theta_i}\right]\bar{R}_m \tag{12}$$

Further, if we denote the "full impact" beta (b_i*) as:

$$\left[\frac{b_i}{1 - \theta_i}\right] = b_i^* \tag{13}$$

then Eq. [12] becomes:

$$\bar{R}_i = \left[\frac{a_i}{1 - \theta_i}\right] + b_i^* \, \bar{R}_m \tag{14}$$

Comparing Eq. [7] with [14]:

$$\frac{a_i}{1 - \theta_i} = \gamma(1 - b_i^*)$$

$$a_i = \gamma(1 - b_i^*)[1 - \theta_i] \tag{15}$$

Finally, by substituting Eq. [13] in [15]:

$$a_i = \gamma\left[1 - \left(\frac{b_i}{1 - \theta_i}\right)\right](1 - \theta_i)$$

which simplifies to provide the null hypothesis:

$$H_0 : a_i = \gamma(1 - \theta_i - b_i) i = 1,2, \ldots, N \tag{16}$$

Thus, the restriction implied by the dynamic CAPM can also be tested using the multivariate testing methodology of Gibbons, Ross and Shanken (1989) and MacKinlay and Richardson (1991).[8] We test the validity of the static and dynamic CAPM in a number of different circumstances. Specifically, we test across the full sample period and also our three other subperiods (pre-float, pre-crash, post-crash). Further, we conduct our testing for the full 24 industries, as well as for the five resources industries and the 19 industrials industries separately.

4. RESULTS

Table 1 presents the results for market model estimation outlined in Eq. [1]. As expected, all of the twenty-four ASX Industries displayed statistically significant betas, which range from 0.4329 in the case of Property Trusts to 1.3542 in the case of Gold. Generally, it appears that the resources industries (Industry 1 to 5) are more risky than industrials with an average resources beta of 1.1876 as compared to an average industrials beta of 0.7666. Exceptions to this general rule are Entrepreneurial Investors with a high beta of 1.2091 and solid fuels, with a low beta of 0.8273.

As revealed in the final column of Table 1, the application of the Durbin-Watson test indicates the presence of significant autocorrelation in half of the industries studied. These twelve industries are Gold; Other Metals; Solid Fuels; Oil and Gas; Alcohol and Tobacco; Food and Household Goods; Chemicals; Media; Banks; Insurance; Investment and Financial Services; and Miscellaneous Industrials.

In response to the serial correlation problem obtained from the standard market model, the model is re-estimated augmented by a lagged dependent variable as outlined in Eq. [2] and the result is reported in Table 2. While we find that half of the industries have a statistically significant lagged dependent variable, they are not exactly the same industries which revealed a significant DW statistic in Table 1. Specifically, eight cases coincide, namely, Gold; Other Metals; Solid Fuels; Oil and Gas; Media; Insurance; Investment and Financial Services; and Miscellaneous Industrials. The other four cases are: Developers and Contractors; Engineering; Entrepreneurial Investors; and Miscellaneous Services. For eleven of these industries the coefficient on the lagged dependent variable is significantly positive – the exception being Other Metals with a significantly negative θ coefficient (at the 10% significance level).

Importantly, Table 2 reveals that the application of the Durbin-Watson test to this dynamic market model indicates there is now very little residual autocorrelation. Only two industries (Other Metals and Alcohol and Tobacco) retain significant residual autocorrelation.[9] It should be noted that the Durbin-Watson test is applicable even in the presence of a lagged dependent variable. Inder (1984, 1986) has demonstrated that the test still has good power although its critical values need adjusting.[10]

Table 1. Estimation of the Standard Market Model: 1974: 01 to 1995: 12.

This table reports the results of estimating the market model for 24 industry portfolios over the period 1974 to 1995. We report the OLS estimates of a_i and b_i, plus t-statistics for the hypothesis that the coefficient equals zero in parentheses. Coefficients significant at the 5% level are indicated by **, while coefficients significant at the 10% level are indicated by *. The Durbin-Watson (DW) Test is also reported and if significant, indicated by ***.

ASX Industry	a_i	b_i	DW Test
1: Gold	− 0.0057	1.3542**	1.5534***
	(1.19)	(17.15)	
2: Other Metals	− 0.0053*	1.3309**	1.7136***
	(1.76)	(26.77)	
3: Solid Fuels	0.0061*	0.8273**	1.4976***
	1.78	(14.87)	
4: Oil and Gas	− 0.0054	1.2658**	1.5371***
	(1.36)	(19.4692)	
5: Diversified Resources	0.0006	1.1599**	1.7983
	(0.24)	(28.82)	
6: Developers and Contractors	− 0.0008	1.0442**	1.7631
	(− 0.34)	(28.15)	
7: Building Materials	0.0010	0.8476**	1.9399
	(0.62)	(32.35)	
8: Alcohol and Tobacco	0.0045	0.8331**	1.6320***
	(1.47)	(16.64)	
9: Food and Household Goods	0.0059**	0.6950**	1.7123***
	(2.88)	(20.68)	
10: Chemicals	0.0052**	0.7572**	2.2721***
	(2.04)	(18.25)	
11: Engineering	0.0036*	0.6699**	1.9188
	(1.72)	(19.53)	
12: Paper and Packaging	0.0032	0.7407**	1.9341
	(1.57)	(22.30)	
13: Retail	0.0013	0.7982**	1.7733
	(0.55)	(20.96)	
14: Transport	0.0012	1.0495**	1.7665
	(0.43)	(23.19)	
15: Media	0.0073*	0.8973**	1.6452***
	(1.90)	(14.31)	
16: Banks	0.0032	0.8354**	1.7359***
	(1.22)	(19.60)	
17: Insurance	0.0028	0.6987**	1.6891***
	(0.82)	(12.42)	
18: Entrepreneurial Investors	− 0.0018	1.2091**	1.8285
	(0.40)	(16.48)	

Table 1. Continued.

This table reports the results of estimating the market model for 24 industry portfolios over the period 1974 to 1995. We report the OLS estimates of a_i and b_i, plus t-statistics for the hypothesis that the coefficient equals zero in parentheses. Coefficients significant at the 5% level are indicated by **, while coefficients significant at the 10% level are indicated by *. The Durbin-Watson (DW) Test is also reported and if significant, indicated by ***.

ASX Industry	a_i	b_i	DW Test
19: Investment and Financial Services	0.0019	0.7139**	1.6810***
	(1.02)	(23.41)	
20: Property Trusts	0.0067**	0.4329**	1.8553
	(3.93)	(15.41)	
21: Miscellaneous Services	0.0056**	0.6972**	1.8122
	(2.38)	(18.18)	
22: Miscellaneous Industrials	0.0006	0.6651**	1.5777***
	(0.33)	(20.93)	
23: Diversifies Industrials	0.0024	0.9096**	1.9102
	(1.32)	(30.81)	
24: Tourism and Leisure	0.0066**	0.5055**	1.9248
	(2.37)	(11.12)	
Number of significant test statistics at 5% (10%) level	**5 (8)**	**24 (24)**	

Now consider the final column of Table 2 which reports the implied long term "full-impact" beta across the industry portfolios. An inspection of the results indicates that Property Trusts is the lowest risk industry with a long-run beta of 0.4520 while Gold is the highest risk industry with a long-run beta of 1.5229. Consistent with the results on the short-run beta, resources industries appear to be more risky than industrials industries. The average long-run beta for the resources industries is 1.2737 as compared to 0.8536 for the industrials sector. Consistent with expectations, industries with significant θ coefficients exhibit the largest change between their short-run and long-run beta.

Table 3 reports the subperiod analysis using the model specified in Eq. [4]. The Oil and Gas industry appears to be less risky over time, as its short-run beta is 1.5904 in the pre-float period but has fallen to 0.8950 in the post-crash period. On the other hand, the Media industry appears to be more risky over time, with a short-run beta of 0.5297 in the pre-float period rising to 1.5206 in the post-crash period. The lagged dependent variable is significant in all three

Table 2. Estimation of the Market Model Augmented with a Lagged Dependent Variable.

This table reports the results of estimating the market model augmented with a lagged dependent variable for 24 industry portfolios over the period 1974 to 1995. We report the OLS estimates of a_i, b_i and θ_i, plus t-statistics for the hypothesis that the coefficient equals zero in parentheses. Coefficients significant at the 5% level are indicated by **, while coefficients significant at the 10% level are indicated by *. The Durbin-Watson (DW) Test (if significant, indicated by ***) and implied long term beta (β_{LTi}) are also reported.

ASX Industry	a_i	b_i	θ_i	DW Test	β_{LTi}
1: Gold	− 0.0068	1.3405**	0.1198**	1.7809	1.5229
	(-1.42)	(17.17)	(2.86)		
2: Other Metals	− 0.0048	1.3364**	− 0.0606*	1.6108***	1.2600
	(1.57)	(26.97)	(1.90)		
3: Solid Fuels	0.0039	0.8088**	0.1484**	1.7876	0.9497
	(1.16)	(14.74)	(3.31)		
4: Oil and Gas	− 0.0065	1.2346**	0.1618**	1.7737	1.4729
	(− 1.69)	(19.44)	(4.19)		
5: Diversified Resources	0.0006	1.1597**	0.0029	1.8048	1.1631
	(0.22)	(28.73)	(0.10)		
6: Developers and Contractors	− 0.0016	1.0367**	0.0785**	1.9256	1.1250
	(-0.69)	(28.15)	(2.57)		
7: Building Materials	0.0006	0.8429**	0.0451	2.0281	0.8827
	(0.35)	(32.06)	(1.63)		
8: Alcohol and Tobacco	0.0037	0.8297**	0.0577	1.7417***	0.8805
	(1.20)	(16.58)	(1.34)		
9: Food and Household Goods	0.0056**	0.6930**	0.0240	1.7584	0.7100
	(2.64)	(20.52)	(0.63)		
10: Chemicals	0.0057**	0.7604**	− 0.0437	2.1752	0.7286
	(2.21)	(18.29)	(− 1.06)		
11: Engineering	0.0022	0.6649**	0.1261**	2.1892	0.7608
	(1.07)	(19.71)	(3.25)		
12: Paper and Packaging	0.0033	0.7416**	− 0.0091	1.9195	0.7349
	(1.59)	(22.16)	(− 0.25)		
13: Retail	0.0008	0.7939**	0.0504	1.8676	0.8360
	(0.34)	(20.81)	(1.33)		
14: Transport	0.0012	1.0496**	− 0.0015	1.7645	1.0480
	(0.43)	(23.11)	(− 0.04)		
15: Media	0.0052	0.8796**	0.1311**	1.8886	1.0123
	(1.34)	(14.15)	(2.86)		
16: Banks	− 0.0027	0.8310**	0.0441	1.8347	0.8693
	(1.01)	(19.42)	(1.12)		
17: Insurance	0.0007	0.6988**	0.2004**	2.1599	0.8739
	(0.20)	(12.81)	(4.21)		

Table 2. Continued.

ASX Industry	a_i	b_i	θ_i	DW Test	β_{LTi}
18: Entrepreneurial Investors	− 0.0032	1.1930**	0.1324**	2.1543	1.3751
	(− 0.73)	(16.49)	(3.10)		
19: Investment and Financial Services	− 0.0002	0.7054**	0.2224**	2.2124	0.9072
	(− 0.14)	(25.05)	(6.84)		
20: Property Trusts	0.0062**	0.4302**	0.0482	1.9507	0.4520
	(3.48)	(15.26)	(1.07)		
21: Miscellaneous Services	0.0044*	0.6875**	0.0920**	1.9562	0.7572
	(1.87)	(17.95)	(2.24)		
22: Miscellaneous Industrials	− 0.0010	0.6550**	0.2093**	1.9931	0.8284
	(− 0.53)	(21.86)	(5.87)		
23: Diversified Industrials	0.0024	0.9095**	0.0002	1.9106	0.9097
	(1.29)	(30.59)	(0.01)		
24: Tourism and Leisure	0.0060**	0.5043**	0.0448	2.0284	0.5280
	(2.12)	(11.09)	(0.88)		
Number of significant test statistics at 5% (10%) level	4 (5)	24 (24)	11 (12)	–	–

subperiods for four industries, specifically, Other Metals; Insurance; Investment and Financial Services; and Miscellaneous Industrials. In contrast, there are seven industries which never have a significant lagged dependent variable, specifically, Alcohol and Tobacco; Chemicals; Paper and Packaging; Retail; Transport; Banks; and Property Trusts. There are nine industries which only have a significant lagged dependent variable in the pre-float period; two industries with significant lagged dependent variables in the pre-float and post-crash periods; and two industries with a significant lagged dependent variable for only the post-crash period.

The results in Table 3 lead to the obvious question of how stable the coefficients are across subperiods. Accordingly, Table 4 reports tests of the stability of short-term beta across sub-periods. The following thirteen industries exhibit considerable evidence of instability of short-term beta: Solid Fuels; Oil and Gas; Diversified Resources; Developers and Contractors; Food and Household Goods; Engineering; Media; Insurance; Entrepreneurial Investors; Investment and Financial Services; Property Trusts; Diversified Industrials; and Tourism and Leisure. In contrast, the following seven industries exhibit little if any evidence of instability over time: Gold; Other Metals; Alcohol and Tobacco; Chemicals; Retail; Transport; and Miscellaneous

Table 3. Augmented Market Model Regression With Subperiod Dummy Variables.

This table reports the results of estimating the augmented market model with subperiod dummy variables for 24 industry portfolios over the period 1974 to 1995. We report OLS parameter estimates, plus t-statistics for the hypothesis that the coefficient equals zero in parentheses. Coefficients significant at the 5% level are indicated with a **, while coefficients significant at the 10% level are indicated by *.

	b_i				θ_i		
ASX Industry	1974/83	1983/87	1988/95	Oct. 1987	1974/83	1983/87	1988/95
1: Gold	1.3213**	1.3046**	1.2683**	1.4008**	0.1476**	0.1300	0.0798
	(11.38)	(5.89)	(6.28)	(9.21)	(2.38)	(1.24)	(1.13)
2: Other Metals	1.3465**	1.4416**	1.3119**	1.2815**	− 0.0331	− 0.0911	− 0.0843
	(18.29)	(10.28)	(10.27)	(13.28)	(0.68)	(1.24)	(1.59)
3: Solid Fuels	0.9666**	0.5498**	0.6525**	0.7186**	0.1663**	0.0499	0.0732
	(11.92)	(3.65)	(4.71)	(6.87)	(2.92)	(0.36)	(0.88)
4: Oil and Gas	1.5904**	0.9850**	0.8950**	0.9742**	0.1105**	0.2060*	0.1950**
	(17.69)	(5.85)	(5.78)	(8.33)	(2.48)	(1.82)	(2.31)
5: Diversified Resources	1.3867**	1.0240**	1.2473**	0.8282**	− 0.1008**	0.1762**	0.0760
	(25.25)	(9.82)	(13.09)	(11.51)	(2.72)	(2.37)	(1.47)
6: Developers and Contractors	0.9604**	0.9284**	0.9195**	1.2213**	0.1762**	0.0685	0.0103
	(17.97)	(8.86)	(10.04)	(17.63)	(3.94)	(0.73)	(0.23)
7: Building Materials	0.8636**	0.9249**	0.8882**	0.7356**	0.1022**	− 0.0708	0.0002
	(22.30)	(12.43)	(13.42)	(14.72)	(2.53)	(− 1.02)	(0.00)
8: Alcohol and Tobacco	0.7976**	0.8285**	0.7362**	0.9217**	0.0176	0.1781	0.0148
	(10.86)	(5.83)	(5.72)	(9.51)	(0.25)	(2.20)	(0.21)
9: Food and Household Goods	0.5329**	0.7577**	0.8314**	0.8355**	0.1851**	0.0865	− 0.0821
	(11.19)	(7.77)	(9.93)	(13.38)	(2.83)	(0.97)	(− 1.59)
10: Chemicals	0.7227**	0.8434**	0.8733**	0.7384**	− 0.0848	− 0.0416	− 0.0105
	(11.79)	(7.10)	(8.06)	(9.16)	(− 1.38)	(− 0.40)	(− 0.16)
11: Engineering	0.5402**	0.8792**	0.8529**	0.6367**	0.2326**	0.0540	0.0865
	(11.00)	(9.28)	(10.13)	(9.99)	(3.59)	(0.59)	(1.57)

Table 3. Continued.

12: Paper and Packaging	0.6861**	0.7208**	0.9342**	0.7405**	0.0357	− 0.0754	− 0.0117
	(13.99)	(7.37)	(11.01)	(11.54)	(0.61)	(− 0.79)	(− 0.22)
13: Retail	0.7209**	0.8585**	0.7518**	0.9008**	0.0098	0.1030	0.0902
	(12.85)	(7.69)	(7.70)	(12.27)	(0.17)	(1.03)	(1.55)
14: Transport	0.9400**	1.1640**	1.0334**	1.1710**	0.0282	0.0751	− 0.0608
	(14.19)	(8.85)	(8.95)	(13.41)	(0.51)	(0.88)	(− 1.10)
15: Media	0.5297**	1.2342**	1.5206**	0.9496**	0.2434**	0.0738	0.1145**
	(6.11)	(7.32)	(10.20)	(8.44)	(2.37)	(0.77)	(2.11)
16: Banks	0.8891**	0.7433**	0.9967**	0.6807**	0.0691	− 0.0182	0.0260
	(14.19)	(6.18)	(9.17)	(8.29)	(1.24)	(− 0.19)	(0.39)
17: Insurance	0.4592**	1.0419**	0.9977**	0.7309**	0.2843**	0.2233**	0.1410*
	(5.90)	(7.01)	(7.30)	(7.13)	(3.72)	(2.32)	(1.90)
18: Entrepreneurial Investors	0.7936**	1.1168**	1.0789**	1.9308**	0.2536**	0.1091	0.1048**
	(8.12)	(5.85)	(6.33)	(15.00)	(3.44)	(0.91)	(2.05)
19: Investment and Financial Services	0.6028**	0.6908**	0.6665**	0.8710**	0.1953**	0.2327**	0.2779**
	(14.93)	(8.97)	(9.56)	(16.54)	(3.70)	(2.79)	(6.22)
20: Property Trusts	0.4065**	0.2766**	0.3595**	0.5569**	0.0725	0.1842	0.0209
	(9.71)	(3.45)	(5.05)	(10.36)	(1.15)	(1.36)	(0.30)
21: Miscellaneous Services	0.7407**	0.6954**	0.6795**	0.6095**	0.0508	0.0277	0.1275**
	(13.00)	(6.31)	(6.90)	(8.23)	(0.76)	(0.23)	(2.24)
22: Miscellaneous Industrials	0.5779**	0.5483**	0.7127**	0.7892**	0.2343**	0.2385**	0.1984**
	(13.28)	(6.61)	(9.40)	(13.78)	(4.00)	(2.30)	(4.07)
23: Diversified Industrials	0.7802**	0.9598**	0.9763**	1.0528**	0.0873*	0.0174	− 0.0514
	(17.99)	(11.73)	(13.22)	(18.89)	(1.86)	(0.23)	(− 1.29)
24: Tourism and Leisure	0.3868**	0.2498**	0.6374**	0.7043**	0.1722**	0.1317	− 0.0675
	(5.90)	(1.99)	(5.58)	(8.20)	(2.31)	(0.96)	(− 1.02)
Number of significant test statistics at 5% (10%) level	24 (24)	24 (24)	24 (24)	24 (24)	13 (15)	4 (4)	6 (8)

Table 4. Testing the Equality of the Short-term Betas.

This table reports the results of testing the equality of the short-term betas across the different subperiods in the augmented market model. We report the calculated values of the test statistic plus p-values in parentheses. If the test statistic is significant at the 5% level it is indicated by **, while if the test statistic is significant at the 10% level it is indicated by *.

ASX Industry	Ho_1: $b_1 = b_2$	Ho_2: $b_1 = b_3$	Ho_3: $b_1 = b_4$	Ho_4: $b_2 = b_3$	Ho_5: $b_2 = b_4$	Ho_6: $b_3 = b_4$	Ho_7: $b_1 = b_2 = b_3$	Ho_8: $b_1 = b_2 = b_4$	Ho_9: $b_1 = b_3 = b_4$	Ho_{10}: $b_2 = b_3 = b_4$	Ho_{11}: $b_1 = b_2 = b_3 = b_4$
1: Gold	0.0046	0.0526	0.1709	0.0150	0.1267	0.2723	0.0528	0.2045	0.3047	0.3039	0.3194
	(0.946)	(0.819)	(0.679)	(0.902)	(0.722)	(0.602)	(0.974)	(0.903)	(0.859)	(0.859)	(0.956)
2: Other Metals	0.3686	0.0559	0.2845	0.4806	0.8734	0.0358	0.2614	0.4423	0.1458	0.4503	0.3116
	(0.544)	(0.813)	(0.594)	(0.489)	(0.351)	(0.850)	(0.770)	(0.643)	(0.864)	(0.638)	(0.817)
3: Solid Fuels	6.0621**	3.8818**	3.4866*	0.2599	0.8377	0.1439	8.0252**	7.6076**	5.6734*	0.8396	8.8787**
	(0.014)	(0.049)	(0.062)	(0.610)	(0.360)	(0.704)	(0.018)	(0.022)	(0.059)	(0.657)	(0.031)
4: Oil and Gas	10.2971**	15.3233**	17.3188*	0.1593	0.0027	0.1655	10.3660**	11.0240**	12.6140**	0.1064	9.3718**
	(0.002)	(0.000)	(0.000)	(0.690)	(0.959)	(0.685)	(0.000)	(0.000)	(0.000)	(0.899)	(0.000)
5: Diversified Resources	9.6866**	1.6340	37.7273**	2.5706	2.3610	12.2115**	4.9945**	20.053**	19.0064**	6.1427**	13.5915**
	(0.002)	(0.202)	(0.000)	(0.110)	(0.126)	(0.001)	(0.007)	(0.000)	(0.000)	(0.002)	(0.000)
6: Developers and Contractors	0.0753	0.1512	8.8228**	0.0042	5.3846**	6.8489**	0.1868	10.0239**	10.5943*	9.0774**	11.3758**
	(0.784)	(0.697)	(0.003)	(0.948)	(0.020)	(0.009)	(0.911)	(0.007)	(0.005)	(0.011)	(0.010)
7: Building Materials	0.5420	0.1042	4.0692**	0.1388	4.4105**	3.3584*	0.5648	5.8230*	5.0106*	5.7510*	6.3891*
	(0.462)	(0.747)	(0.044)	(0.709)	(0.036)	(0.067)	(0.754)	(0.054)	(0.082)	(0.056)	(0.094)
8: Alcohol and Tobacco	0.0380	0.1748	1.0336	0.2378	0.2904	1.3164	0.2651	1.0388	1.5853	1.3341	1.5884
	(0.845)	(0.676)	(0.309)	(0.626)	(0.590)	(0.251)	(0.876)	(0.595)	(0.453)	(0.513)	(0.662)
9: Food and Household Goods	4.3430**	9.7379**	14.7416**	0.3342	0.4482	0.0016	11.7845**	16.0471**	19.1494**	0.4880	19.8214**
	(0.037)	(0.002)	(0.000)	(0.563)	(0.503)	(0.968)	(0.003)	(0.000)	(0.000)	(0.784)	(0.000)
10: Chemicals	0.8299	1.4849	0.0236	0.0355	0.5290	0.9919	1.9052	0.8375	1.5276	1.1512	2.0277
	(0.362)	(0.223)	(0.878)	(0.850)	(0.467)	(0.319)	(0.386)	(0.658)	(0.466)	(0.562)	(0.567)
11: Engineering	10.2401**	10.4173**	1.4301	0.0441	4.4607**	4.1556**	16.6566**	10.3672**	10.4782**	6.3408**	16.7376**
	(0.001)	(0.001)	(0.232)	(0.834)	(0.035)	(0.041)	(0.000)	(0.006)	(0.005)	(0.042)	(0.001)

Table 4. Continued.

12: Paper and Packaging	0.1022	6.5054**	0.4511	2.7760*	0.0282	3.2856*	6.5796**	0.4740	6.5117**	4.0244	6.5802*
	(0.749)	(0.011)	(0.502)	(0.096)	(0.867)	(0.070)	(0.037)	(0.789)	(0.039)	(0.134)	(0.087)
13: Retail	1.2321	0.0763	3.7580*	0.5287	0.0994	1.4738	1.2327	4.1558	3.8491	1.4878	4.3258
	(0.267)	(0.782)	(0.053)	(0.467)	(0.753)	(0.225)	(0.540)	(0.125)	(0.146)	(0.475)	(0.228)
14: Transport	2.3559	0.4997	4.4071**	0.5711	0.0019	0.8972	2.4890	5.4703*	4.4177	0.9953	5.4753
	(0.125)	(0.480)	(0.036)	(0.450)	(0.965)	(0.344)	(0.288)	(0.065)	(0.110)	(0.608)	(0.140)
15: Media	13.9810**	33.4079**	8.6826**	1.6535	1.9497	9.2571**	39.3667**	18.0829**	35.0750**	9.3719**	39.997**
	(0.000)	(0.000)	(0.003)	(0.198)	(0.163)	(0.002)	(0.000)	(0.000)	(0.000)	(0.009)	(0.000)
16: Banks	1.1788	0.7470	4.0388**	2.5043	0.1826	5.3387**	2.5086	4.3735	6.3452**	5.5327*	7.0736*
	(0.278)	(0.387)	(0.044)	(0.114)	(0.669)	(0.021)	(0.285)	(0.112)	(0.042)	(0.063)	(0.070)
17: Insurance	12.3042**	11.6153**	4.4188**	0.075	2.9346*	2.2654	19.3407**	13.7919**	12.9673**	3.8266	19.6919*
	(0.000)	(0.001)	(0.036)	(0.785)	(0.087)	(0.132)	(0.000)	(0.001)	(0.002)	(0.148)*	(0.000)
18: Entrepreneurial Investors	2.3129	2.1425	49.1679**	0.0225	12.3870**	15.7708**	3.6138	49.2688**	49.5144**	20.7955**	49.5630**
	(0.128)	(0.143)	(0.000)	(0.881)	(0.000)	(0.000)	(0.164)	(0.000)	(0.000)	(0.000)	(0.000)
19: Investment and Financial Services	1.0420	0.6342	16.2218**	0.0562	3.6953*	5.4348**	1.3647	16.2255**	16.3867**	6.7696**	16.3868**
	(0.307)	(0.426)	(0.000)	(0.813)	(0.055)	(0.020)	(0.505)	(0.000)	(0.000)	(0.034)	(0.001)
20: Property Trusts	2.0846	0.3280	4.8433**	0.6102	8.3498**	4.8598**	2.1379	9.3580**	6.5269**	9.9394**	10.2475**
	(0.149)	(0.567)	(0.028)	(0.435)	(0.004)	(0.027)	(0.343)	(0.009)	(0.038)	(0.007)	(0.017)
21: Miscellaneous Services	0.1359	0.2934	1.9583	0.0118	0.4145	0.3205	0.3549	1.9583	1.9762	0.5445	1.9769
	(0.712)	(0.588)	(0.162)	(0.914)	(0.520)	(0.571)	(0.837)	(0.376)	(0.372)	(0.762)	(0.577)
22: Miscellaneous Industrials	0.1019	2.4126	8.5608**	2.1954	5.6397**	0.6427	2.9155	9.8410**	9.1044**	5.6451*	10.7433**
	(0.750)	(0.120)	(0.003)	(0.138)	(0.018)	(0.423)	(0.233)	(0.007)	(0.011)	(0.059)	(0.013)
23: Diversified Industrials	3.8158*	5.3106**	14.7986**	0.0228	0.8726	0.6784	7.3678**	15.6750**	16.2452**	1.1451	16.7864**
	(0.051)	(0.021)	(0.000)	(0.880)	(0.350)	(0.410)	(0.025)	(0.000)	(0.000)	(0.564)	(0.001)
24: Tourism and Leisure	0.9518	3.6711*	8.5671**	5.3416**	8.7997**	0.2178	5.8361*	11.8396**	9.7935**	9.2247**	13.9213**
	(0.329)	(0.055)	(0.003)	(0.021)	(0.003)	(0.641)	(0.054)	(0.003)	(0.007)	(0.010)	(0.003)
Number of significant test statistic at 5% (10%) level	7 (8)	8 (9)	15 (17)	1 (3)	7 (10)	8 (10)	9 (10)	14(16)	14(16)	8(11)	14(17)

Table 5. Testing the Equality of the Coefficient on the Lagged Dependent Variable.

This table reports the results of testing the equality and significance of the lagged dependent variables in the augmented market model. We report the calculated values of the Wald test statistic plus p-values in parentheses. If the test statistic is significant at the 5% level it is indicated by **, while if the test statistic is significant at the 10% level it is indicated by *.

ASX Industry	Ho_{12}:$\theta_1 = \theta_2$	Ho_{13}:$\theta_1 = \theta_3$	Ho_{14}:$\theta_2 = \theta_3$	Ho_{15}:$\theta_1 = \theta_2 = \theta_3$	Ho_{16}:$\theta_1 = \theta_2 = \theta_3 = 0$
1: Gold	0.0212	0.5201	0.1576	0.5299	8.4695**
	(0.884)	(0.471)	(0.691)	(0.767)	(0.037)
2: Other Metals	0.4350	0.5072	0.0057	0.3440	1.5049
	(0.510)	(0.477)	(0.940)	(0.709)	(0.214)
3: Solid Fuels	0.5977	0.8644	0.0207	1.2150	9.3637**
	(0.439)	(0.353)	(0.886)	(0.545)	(0.025)
4: Oil and Gas	0.6162	0.7849	0.0061	0.5996	4.9199**
	(0.433)	(0.376)	(0.938)	(0.550)	(0.002)
5: Diversified Resources	11.1860**	7.7608**	1.2351	7.5440**	5.083**
	(0.000)	(0.005)	(0.267)	(0.001)	(0.002)
6: Developers and Contractors	1.0875	6.8123**	0.3126	6.8933**	16.0304**
	(0.297)	(0.009)	(0.576)	(0.032)	(0.001)
7: Building Materials	4.6626**	2.8831*	0.7389	5.7421*	7.5059*
	(0.031)	(0.090)	(0.390)	(0.057)	(0.057)
8: Alcohol and Tobacco	2.2368	0.0008	2.2810	2.8985	4.9016
	(0.135)	(0.978)	(0.131)	(0.235)	(0.179)
9: Food and Household Goods	0.8182	10.3523**	2.6732	10.7239**	11.3638**
	(0.366)	(0.001)	(0.102)	(0.005)	(0.010)
10: Chemicals	0.1294	0.6911	0.0644	0.6971	2.0703
	(0.719)	(0.406)	(0.800)	(0.706)	(0.558)
11. Engineering	2.5951	2.9607*	0.0925	3.8844	15.5559**
	(0.107)	(0.085)	(0.761)	(0.143)	(0.001)

Table 5. Continued.

12: Paper and Packaging	0.9941	0.3667	0.3403	1.0535	1.0577
	(0.319)	(0.545)	(0.560)	(0.591)	(0.787)
13: Retail	0.6596	0.9503	0.0124	1.2009	3.5131
	(0.417)	(0.330)	(0.911)	(0.549)	(0.319)
14: Transport	0.2181	1.3115	1.8016	2.2508	2.2512
	(0.640)	(0.252)	(0.180)	(0.325)	(0.522)
15: Media	1.5016	1.2364	0.1364	1.6867	10.5160**
	(0.220)	(0.266)	(0.712)	(0.430)	(0.015)
16: Banks	0.6158	0.2482	0.1424	0.6844	1.7310
	(0.433)	(0.618)	(0.706)	(0.710)	(0.630)
17: Insurance	0.2517	1.7962	0.4528	1.8038	22.6517**
	(0.616)	(0.180)	(0.501)	(0.406)	(0.000)
18: Entrepreneurial Investors	1.0893	2.7086	0.0011	2.8743	16.9839**
	(0.297)	(0.100)	(0.974)	(0.238)	(0.001)
19: Investment and Financial	0.1481	1.4151	0.2259	1.4221	59.9178**
Services	(0.700)	(0.234)	(0.635)	(0.491)	(0.000)
20: Property Trusts	0.5754	0.3088	1.1637	1.2000	3.1058
	(0.448)	(0.578)	(0.281)	(0.549)	(0.376)
21: Miscellaneous Services	0.0284	0.7689	0.5665	1.0475	5.5905
	(0.866)	(0.381)	(0.452)	(0.592)	(0.133)
22: Miscellaneous Industrials	0.0013	0.2214	0.1218	0.2697	37.6128**
	(0.972)	(0.638)	(0.727)	(0.874)	(0.000)
23: Diversified Industrials	0.6504	5.1007**	0.6691	5.1097*	5.1873
	(0.420)	(0.024)	(0.413)	(0.078)	(0.159)
24: Tourism and Leisure	0.0678	5.8745**	1.7167	6.2834**	7.3425*
	(0.794)	(0.015)	(0.190)	(0.043)	(0.062)
Number of significant test statistic at 5% (10%) level	2 (2)	5 (7)	0 (0)	4 (6)	12(14)

Table 6. Testing the Equality of the Long-term Betas.

This table reports the results of testing the equality of the long-term betas across the different subperiods in the augmented market model. We report the calculated values of the test statistic plus p-values in parentheses. If the test statistic is significant at the 5% level it is indicated by **, while if the test statistic is significant at the 10% level it is indicated by *.

ASX Industry	$Ho_{17}: b_1/(1-\theta_1) = b_2/(1-\theta_2)$	$Ho_{18}: b_1/(1-\theta_1) = b_3/(1-\theta_3)$	$Ho_{19}: b_2/(1-\theta_2) = b_3/(1-\theta_3)$	$Ho_{20}: b_1/(1-\theta_1) = b_2/(1-\theta_2) = b_3/(1-\theta_3)$
1: Gold	0.0242	0.3348	0.1026	0.3349
	(0.876)	(0.563)	(0.749)	(0.846)
2: Other Metals	0.0115	0.3477	0.3246	0.2136
	(0.914)	(0.556)	(0.5693)	(0.808)
3: Solid Fuels	8.0233**	5.6983**	0.2856	10.6368**
	(0.005)	(0.017)	(0.593)	(0.005)
4: Oil and Gas	3.559*	6.797**	0.1356	4.3896**
	(0.060)	(0.010)	(0.713)	(0.013)
5: Diversified Resources	0.0010	0.3994	0.2808	0.2207
	(0.921)	(0.528)	(0.597)	(0.802)
6: Developers and Contractors	1.3476	3.4102*	0.1809	3.7464
	(0.246)	(0.065)	(0.671)	(0.154)
7: Building Materials	1.1411	0.6498	0.0542	1.3719
	(0.285)	(0.420)	(0.816)	(0.504)
8: Alcohol and Tobacco	0.9710	0.1598	1.3608	1.4089
	(0.324)	(0.689)	(0.243)	(0.494)
9: Food and Household Goods	1.9826	1.1639	0.2195	2.3450
	(0.159)	(0.281)	(0.639)	(0.310)
10: Chemicals	1.0912	2.4317	0.1105	2.9283
	(0.296)	(0.119)	(0.740)	(0.231)
11: Engineering	2.7469*	3.0895*	0.0007	4.4321
	(0.097)	(0.079)	(0.979)	(0.109)

Table 6. Continued.

12: Paper and Packaging	0.1487	3.5512*	3.7916*	4.5371
	(0.700)	(0.060)	(0.052)	(0.103)
13: Retail	2.4214	0.5700	0.5637	2.6051
	(0.1197)	(0.450)	(0.453)	(0.272)
14: Transport	2.8952*	0.0022	2.1725	3.0639
	(0.089)	(0.962)	(0.140)	(0.216)
15: Media	7.6018**	19.5222**	2.0431	21.6035**
	(0.006)	(0.000)	(0.153)	(0.000)
16: Banks	2.3746	0.2085	2.7955*	3.2852
	(0.123)	(0.648)	(0.095)	(0.193)
17: Insurance	7.2591**	5.2288**	0.4106	10.1349**
	(0.007)	(0.022)	(0.522)	(0.006)
18: Entrepreneurial Investors	0.4699	0.3212	0.0249	0.5996
	(0.493)	(0.571)	(0.875)	(0.741)
19: Investment and Financial Services	1.1957	1.9165	0.0187	2.5773
	(0.274)	(0.166)	(0.891)	(0.276)
20: Property Trusts	0.8466	0.6458	0.0524	1.2179
	(0.358)	(0.422)	(0.819)	(0.544)
21: Miscellaneous Services	0.2126	0.0001	0.1421	0.2260
	(0.645)	(0.991)	(0.706)	(0.893)
22: Miscellaneous Industrials	0.0518	1.0861	0.9769	1.3669
	(0.820)	(0.297)	(0.323)	(0.505)
23: Diversified Industrials	1.1864	0.6110	0.1540	1.4144
	(0.276)	(0.434)	(0.695)	(0.493)
24: Tourism and Leisure	1.2165	0.9187	3.0452*	3.0618
	(0.270)	(0.338)	(0.081)	(0.216)
Number of significant test statistic at 5% (10%) level	3 (6)	4 (7)	0 (3)	4 (4)

Services. Generally, the short-term betas are found not to be stable across subperiods.

Table 5 reports the results for testing the stability of the lagged dependent variable across subperiods. The following six industries exhibit considerable evidence of instability: Diversified Resources; Developers and Contractors; Building Materials; Food and Household Goods; Diversified Industrials; and Tourism and Leisure. The results for the remaining eighteen industries supports the stability of the θ coefficients over time. In terms of the joint significance of the θ coefficients across the three sub-periods, there are fourteen industries where it can be concluded that the θ coefficients are jointly significant.

The results reported in Table 6 combine the analysis of Tables 4 and 5 by examining the stability of long-term beta. There are only four industries which exhibit strong evidence of instability in the long-term beta, namely, Solid Fuels; Oil and Gas; Media; and Insurance. It is interesting to note that the instability for these four industries appears to be caused by differences in the pre-deregulation period. That is, in all four cases we find that the long run pre-deregulation beta is significantly different from the long run pre-crash and post-crash betas respectively. In contrast, none of these four industries exhibits a significant difference in the long run beta when comparing the pre-crash and post-crash periods.

Table 7 reports the likelihood ratio test [Gibbons (1982)] results for the three subperiods and the full sample period (January 1974 to December 1995), as outlined earlier. The table reveals that most of the estimated gamma coefficients (a potential proxy for risk-free rate) for both the static and dynamic CAPMs are significant at 5% level of significance. The sole exception is the subperiod 1988 to 1995 for the resources industries group. It is interesting to note that the estimates obtained are not very sensitive to which model is estimated. There is however a question of whether the estimated gamma is a credible estimate of risk-free rate. With the exception of the period 1988 to 1995 for resources portfolios, the estimated gamma appears to be too high. For example, an estimate of 1.1% per month (approximately 14% per annum) is obtained for the full sample period, when using all 24 industries.

Now consider the results for testing the restrictions imposed by the static and dynamic versions of the CAPM, contained in Table 7. Apart from one case of weak rejection (at the 10% level) for Resources during the period 1983:12 to 1987:09, the likelihood ratio test indicates strong support for the static CAPM. Likewise, the results regarding the Dynamic CAPM show that it cannot be rejected across any subperiod or industry subset.[11]

This leads to the important question of which of the two models is to be preferred. While the models cannot be compared directly in a nested

Table 7. Likelihood Ratio Tests of the Static and Dynamic Versions of the CAPM.

This table reports the results of multivariate likelihood ratio testing (LRT) of the static and dynamic CAPM. The LRT statistics are adjusted for small sample bias following Gibbons, Ross and Shanken (1989) and MacKinlay and Richardson (1991). We report the calculated values of the test statistic plus p-values in parentheses. If the test statistic is significant at the 5% level it is indicated by **, while if the test statistic is significant at the 10% level it is indicated by *. The estimated coefficient for the expected return on the zero-beta portfolio (gamma) and its associated t-statistic (in parentheses) are also reported.

		STATIC CAPM		DYNAMIC CAPM	
	Degrees of Freedom	Estimated Gamma	Adjusted LRT Statistic	Estimated Gamma	Adjusted LRT Statistic
Panel A: Monthly Returns 1974: 01 to 1983: 11					
Resources	4	0.0125**	2.5879	0.0112*	2.1024
		(2.081)	(0.629)	(1.783)	(0.717)
Industrials	18	0.0123**	8.5174	0.0126**	7.4285
		(4.364)	(0.970)	(4.165)	(0.986)
All Industries	23	0.0120**	13.2014	0.0122**	11.8799
		(4.617)	(0.947)	(4.441)	(0.972)
Panel B: Monthly Returns 1983: 12 to 1987: 09					
Resources	4	0.0312**	8.7089*	0.0510**	4.9753
		(3.110)	(0.069)	(3.419)	(0.290)
Industrials	18	0.0144**	14.418	0.0138**	13.3813
		(6.202)	(0.701)	(6.247)	(0.768)
All Industries	23	0.0145**	18.3518	0.0139**	15.6915
		(6.882)	(0.738)	(7.602)	(0.868)
Panel C: Monthly Returns 1988: 01 to 1995: 12					
Resources	4	0.0058	3.3031	0.0060	2.8544
		(0.919)	(0.508)	(0.991)	(0.582)
Industrials	18	0.0114**	21.9307	0.0124**	18.8831
		(4.732)	(0.235)	(4.896)	(0.399)
All Industries	23	0.0125**	28.4539	0.0125**	25.8125
		(5.681)	(0.199)	(5.782)	(0.310)
Panel D: Monthly Returns 1974: 01 to 1995: 12					
Resources	4	0.0173**	2.9827	0.0154**	3.5343
		(2.765)	(0.561)	(2.571)	(0.473)
Industrials	18	0.0106**	10.6878	0.0114**	9.3141
		(5.460)	(0.907)	(5.539)	(0.952)
All Industries	23	0.0113**	17.6301	0.0115**	16.7055
		(6.110)	(0.777)	(5.980)	(0.823)

framework, we are able to perform a multivariate test of the joint equality of the θ_i coefficients to zero. Rejection of this hypothesis would provide indirect evidence favouring the dynamic CAPM. In unreported analysis, this hypothesis is soundly rejected across the twenty-four portfolio system.

5. CONCLUSION

In the current paper we have developed a new version of a dynamic CAPM, primarily testing within the framework of a dynamic market model. The dynamic market model (see Cartwright and Lee (1987)) is obtained by simply augmenting the standard market model with a lagged dependent variable. The primary (but not sole) justification for this approach was a recognition of the thin trading phenomena that is characteristic of many modern day markets. Notably, the dynamic CAPM accommodates the concept of both a short-run and a long-run (full impact) beta risk measure. We subject the static and dynamic versions of the CAPM to empirical scrutiny. Specifically, we conduct these tests using Australian industry portfolios over the period January 1974 to December 1995. In addition to analysing the full sample period, we also considered a pre-deregulation subperiod, a pre-crash subperiod and a post-crash subperiod.

In summary, our key findings are as follows. First, a number industry portfolios reveal a problem with serial correlation in a standard market model setting, thereby providing justification to investigate models including a lagged dependent variable. Second, there is a range of industries with a significant coefficient on the lagged dependent variable, thus supporting exploration of the dynamic CAPM. Third, there is some evidence suggesting instability in the short-run and long-run betas across the subperiods. Fourth, we find strong evidence in favour of both versions of the model, leading to the question of which one is preferred. Finally on this question, we are able to perform a multivariate test of the joint equality of the θ_i coefficients to zero – which is rejected – thus providing indirect evidence favouring the dynamic CAPM.

NOTES

1. See Jagannathan, Kubota and Takehara (1998, p. 320).
2. Emerging Stock Markets Fact Book (1993).
3. Hathaway (1986) reports that from a sample of 385 Australian stocks, over a nine-year period, about 30% reveal little or no trading. Moreover, he found that even the top fifty stocks do not trade every day.
4. Emerging Stock Markets Fact Book (1997).
5. For another example of this Lo and MacKinlay argument, see Clare, Priestley and Thomas (1998, p. 1209).

6. Gibbons' (1982) multivariate testing methodology requires that returns are stationary and have a multivariate normal distribution. Consequently, the results reported in this paper are necessarily joint tests of mean-variance efficiency and of the multivariate distributional assumption about returns.

7. It could be argued that a very thinly traded stock will have returns which will be highly correlated with lagged market returns but uncorrelated with future market returns. In such a circumstance b_{ij} would be significant for negative j, but zero for positive j. In this case a simple geometric progression will not capture the features of the data, however, for reasons of simplicity we retain our assumption of a simple geometric progression. We wish to thank an anonymous referee for bringing this issue to our attention.

8. The standard multivariate-testing framework assumes that the disturbances are independent of the independent variables. In the present setting this will only be the case where the regression error terms are uncorrelated. This is clearly an empirical question. We wish to thank an anonymous referee for drawing our attention to this point.

9. Following on from the discussion of footnote 8, this implies that for twenty-two of our twenty-four industry portfolios, the standard multivariate-testing framework will be valid. This issue will be addressed further in the later analysis.

10. For a detailed discussion see Inder (1984, 1986) and King (1987).

11. Following the comments made in footnotes 8 and 9, we have re-run our analysis removing two industries due to potential concerns about endogeneity (ie. the two cases that revealed significant serial correlation in the residuals of the dynamic market model of [2]) – namely, Other Metals and Alcohol and Tobacco. Our results are not impacted by this robustness check and, hence, these additional results are not reported to conserve space.

ACKNOWLEDGMENT

We would like to thank an anonymous referee for making valuable comments and suggestions that greatly improved the quality of this paper.

REFERENCES

Ball, R., Brown, P., & Officer, R. (1976). Asset Pricing in the Australian Industrial Equity Market. *Australian Journal of Management*, *1*, 1–32.

Black, F. (1993). Beta and Return. *Journal of Portfolio Management*, Fall, 8–18.

Brailsford, T., & Faff, R. (1997). Testing the conditional CAPM and the effect of Intervaling: A Note. *Pacific-Basin Finance Journal*, *5*, 527–537.

Brooks, R. D., & Faff, R. W. (1995). Financial Market Deregulation and Bank Risk: Testing for Beta Instability. *Australian Economic Papers*, *34*, 180–199.

Brooks, R. D., & Faff, R. W. (1997). Financial Market Deregulation and Relative Risk of Australian Industry. *Australian Economic Papers*, *36*, 308–320.

Brooks, R. D., Faff, R. W., & McKenzie, M. D. (1997). Bivariate GARCH Estimation of Beta Risk in the Australian Banking Industry. *Accountability and Performance*, *3*, 81–101.

Brooks, R. D., Faff, R. W., & McKenzie, M. D. (1998). Time-Varying Beta Risk of Australian Industry Portfolios: A Comparison of Modelling Techniques. *Australian Journal of Management*, *23*, 1–22.

Cartwright, P., & Lee, C. (1987). Time Aggregation and the Estimation of the Market Model: Empirical Evidence. *Journal of Business and Economic Statistics*, *5*, 131–143.

Clare, A., Priestley, R., & Thomas, S. (1998). Reports of Beta's Death are Premature: Evidence from the U.K. *Journal of Banking and Finance*, *22*, 1207–1229.

Davis, J. (1994). The Cross-Section of Realised Stock Returns: The pre-COMPUSTAT Evidence. *Journal of Finance*, *49*, 1579–1593.

Dimson, E. (1979). Risk Measurement when Shares are Subject to Infrequent Trading. *Journal of Financial Economics*, *7*, 197–226.

Durbin, J., & Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression I. *Biometrika*, *37*, 409–428.

Durbin, J., & Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression II. *Biometrika*, *38*, 159–178.

Faff, R. W. (1991). A Likelihood Ratio Test of the Zero-Beta CAPM in Australian Equity Returns. *Accounting and Finance*, *31*, 88–95.

Faff, R. W. (1992). A Multivariate Test of an Equilibrium APT with Time-Varying Risk Premia in the Australian Equity Market. *Australian Journal of Management*, *17*, 233–258.

Faff, R. W., & Lau, S. (1997). A Generalised Method of Moments Test of Mean Variance Efficiency in the Australian Stock Market. *Pacific Accounting Review*, *9*, 2–16.

Fama, E., & French, K. (1992). The Cross-section of Expected Stock Returns. *Journal of Finance*, *47*, 427–465.

Fama, E., & French, K. (1993). Common Risk Factors in the returns on Stocks and Bonds. *Journal of Financial Economics*, *33*, 3–56.

Fama, E., & French, K. (1995). Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance*, *50*, 131–155.

Fama, E., & French, K. (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, *51*, 55–84.

Fama, E. & French, K. (1997). Industry Costs of Equity. *Journal of Financial Economics*, *43*, 153–193.

Flannery, M., Hameed, A., & Harjes, R. (1997). Asset Pricing, Time-Varying Risk Premia and Interest Rate Risk, *Journal of Banking and Finance*, *21*, 315–335.

Gibbons, M. R. (1982). Multivariate Tests of Financial Models : A New Approach. *Journal of Financial Economics*, *10*, 3–27.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*, 1121–1152.

Harper, I., & Scheit, T. (1992). The Effects of Financial Market Deregulation on Bank Risk and Profitability. *Australian Economic Papers*, *31*, 260–271.

Hathaway, N. (1986). The Non-Stationarity of Share Price Volatility. *Accounting and Finance*, *26*, 35–54.

He, J., & Ng, L. (1994). Economic Forces, Fundamental Variables and Equity Returns. *Journal of Business*, *67*, 599–639.

Inder, B. A. (1984). Finite Sample Power of Tests for Autocorrelation in Models Containing Lagged Dependent Variables. *Economics Letters*, *14*, 179–185.

Inder, B. A. (1986). An Approximation to the Null Distribution of the Durbin-Watson Statistic in Models Containing Lagged Dependent Variables. *Econometric Theory*, *2*, 413–428.

International Finance Corporation (1993). *Emerging Stock Markets Factbook*. IFC, Washington.

International Finance Corporation (1997). *Emerging Stock Markets Factbook*. IFC, Washington.

Jagannathan, R., Kubota, K., & Takehara, H. (1998). Relationship Between Labor-Income Risk and Average Return: Empirical Evidence from the Japanese Stock Market. *Journal of Business*, *71*, 319–347.

Kariya, T. (1988). The Class of Models for which the Durbin-Watson Test is Locally Optimal. *International Economic Review*, *29*, 167–176.

King, M. L. (1987). *Testing for Autocorrelation in Linear Regression Models: A Survey, Specification Analysis in the Linear Model*. M. L. King & D. E. A. Giles (Eds). Routledge and Kegan Paul, London.

King M. L., & Evans, M. A. (1988). Locally Optimal Properties of the Durbin-Watson Test. *Econometric Theory*, *4*, 509–516.

Kothari, S., & Shanken, J. (1997). Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis. *Journal of Financial Economics*, *44*, 169–203.

Kothari, S., Shanken, J., & Sloan, R. (1995). Another look at the Cross-Section of Expected Returns. *Journal of Finance*, *50*, 185–224.

Lo, A., & MacKinlay, C. (1990). Data Snooping Biases in Tests of Financial Asset Pricing Models. *Review of Financial Studies*, *3*, 431–467.

Loughran, T. (1997). Book-to-Market across Firm Size, Exchange, and Seasonality: Is there an Effect? *Journal of Financial and Quantitative Analysis*, *32*, 249–268.

MacKinlay, A. C., & Richardson, M. (1991). Using Generalised Methods of Moments to Test Mean-Variance Efficiency. *Journal of Finance*, *46*, 511–527.

Pettengill, G., Sundaram, S., & Mathur, I. (1995). The Conditional Relation between Beta and Returns. *Journal of Financial and Quantitative Analysis*, *30*, 101–116.

Shanken, J. (1985). Multivariate Tests of the Zero-Beta CAPM. *Journal of Financial Economics*, *14*, 327–348.

Wood, J. (1991). A Cross-Sectional Regression Test of the Mean-Variance Efficiency of an Australian Value Weighted Market Portfolio. *Accounting and Finance*, *18*, 96–109.

This Page Intentionally Left Blank

BIASES IN USING JENSEN'S ALPHA

Yexiao Xu

ABSTRACT

Jensen's alpha is widely used in the evaluation of mutual fund performance. This paper investigates the accuracy of this measure and reveals a potential return measurement bias both theoretically and empirically due to the nonlinear geometric compounding of the return data. It is also shown that this source of bias can cause problems in the beta estimation of mutual funds.

INTRODUCTION

Whether or not average mutual funds have outperformed a popular benchmark still remains an open question. The first generation of empirical finance researchers, including Treynor (1965), Sharpe (1966), and Jensen (1968), generally believed, from both a theoretical perspective and their empirical results, that after relevant expenses, mutual funds did not systematically outperform a benchmark portfolio (such as the S&P 500). This conclusion is consistent with an efficient market hypothesis. While some recent papers confirm the earlier findings (for example, Carhart, 1997; Connor & Korajczyk, 1991; Grinblatt & Titman, 1989; Malkiel, 1995), other researchers have provided evidence that appears to be inconsistent with these previous conclusions. One line of research (Ippolito, 1989) justifies superior performance of mutual funds from the perspective of costly information collection. He argues that some mutual fund managers possess the ability to find inefficiencies in market prices and, thus, justify the costs of collecting information. Another

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 161–181.

ISBN: 0-7623-0798-6

line of research, while looking at the market as whole, claims there is persistency (a "hot hand" phenomenon) in the performance of the mutual fund industry. This means that a fund which performed well in one period is likely to perform well in the subsequent period (see Hendricks, Patel & Zeckhauser, 1993). Using portfolio composition data, Wermers (2000) shows that abnormal returns are highly correlated with fund styles. Recently, Day, Wang, and Xu (2000) have demonstrated that underperformance of a mutual fund does not necessarily mean that lack of abilities of managers to identify winning stocks. An inefficient allocation of assets in a fund could result in a low expected return as opposed to what modern portfolio theory suggests.

Setting aside the issue of portfolio inefficiency, measuring mutual fund performance relies on both the data sample and the evaluation method. This paper examines the relevance of evaluation measures that have been applied in practice. Generally speaking, there are three approaches used in empirical studies of mutual fund performance. First, Jensen's alpha is the most widely used measure and is related to the CAPM theory. Basically, it is the intercept from regressing the excess returns of a fund on the excess returns of a market portfolio. The style based approach, proposed by Daniel, Grinblatt, Titman and Wermers (1997), constructs benchmark according to the characteristics of individual stocks held in each mutual fund portfolio. This methodology relies on the knowledge of individual mutual funds' portfolio compositions, which is not widely available in practice. The third approach is to use either numerical simulation or strategy simulation (see Malkiel, 1995; Brown, Goetzmann, Ibboston & Ross, 1992). In this paper, we focus on Jensen's alpha since it is easy to apply and can be considered as the risk-adjusted return method. Although the CAPM model has been rejected in many empirical studies especially from the cross-sectional evidence, the market model is still useful to describe the risk and return relationship in the case of mutual funds because of the high R^2 for each fund. This is largely attributed to the fact that the beta of any portfolio is much more stable than that of the individual stocks. Presumably, one could use Carhart's (1997) multi-factor regression approach to compute alphas. However, as Merville, and Xu (2000) have demonstrated, a single factor model is sufficient for a well-diversified portfolio due to automatic "factor hedging." Moreover, the market model approach makes it possible to compare the performance of mutual funds to a common standard – the market portfolio, and it is commonly reported in financial publications. We take as given the validity of the CAPM model, and argue that the Jensen's alpha measure could still be seriously biased due to nonlinear compounding.

There are several studies have investigated the issue of "intervalling" effect of asset returns in asset pricing. Levhari and Levy (1977) are the first to show mathematically that the beta of defensive stocks decreases with the assumed investment horizon and the opposite holds for aggressive stocks. In an empirical study by Handa, Kothari, and Wasley (1989), they confirm that beta changes with the return interval. Furthermore, they provide evidence from cross-sectional regression that the size effect becomes statistically insignificant when risk is measured by betas estimated using annual returns. A follow-up study by Kothari, Shanken and Sloan (1995) suggests that the relation between book-to-market equity and returns is weaker and less consistent than that in Fama and French (1992) once annual betas are used. Longstaff (1989) investigated the same issue by comparing the CAPM with a discretized continuous time CAPM – a three-factor model. He finds that the former is easily rejected while the latter cannot be rejected.

Because of the popular use of Jensen's alpha, we investigate a similar bias in this measure due to the temporal aggregation effect of returns. This is an equally important issue relative to beta estimates especially in mutual fund performance evaluation. Based on a similar idea as in Levhari and Levy (1977), we derive the bias measure in Jensen's alpha without assuming independence in asset returns. Using actual stock returns and simulation, we find that the alpha estimates increase with the return horizon with no persistent predictive power. At the same time since the volatility measure increases much faster than the mean, the percentage of significant alphas may actually decrease with the return horizon. We also demonstrate how a risk measure such as beta could become less useful in cross-sectional regression using mutual fund returns where an error-in-variable problem is not an issue.

If Jensen's alpha measures the (superior) performance of mutual funds, it should be useful information for investors. In particular, we examine the predictability of Jensen's alpha in the next section. We then present a theoretical analysis of the potential bias in both the Jensen's alpha and the beta estimator as a result of nonlinear compounding in return calculations in Section 2. In order to study and assess the magnitude of such bias, we investigate the issue using all CRSP stocks in Section 3. Finally, we examine actual mutual fund returns for bias in Section 4. Section 5 concludes the paper. The main issue is not whether there exists a uniquely correct way to compute performance but to point out that there is always a bias whenever we use a nonlinear approximation in applying the CAPM model.

1. THE PREDICTABILITY OF JENSEN'S ALPHA

Despite the fact that the CAPM has been rejected in many recent empirical studies, especially in the cross sectional tests of risk and return relationship (see

Fama & French, 1992), the market model (without imposing zero alpha in the CAPM) is still widely applied in practice because of its simplicity. Furthermore, the model does capture the high correlations between individual funds' returns and market index returns. We have estimated the market model for each of the 206 equity mutual funds that survived over the period from 1971 to 1991.[1] The results show that the average estimated R^2 for individual mutual funds[2] is as high as 0.807 with a standard deviation of 0.109. This means that the linear risk and return relationship prescribed by the market model fits our mutual fund sample well. However, at the same time, empirical research has typically found a non-zero intercept term when estimating the CAPM for individual funds, which gives rise to the issue of a correct interpretation of this term.

Jensen (1968) first interpreted this non-zero intercept term as a measure of risk-adjusted performance in the case of mutual fund evaluation. Usually positive alphas are associated with superior performance, negative alphas with inferior performance. Because there were 14 significant negative alphas and 3 significant positive alphas among 115 mutual funds studied, he concluded that the mutual fund industry did not consistently outperform the market. In the subsequent literature, this intercept term is therefore labeled as "Jensen's alpha." Ippolito (1989) reached a different conclusion[3] after studying 143 mutual funds over a different period. He found that there were 12 funds with positive alphas and only 4 with negative alphas.

If there is mispricing in individual stocks relative to the CAPM, a consistent alpha over different time period for a fund may signal the fund managers' abilities to identify winning stocks. It is important to note that such a significant alpha should persist *ex post*. Otherwise, observing a positive alpha is pure luck. Christopherson and Turner (1991) find that alpha does not persist over different periods for pension funds, while Lakonishok, Shleifer and Vishny (1992) find some persistence of the relative returns of pension funds over two to three year period. Using conditional approach, Christopherson, Ferson and Glassman (1998) find that persistence in alpha is concentrated in underperforming pension funds. In this study, we investigate whether alpha captures important information by exploring an investment strategy based on picking funds with large alphas and controlling for risk. Presumably such an investment strategy would systematically beat a market index fund if Jensen's alpha were informative.

In Panel A of Table 1, we have implemented such a strategy. At the beginning of each year, we select an equally weighted portfolio of n funds from our sample with the largest alphas estimated from their previous five-year returns. Therefore, our performance sample runs from 1976 to 1991. We divided this sample period into three equal sub-periods. Although the average annual return of all equity mutual funds beats the S&P 500 index by almost 5.5% from 1976 to 1981, the portfolio of the top 5 funds using the strategy only outperformed the index by 0.4%. The performance improved slightly when the portfolio includes the top 10 funds, top 15 funds, or top 20 funds. During the other two sub-periods, however, the performance of such strategies was lagging behind substantially while the general equity fund as a whole underperforms the S&P 500 index portfolio. For example, from 1982 to 1986, the portfolio of the top 5 funds returned 7% less than the S&P 500 index portfolio. The gap is still more than 4% during the last sub-period from 1987–1991. Over the whole sample period, when the general equity funds slightly outperform the index fund, the strategy of buying the top *n* funds underperformed the index fund by more than 2.0%.

Perhaps the strategy of buying the top *n* funds with the largest significant alphas would have worked if used over a longer period. In Panel B of Table 1,

Table 1. Performance Comparison of Different Trading Strategies.

This table shows annual returns over each period using a strategy of buying funds with the most significant positive alphas calculated from previous five-year returns. In Panel A, portfolios are rebalanced every year, while portfolios are rebalanced every two years in Panel B.

	1976–1981	1982–1986	1987–1991	1976–1991
		Panel A: Hold for one year		
Top 5 Funds	10.96	12.59	10.46	11.31
Top 10 Funds	11.31	13.76	11.93	12.26
Top 15 Funds	11.11	14.87	12.58	12.73
Top 20 Funds	10.98	15.07	12.66	12.77
All Funds	16.14	17.41	13.63	15.74
S&P 500	10.56	19.80	15.30	14.87
		Panel B: Hold for two years		
Top 5 Funds	10.54	14.83	10.99	12.00
Top 10 Funds	11.21	14.65	12.78	12.77
Top 15 Funds	11.14	15.74	13.31	13.24
Top 20 Funds	11.34	15.67	13.31	13.29
All Funds	16.14	17.41	13.63	15.74
S&P 500	10.56	19.80	15.30	14.87

we have reported results based on the same strategy but held for the subsequent two years.[4] In general, the results are better than those in Panel A. However, the basic pattern remains. For example, over the whole sample period, the return from the portfolio of the top 5 funds was 12.00%, which is 2.87% lower than that of the index return. It is also possible that buying extreme CAPM-alpha funds might result in holding more volatile funds. For the portfolio of the top 5 funds, the average beta during the portfolio formation period is 1.02 with a standard deviation of 0.115. Therefore, under performance is not due to volatility. At the same time, if betas of the constructed portfolios are different from that of the index, our simple comparison will be invalid. On a risk-adjusted basis, we have also plotted the excess annual returns of holding the top 10 funds with the largest alpha in Fig. 1. The solid line and dotted line indicate results from rebalancing the portfolios every year and every other year, respectively. Apparently, the strategy only worked for five (six) years out of sixteen years when holding for one (two) year(s). Furthermore, these better performing years occurred mostly in the 1970s, which is consistent with Malkiel's (1995) finding.

Based on the evidence, one should question the overall consistency and the ability of alpha to predict excess returns of mutual funds. Ferson and Warther (1996) have argued that a non-zero alpha could also result from a shift in the risk exposure of a fund. When conditional expected future returns are high, managers would increase their holdings in risky stocks and would scale back their risk exposure when the expected future returns are low. If this is the case, our simple strategy should have also worked well. Only the interpretation of our performance, if there were any, would be different. Therefore, we suggest a different reason that could result in positive Jensen's alpha, the nonlinear compounding returns. In fact, the failure of the above strategy is consistent with this explanation.

2. COMPOUNDING RETURN AND JENSEN'S ALPHA

The CAPM is a linear model that imposes the following linear relationships between the excess return of a security and the market excess return:

$$r_{t,k} - r^f_{t,k} = \alpha_0 + \beta_0 (r^m_{t,k} - r^f_{t,k}) + \epsilon_{t,k} \tag{1}$$

where $\alpha_0 = 0$, and k corresponds to a short time period or high frequency, such as a day, a week, or a month, t corresponds to a regular (or low frequency) time period (a quarter or a year), and $r_{t,k}$, $r^f_{t,k}$, and $r^m_{t,k}$ are the short period rate of returns for a risky security, the riskless security, and the market portfolio, respectively.

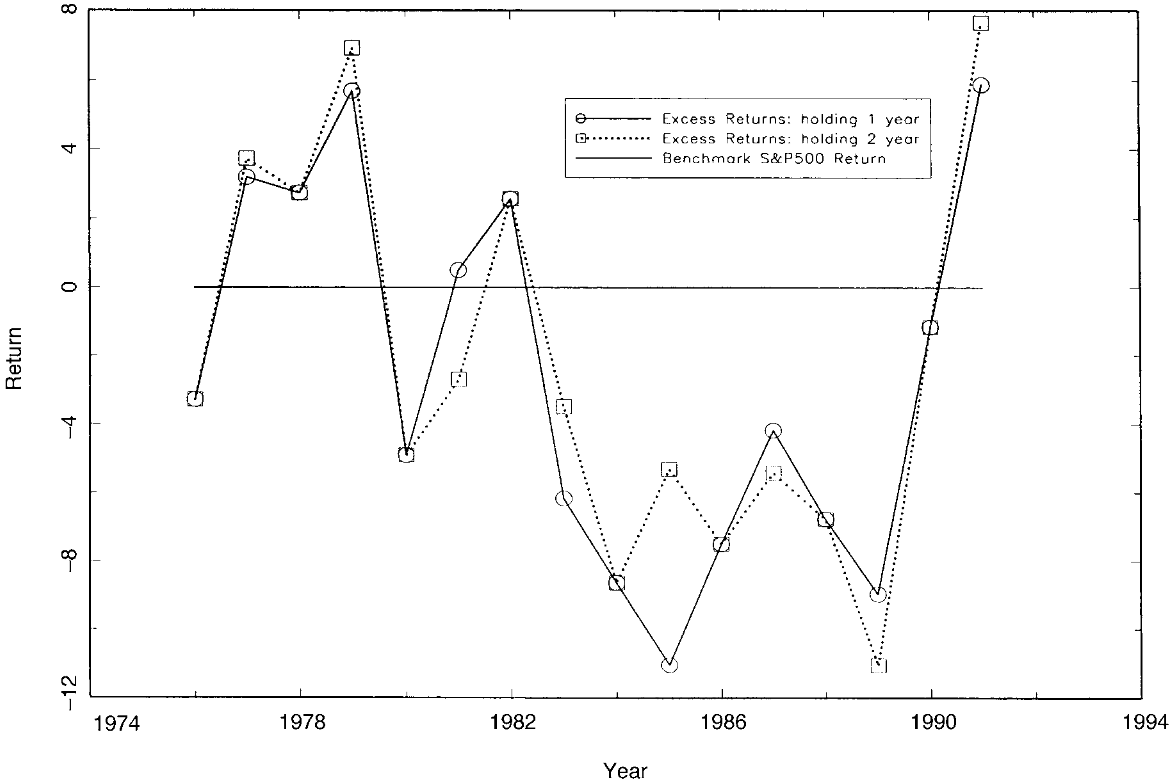


Fig. 1. Excess Annual Returns of Top 10 Mutual Funds (1976–1991).

The CAPM theory does not provide guidance in choosing a return horizon. For ease of exposition, we assume that the CAPM holds exactly for high frequency returns. When the model is estimated using low-frequency returns (such as annual data), will one still obtain unbiased estimates of α_0 (which should be zero) and β_0? The answer depends on the way annual returns are computed. If arithmetic sums of the high frequency returns are used, there is no bias because the arithmetic summation preserves the linearity of the model. However, if geometric (compounded) returns are used, which is how performance figures are generally stated, this nonlinear transformation will result in biased parameter estimators.

To illustrate this point, we take the special case of $k = 1, 2$ (think of $k = 1$ as the first half of year t and $k = 2$ as the second half of year t), and $\epsilon_{t,k}$ are *i.i.d.* normal. Define x_t to be the low frequency rate of return (for example, the annual return) computed using the geometric sum, i.e. $x_t = (1 + x_{t,1})(1 + x_{t,2}) - 1$, where x_t could be r_t, r^f_t, or r^m_t, and $x_{t,k}$ could be $r_{t,k}$, $r^f_{t,k}$, or $r^m_{t,k}$. If indeed $x_t = \log(1 + x_t)$, it is easy to show that the same relationship given in (1) will hold no matter what frequency returns are applied. Since x_t is just a first order approximation of $\log(1 + x_t)$, the assumed relationship is not true in general. If x_t is less than 10%, a magnitude roughly equivalent to the average long-term interest rate during the period covered, the approximation error between x_t and $\log(1 + x_t)$ is at most 0.47%. In this case, the arithmetic sum is a good approximation for the geometric sum. Yet as stock market prices fluctuate widely, returns can regularly fall to -20% or rise to $+20\%$. In this case, the approximation error is sizeable, about 2%. Through some algebraic manipulation, one can show the following relationship for low frequency returns:

$$r_t - r^f_t = \alpha + \beta(r^m_t - r^f_t) + u_t \tag{2}$$

where:

$$\alpha = \beta_0(\beta_0 - 1)(r^m_{t,1} - r^f_{t,1})(r^m_{t,2} - r^f_{t,2}) \text{ and}$$
$$u_t = (\beta_0(r^m_{t,2} - r^f_{t,2}) + r^f_{t,2} + 1)\epsilon_{t,1} + (\beta_0(r^m_{t,1} - r^f_{t,1}) + r^f_{t,1} + 1)\epsilon_{t,2} + \epsilon_{t,1}\epsilon_{t,2}.$$

Apparently, there is a non-zero intercept term α if β is assumed to be known, i.e. $\beta = \beta_0$. In particular, α depends not only on the β_0 coefficient, but also on the excess market return. There are two special cases where α is negligible. First, if β_0 equals one, the CAPM model reduces to an identity. Therefore, the model holds in any circumstance. In the second case, when the excess market returns are relatively small, the approximation discussed above is fairly good. If none of the above cases holds, we should expect to observe a significant intercept term α in the model estimation.

In practice, β is unknown in advance and is commonly estimated using OLS. It is useful to investigate the effect of nonlinear compounding on the OLS estimation of Eq. (2). Presumably, the estimated β will also change because it is impossible to constrain it to be the true value. Denote $r_t^{er} = (r_{t,1}^m - r_{t,2}^f)(r_{t,2}^m - r_{t,2}^f)$ as the compounded excess market return, and let $\hat{\alpha}$ and $\hat{\beta}$ be the OLS estimators of the corresponding α and β in Eq. (2). It can be shown that:

$$E(\hat{\alpha}) = \beta_0(\beta_0 - 1)\rho_\alpha \text{ and}$$
$$E(\hat{\beta}) = \beta_0 + \beta_0(\beta_0 - 1)\rho_{er.rm},$$

where

$$\rho_{er.rm} = \frac{Cov(r_t^{er}, r_t^m - r_t^f)}{Var(r_t^m - r_t^f)}, \quad \text{and}$$
$$\rho_\alpha = \bar{r}^{er} - \rho_{er.rm}(\bar{r}^m - r^f).$$

If $\rho_{er,rm}$ is positive, the OLS estimator of β will be either biased upward when β_0 is greater than one or biased downward when β_0 is less than one. In other words, we would mistakenly conclude that high risky assets (corresponding to large β) do not pay as much as they should, while low risky assets (corresponding to small β) pay more than they should according to their estimated betas. In this case, the estimated capital market line will be flatter than the theoretically proper relationship in Eq. (1). This is certainly a misleading inference due to the bias from the difference between geometric sum and arithmetic sum of stock returns.

Most significantly, starting from a model with a zero intercept, one could obtain a non-zero α estimate just by switching from, say, daily returns to quarterly returns, provided $\beta_0 \neq 1$. This offers a possible explanation for findings of non-zero Jensen's alphas in the previous section. Therefore, the observed Jensen's alpha could be attributed to the non-linearity in the geometric sum of returns (or equivalently the log-linear approximation). Based on this analysis, we should be cautious when interpreting Jensen's alpha as a measure of the abnormal performance of a fund. Instead, it may simply indicate a possible existence of bias in the beta estimator.

For the same reason, when we try to test the CAPM from the hypothesis of a zero intercept (i.e. $\alpha = 0$), we could reject the model statistically using certain frequency of returns while the model may actually hold using another frequency of returns. In recent years, the CAPM has been under serious debate. Since Fama and French (1992) have shown that the slope estimate from the cross-sectional regression of expected returns on betas is flat, they conclude that beta is not informative about risk inherited in a stock. The discussion in the previous section also sheds light on this issue. As small (large) betas tend to be even smaller (larger) after compounding returns, we may find a weak relationship between expected return and beta measure when the expected returns do not change. However, this is purely due to the nonlinear compounded returns.

The discussion about Eq. (1) is based on a special case of aggregating two periods. It can easily be extended to the case of aggregating l periods with:

$$r_t - r_t^f = \alpha + \beta(r_t^m - r_t^f) + u_t \tag{3}$$

where:

$$\alpha = \sum_{i=2}^{l} \sum_{\substack{k_1 < k_2 < \cdots < k_i \\ k_1, \cdots, k_i = 1, \cdots, l}} \beta_0(\beta_0^{i-1} - 1)(r_{t,k_1}^m - r_{t,k_1}^f) \cdots (r_{t,k_i}^m - r_{t,k_i}^f)$$

$$u_t = f(\epsilon_{t,1}, \epsilon_{t,2}, \cdots, \epsilon_{t,1})$$

Obviously, the high frequency returns $r_{t,k}^m$'s and $r_{t,k}^f$'s are unobservable when we only have low frequency returns such as r_t^ms and $r_{t,k}^f$s. Furthermore, the high frequency returns are correlated with the low frequency market excess return $(r_t^m - r_t^f)$, but are left in the residual term. Therefore, we have the classic faulty omission case in a regression context. Econometric theory suggests that the estimators of both α and β will be biased where,

$$E(\hat{\alpha}) = \sum_i \beta_0(\beta_0^{i-1} - 1) \sum_{k(i)} \rho_{k(i),\alpha}$$

$$\mathrm{E}(\hat{\beta}) = \beta_0 + \sum_i \beta_0(\beta_0^{i-1} - 1) \sum_{k(i)} \rho_{k(i),m}$$

with $\rho_{k(i),\alpha}$s being the correlation coefficients of missing regressors with the constant term, and $\rho_{k(i),m}$ being the correlation coefficients of missing regressors with the excess market return $(r_t^m - r_t^f)$. The same kind of bias emerges in the estimates. The above discussion continues to be valid. Therefore, we observe Jensen's α's in general.

In order to assess the empirical importance of the potential biases in both alpha and beta, we study first daily stock returns on the CRSP tape in the next section.

3. A COMPREHENSIVE LOOK AT BIAS IN ALPHA AND BETA ESTIMATES FROM STOCK RETURNS

Due to a limitation in the return frequency of the mutual fund sample, it is impossible to obtain an unbiased estimate of Jensen's alpha, if a non-zero alpha, in fact, exists. Therefore, it is difficult to study the significance of the bias from mutual fund data alone. One way to examine the issue is to use the daily stock returns recorded on the 1999 version of daily CRSP tape, which includes all the stocks traded on NYSE, AMEX, and NASDAQ. In particular, we study how alpha and beta estimates change when using ten years of monthly or quarterly compounded returns from daily stock returns. When a stock has twenty years of daily returns, we treat it as "two" stocks with ten-year records of returns. Similar treatment for stocks with thirty years of returns is employed. The analysis time period of ten years is chosen for the following reasons: first, as prior empirical studies have found that beta changes over time, we hope to obtain as stable beta estimates as possible; second, a ten-year period corresponds to forty quarters, which is reasonable to produce reliable OLS estimates; finally, we want to have as many stocks as possible included in our study. In order to estimate the market model, we use the value weighted NYSE/AMEX/NASDAQ composite index as the proxy for market portfolio and the three-month T-bill rate as the risk-free rate.

Some stocks have a very poor fit to the market model when using daily returns. At the same time, microstructure effects, such as non-synchronous trading and bid-ask bouncing, could severely alter the true distribution of returns. In order to reduce such effects, we use weekly returns computed by compounding five-day returns as the base case. Moreover, many small stocks are infrequently traded with zero recorded returns for some trading days. In these cases, it does not make sense to study the reliability issue of the alpha estimate. Therefore, we eliminate stocks with R^2s from the market model using weekly returns below 10%. With this criterion, we still have 5314 "stocks" in our sample over some ten-year period from July 1962 to December 1998. The monthly and quarterly returns are computed using 4 week and 12 week returns, respectively for simplicity.

Over the sample period, the average weekly return for the 5314 "stocks" is 0.348% as shown in the first row of Table 2. The average total volatility for those "stocks" is more than six times as high as the average return. The average idiosyncratic volatility, which is calculated from the residuals of the market model, is in turn twice as volatile as that of the index return. For monthly and quarterly index returns, the volatility is only about three times and two times

Table 2. Comparison of Alpha and Beta Estimates Based on Different Frequency of Returns.

This table shows average estimates from the market model fitted to stocks with ten years of weekly returns compounded using five daily returns (1999 version of CRSP tape). In order to obtain reliable estimates, only stocks with R^2 exceeding 10% from the market model using weekly returns are included. The market return and the risk-free rate are taken as the NYSE/AMEX/NASDAQ index return and the three-month T-Bill rate, respectively. When a stock has twenty years of daily returns, it is treated as two stocks with ten-year returns. Similar treatment for stocks with thirty years of returns is applied. The model is also fitted to corresponding compounded monthly (using four weekly returns) and quarterly returns (using three monthly returns). As a comparison, the same market model is run for monthly and quarterly returns using simple sum of weekly returns. All numbers reported in this table are in percentage.

	Geometric Sum, N = 5314			Arithmetic Sum, N = 5314	
	Weekly	Monthly	Quarterly	Monthly	Quarterly
Average Return	0.348	1.369	4.146	1.391	4.172
Index Std. D.	2.124	4.499	7.931	4.493	7.766
Avg. Idio.	4.863	9.008	15.10	8.875	14.34
Panel A: Individual Regression					
α	0.114	0.375	1.010	0.414	1.152
	(0.0027)	(0.0111)	(0.0346)	(0.0109)	(0.0334)
Significant	5.137	5.815	5.363	6.662	7.207
β	1.131	1.273	1.374	1.274	1.347
	(0.0054)	(0.0060)	(0.0079)	(0.0060)	(0.0075)
R^2	20.64	30.04	35.83	30.48	36.06
	(0.1146)	(0.1443)	(0.1947)	(0.1449)	(0.1971)
Panel B: Cross-Sectional Regression $\alpha_{month(or\ quarter)} = \gamma_0 + \gamma_1 \alpha_{week(or\ month)}$					
γ_1		1.0141	1.0233	1.0068	1.0024
		(0.00247)	(0.00284)	(0.00205)	(0.00264)
R^2		97.0	96.1	97.9	96.4
$\beta_{month(or\ quarter)} = \gamma_0 + \gamma_1 \beta_{week(or\ month)}$					
γ_1		0.978	1.104	0.978	1.053
		(0.0077)	(0.0097)	(0.0077)	(0.0093)
R^2		75.3	71.0	75.4	70.9

larger, respectively, than the corresponding average index returns. This is why the average R^2s from the CAPM model are relatively low with 20% for weekly returns, 30% for monthly returns, and 36% for quarterly returns. With such large volatilities, alpha estimates and beta estimates are likely to be biased using compounding returns.

On average, there is a highly significant alpha of 0.114% when weekly returns are fitted to the market model as indicated in Panel A of Table 2. If we examine individual stocks, 5.137% of stocks have significant alpha estimates at a 5% level. From a statistical perspective, alphas for individual stocks may not be significant. This is because, one would expect to see less than 5% of the individual stocks with significant alphas when choosing a 5% significance level. The average alpha remains positive and highly significant when either compounded monthly or compounded quarterly returns are used. However, the percentage of significant alphas does not increase much. In contrast, when arithmetic sum is used in computing monthly and quarterly returns, the percentages of significant alphas increase noticeably to 6.6% and 7.20%, respectively. Under normal distribution assumption, these numbers should be around 5% in theory when arithmetic sum is used. The observed phenomenon is attributable to large kurtosis in the return distribution. This is because kurtosis decreases substantially from high frequency returns to low frequency returns, which will in turn reduce the volatility. In the case of individual stocks, high frequency returns exhibit extremely large kurtosis. Therefore, we expect to see the percentage of significant alphas will increase when arithmetic sum is used as confirmed in our simulations (not reported). Compounded returns, on the other hand, are subject to a second effect as pointed out by Levhari and Levy (1977). The estimated volatilities increase faster than the means when return are compounded. This is evident when comparing the average idiosyncratic volatilities under geometric sum to that of under arithmetic sum. Therefore, without kurtosis in return, the percentage of significant alphas estimated from compounded returns should actually decrease. In the presence of both the kurtosis effect as evident from that of the arithmetic returns and the compounding effect, we only observe a slightly increase in the significant alphas when using compounded monthly or quarterly returns.

What is more important is to recognize that when we do observe significant alphas estimated from compounded returns, they tend to be "huge." In other words, it is interesting to estimate the magnitude of changing in alpha estimates. For that purpose, we perform cross-sectional regressions of monthly alphas on equivalent weekly alphas and quarterly alphas on equivalent monthly alphas. From Panel B of Table 2, we see that the relationship is very tight with R^2s more than 96%. Moreover, with compounded returns, the monthly alphas are 1.41% higher than that of the equivalent weekly alphas comparing with that of 0.68% when using arithmetic sum. Similarly, quarterly alphas are 2.33% larger than that of the equivalent monthly alphas using compounded returns while no changes can be found using arithmetic returns.

A more direct approach to study the significance of changing in alpha is by implementing a trading strategy as in the previous section. In particular, starting 1968, we estimate Jensen's alpha for each stock using previous five years of monthly returns. We then form a portfolio that includes 20 stocks with the most significant positive alphas and compute the compounded annual return for the portfolio during the current year. We repeat this process for the 32 years in our sample period. For comparison, we implement the same strategy based on Jensen's alphas estimated using previous two-year weekly returns.[5] The risk adjusted returns calculated using betas from the estimation period are shown in Fig. 2. Over the 32 years, excess returns from the trading strategy based on weekly returns (light solid line) are positive in half of time. Since significant Jensen's alphas estimated using monthly returns are more likely to be even greater in magnitude, the performance of our strategy ought to be improved. As the solid line demonstrates that we only observe positive excess returns in 12 of the 32 years. Moreover, when the whole sample period is evenly split into three sub-samples, the average excess returns are negative during the '70s and '90s and positive for the '80s. In summary, all the evidence has suggested that nonlinear compounding is an important factor that creates large alphas in practice. Therefore, it may not be beneficial to heavily rely on significant alphas in selecting mutual funds.

Compounded returns also affect the beta estimates as previous research has shown. Since we have eliminated stocks with small R^2s, the average beta estimates tend to be large, 1.131 for weekly returns, 1.273 for monthly returns, and 1.373 for quarterly returns as indicated in Table 2. In general, the beta estimate changes for any particular stock when using different frequency returns. We also report cross-sectional regression results of monthly betas on weekly betas and quarterly betas on monthly betas in Panel B of Table 2. The most significant changes in beta estimates occur when quarterly returns are compounded from monthly returns. In particular, quarterly beta estimates increase (decrease) by more than 10% from monthly betas when betas are greater (less) than one.

Of course, changes in the beta estimates can also be attributed to other factors. In the presence of the microstructure effect of non-trading discussed above, the average positive returns tend to be low and the average negative returns appear to be large, which will bias beta estimates downwards. This may have some effect when going from weekly returns to quarterly returns as

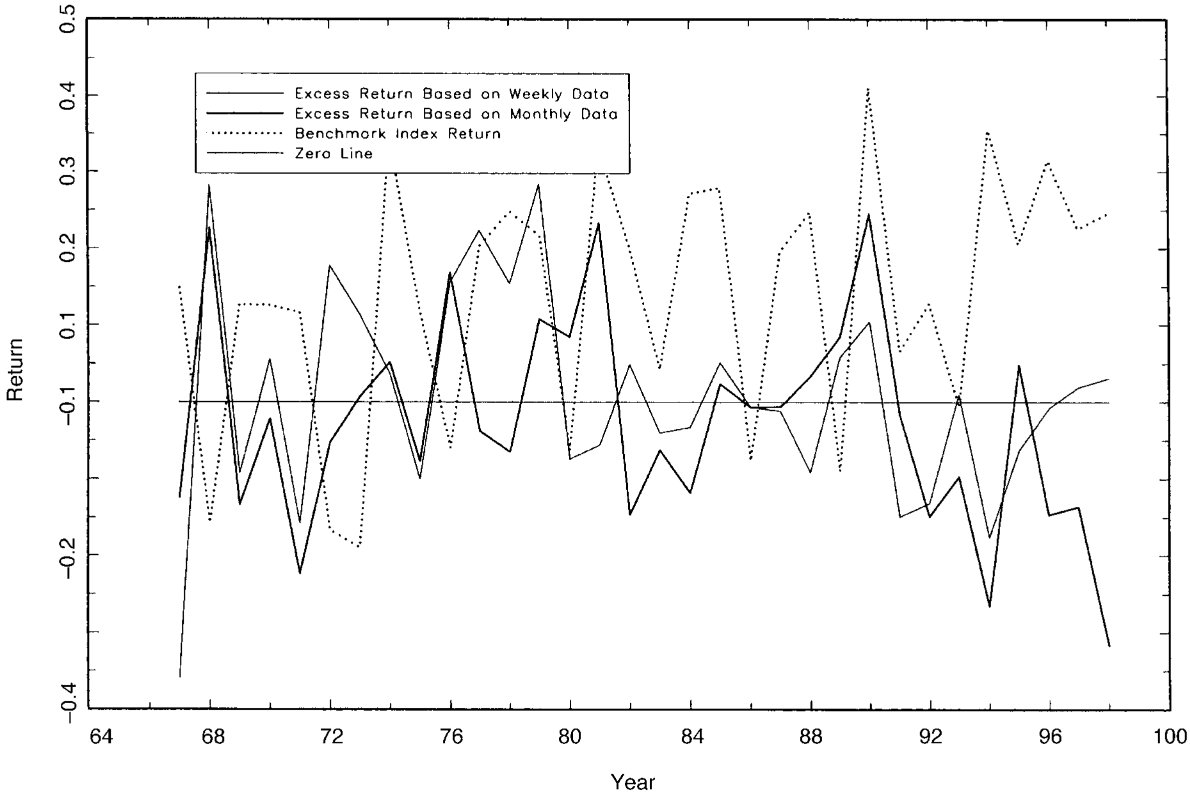


Fig. 2. Average Annual Performance of Top 20 Stocks (1968–1999).

indicated in cross-sectional regression of monthly betas on weekly betas. However, by comparing the cross-sectional regression results of quarterly betas on daily betas under geometric sum with that under arithmetic sum, we can be ensured that the microstructure effect is not an important factor. Big jumps in monthly beta estimates can also be explained if daily returns are positively correlated. Since there is little persistence in daily returns, this is probably not a major factor.

4. SOME OBSERVATIONS FROM MUTUAL FUND DATA

Results from the prior section are based on individual stocks which carry more idiosyncratic risks. In this section we study the actual return data for 206 mutual funds for which we have data to see if similar patterns exist. This sample corresponds to all general equity mutual funds that have consecutive quarterly records from the first quarter of 1971 to the last quarter of 1991.

Since we are unable to obtain high frequency return data, such as daily return or weekly return observations, we cannot determine whether the Jensen's alphas in our mutual fund sample are solely due to nonlinear compounding. Nevertheless, one can still study the nonlinear compounding effect by comparing the results from quarterly return data to those from annual returns; i.e. first estimate alphas and betas using quarterly return data then estimate the corresponding alphas and betas using annual return data, where annual returns are compounded from quarterly return figures. By comparing these two sets of alphas we can get a partial picture of the bias.

The scatter plot of differences in alphas between those based on annual returns and those from quarterly returns against quarterly alphas is shown in Fig. 3. Since the annual returns are compounded from quarterly return data, the quarterly alphas have been converted using $(1+\alpha)^4-1$ in order to be comparable. The solid line is the regression line. If geometric compounding has no effect on alpha estimates, one would expect to see a horizontal line. To the contrary, we observe a positively sloped line. This indicates that positive alphas found in regressions using quarterly returns are generally found to be more positive in regressions using annual return data. At the same time negative quarterly alphas induce more negative annual alphas. This effect could not be due to accumulating quarterly alphas, since the quarterly alphas have been "compounded" on the same scale as annual alphas. The patterns shown in the graph are very similar to those we find in the individual stock return discussion.

Statistically, we can relate changes in the estimates through regressing annual alphas against quarterly alphas,

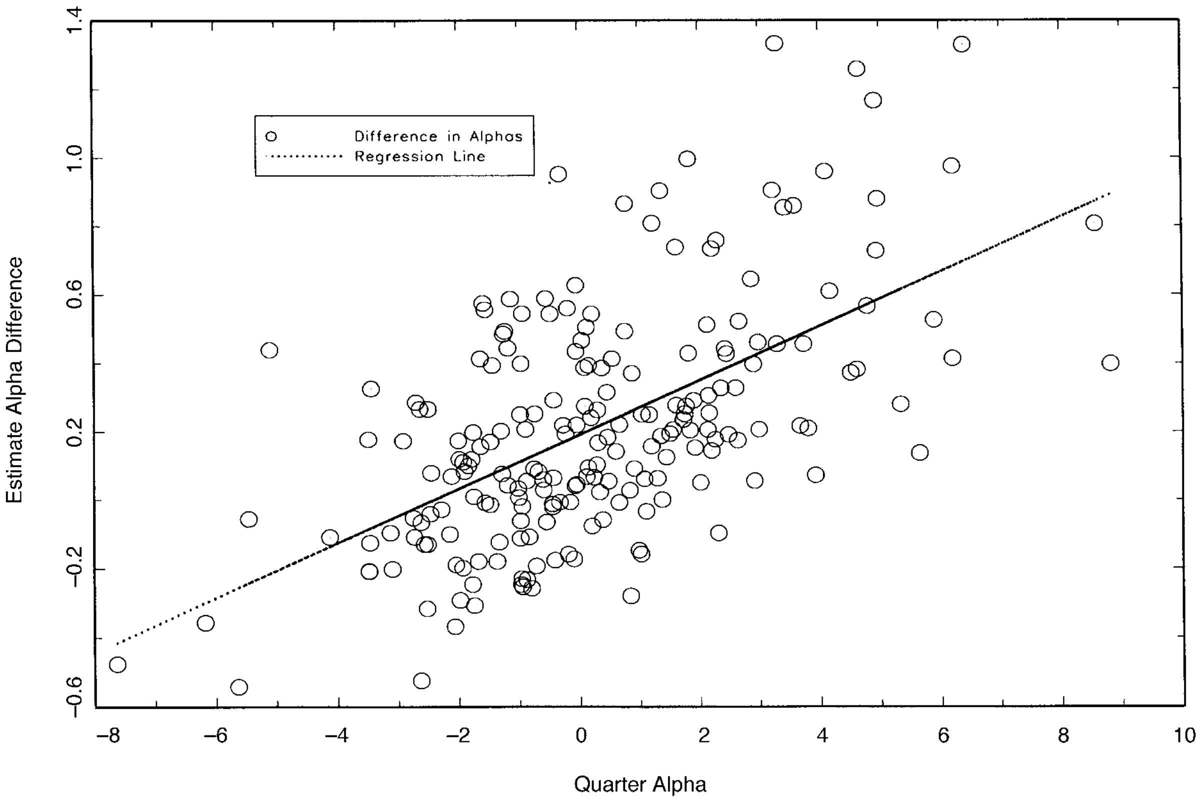


Fig. 3. Estimated Difference Between Annual Alpha and Quarterly Alpha.

$$\alpha_{annual,i} = 0.191 + 1.079 \ \alpha_{quarter,i} \quad R^2 = 0.990,$$
$$(0.0194) \quad (0.0076)$$

where, $\alpha_{quarter,i}$ is estimated from quarterly returns but converted to an annual scale, and the number in the parentheses are standard deviations. Again the slope coefficient would be one if there were no compounding effect. Obviously, the slope coefficient estimate is statistically significant and above one. There has been a 7.9% shift in the alpha. About 12% of the funds have statistically significant negative alphas and about 11% of the funds have significant positive alphas. Based on this finding, one may conclude that the nonlinear compounding effect does in fact exist in real return data for mutual funds.

The geometric compounding not only affects Jensen's alpha but also alters the magnitude of the beta estimate. We would have seen a weak relationship between geometric average returns and beta estimates if we had actually plotted them. This is not true for the arithmetic returns as shown in the following regressions since arithmetic sum preserves linear relationship. Again the numbers in the parentheses are standard deviations.

$$R_i^G = 8.03 + 3.91 \ \beta_i^G \quad R^2 = 0.064,$$
$$(1.03) \quad (1.04)$$
$$R_i^A = 7.41 + 6.02 \ \beta_i^A \quad R^2 = 0.150,$$
$$(0.996) \quad (1.00)$$

where R_i^G and R_i^A are the geometric compounded annual return and the arithmetic annual return, respectively. Similarly, β_i^G *and* β_i^A are estimated from compounded annual returns and simple sum of quarterly returns. Apparently, the R^2 in the first equation is half the magnitude as that in the second equation, although both the slope coefficients are statistically significant from zero. Thus, we would have concluded that there is a weak relationship between return and beta if the geometric average of the return is used. However, we can say that there is a strong positive relationship between beta and return if we examine the second estimated equation where the arithmetic average return is used. Again the problem comes from the conflict between nonlinear compounding and the linear CAPM.

5. CONCLUDING COMMENTS

In this paper attention have been called to an inconsistency in the practice of applying the CAPM. The inconsistency comes about from the practice of applying a linear model to nonlinear compounded return data. Through theoretical discussion and investigation in both individual stock returns and mutual fund returns, it is shown that the non-linearity in compounded returns can both create a Jensen's alpha and cause a shift in the risk measure beta. Although it cannot be concluded that the Jensen's alpha observed in mutual fund data is solely caused by this effect, cross-sectional regression suggests that the equivalent alpha estimates increase by about 8% from quarterly returns to annual returns. Therefore, it is certainly a driving factor. In order to be cautious, other measures are also needed in evaluating mutual fund performance. It is also interesting to point out that strategies based on picking stocks or funds with large alphas do not work as we have demonstrated. Furthermore, the finding of a poor relationship between average returns and betas may also partially be caused by this effect. Therefore, it is necessary to distinguish between short-run and long-run measures of risk. Beta may well be a more useful measure for short-run risk than for longer-run risk.

NOTES

1. This sample is provided by Lipper Analytic Services, which consists of quarterly returns of all the equity mutual funds that survived over the period.
2. In contrast, the R^2 from quarterly returns of individual stocks is relatively small (about 36%).
3. As pointed out by Grossman and Stiglitz (1980), it is perfectly natural for a mutual fund to earn an excess return which just compensates for information acquiring expenses. Therefore, in this paper, as in similar research, we use after-expense returns.
4. Many other strategies have been tried such as buying one year best, two year best, three year best, and four year best and then holding for the next three years, next four years, or next five years, but the general patterns are the same.
5. For simplicity, we assume that there are 21 trading day when compounding for monthly returns. In order to be compatible, we assume there are 50 trading weeks in a year and 5 trading days in a week except for the last week of a given year where 7 trading days are used.

ACKNOWLEDGMENTS

I am indebted to C. F. Lee (the editor), Larry Merville, Georg Noldeke, Chris Paxson, Sanjai Bhagat, Harald Uhlig and especially Burton Malkiel and the anonymous referee for their invaluable comments and kind help as well as to Lipper Analytic Services for providing the data sets.

REFERENCES

Black, F., Jensen, M., & Scholes, M. (1972). The capital Asset Pricing Model: Some empirical tests. In: M. Jensen (Ed.), *Studies in the Theory of Capital Markets* (pp. 79–121). New York: Praeger.

Brown, S., Goetzmann, W., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship Bias in Performance Studies. *The Review of Financial Studies*, *5*, 553–580.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, *52*, 57–82.

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

Christopherson, J. A., & Turner, A. L. (1991). Volatility and Predictability of Manager Alpha: Learning the Lessons of History. *Journal of Portfolio Management*, Fall, 5–12.

Christopherson, J. A., Ferson, W. E., & Glassman, D. A. (1998). Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance. *Review of Financial Studies*, *11*, 111–142.

Connor, G., & Korajczyk, R. (1991). The Attributes, Behavior, and Performance of U.S. Mutual funds. *Review of Quantitative Finance and Accounting*, *1*, 5–26.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring Mutual Fund Performance with Characteristic-Based Benchmarks. *Journal of Finance*, *52*, 1035–1058.

Day, T., Wang, Y., & Xu, Y. Investigating Underperformance by Equity Mutual Fund Portfolios. Working Paper, 2000, School of Management, The University of Texas at Dallas.

Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios. *Review of Financial Studies*, 1*6*, 1–22.

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, *46*, 1575–1617.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance*, *47*, 427–465.

Ferson, W. E., & Warther, V. A. (1996). Evaluating Fund Performance in a Dynamic Market. *Financial Analysts Journal*, *52*(6), 20–28.

Grinblatt, M., & Titman, S. (1989). Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings. *Journal of Business*, *62*, 393–416.

Grossman, S., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, *70*, 393–408.

Handa, P., Kothari, S. P., & Wasley, C. (1989). The Relation Between the Return Interval and Betas: Implications for the Size Effect. *Journal of Financial Economics*, *23*, 79–100.

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974–1988. *Journal of Finance*, *48*, 93–130.

Ippolito, R. A. (1989). Efficiency with Costly Information: A Study of Mutual Fund Performance, 1965–84. *Quarterly Journal of Economics*, *104*, 1–23.

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945–1964. *Journal of Finance*, *23*, 389–416.

Kothari, S. P., Shanken, J., & Sloan, R. G. (1995). Another Look at the Cross-Section of Expected Stock Returns. *Journal of Finance*, *50*, 185–224.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992). The Structure and Performance of the Money Management Industry. *Brookings Papers: Microeconomics*, 339–391.

Levhari, D., & Levy, H. (1977). The Capital Asset Pricing Model and the Investment Horizon. *Review of Economics and Statistics*, *59*, 92–104.

Longstaff, F. A. (1989). Temporal Aggregation and the Continuous-Time Capital Asset Pricing Model. *Journal of Finance*, *45*, 871–887.

Malkiel, B. G. (1995). Returns From Investing in Equity Mutual Funds 1971–1991. *Journal of Finance*, *50*, 549–572.

Merville, L. J., & Xu, Y. (2000). The Changing Factor Structure of Equity Returns. Working Paper, School of Management, The University of Texas at Dallas.
Sharpe, W. (1966). Mutual Fund Performance. *Journal of Business*, *39*, 119–138.
Treynor, J. (1965). How to Rate Management of Investment Funds. *Harvard Business Review*, *43*, 63–75.
Wermers, R. (2000). Mutual Fund Performance: An Empirical Decomposition into Stock-Piling Talent, Style, Transaction Costs and Expenses. *Journal of Finance*, *55*, 1655–95.

This Page Intentionally Left Blank

MARKET TIMING SKILL, EXPECTED RETURNS, AND MUTUAL FUND PERFORMANCE

Jason T. Greene and Charles W. Hodges

ABSTRACT

This paper examines how the interaction between imperfect market timing skill and trading frequency affects the return distribution for market timing strategies in open-end mutual funds. We demonstrate that traders need only a modest level of skill (to be correct only 54% of the time) in order to beat the market when employing a daily *timing strategy. We then examine how market timers may dilute the returns of a fund's buy-and-hold investors. Our results suggest that market timers who trade mutual funds can have a significant impact on a fund's reported performance.*

1. INTRODUCTION

Recently several authors have examined whether some mutual fund investors are actively timing open-end mutual funds (Chalmers, Edelen & Kadlec (1999), Bhargava, Bose & Dubofsky (1998), Bhargava & Dubofsky (2000), and Greene & Hodges (2000)). Chalmers, Edelen and Kadlec (1999) argue that a "wildcard option" exists in a broad cross-section of domestic and international open-end mutual funds as a result of stale prices'.[1] These stale prices give rise to profitable trading strategies in which active traders could earn abnormal returns of 10% to 20% annually. Bhargava and Dubofsky (1998, 2000) show

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 183–204.

ISBN: 0-7623-0798-6

that skilled trading is possible in at least three U.S. based equity funds holding substantial amounts of foreign equities (from now on, international) mutual funds. Greene and Hodges (2000), in their examination of over 900 open-end mutual funds of all types, have four key findings related to market timing of open end mutual funds:

(1) Abnormal returns averaging about 20% per year from a market timing strategy exist in their sample of 109 international mutual funds.
(2) Daily net mutual fund flows in international mutual funds, in particular, are consistent with skilled active traders who take advantage of stale prices. International equity funds have about $2\frac{1}{2}$ times greater daily fund flows than domestic equity funds, with average net fund flows of about 0.9% per day for the international funds.
(3) These skilled market timers earn their abnormal returns at the expense of a fund's buy-and-hold investors, something they call a dilution impact. The Green and Hodges results show a significant negative effect on the reported returns for international funds of nearly 0.50% on an annualized basis.
(4) Efforts by mutual funds to restrict market timing showed mixed results. The level of daily fund flows in international mutual funds was not significantly affected by the following; language stating market timing was not allowed, redemption fees, or limits on the number of fund exchanges per year. Minimum holding periods were associated with reduced fund flows, while front-end loads were associated with higher fund flows.

In summary, these authors present persuasive evidence than active market timers can use open end mutual funds, and simple market timing rules, to earn abnormal returns.[2]

While the above papers demonstrate the feasibility of actively timing open-end mutual funds, and provide evidence that frequent, possibly daily, fund exchanges are occurring, the issue of when investors should follow market timing trading strategies remains in dispute. It is clear that in the absence of transactions costs, a perfectly skilled timer who can predict all up and down markets should engage in market timing. However, little evidence exists that characterizes the expected rewards or risks involved in *imperfect* market timing. Furthermore, the related issue of how often market timers should trade on their imperfect timing information also has received little attention. To resolve these issues, we address the following general questions:

(1) What level of market timing skill is needed in order to expect to earn above average returns compared with a buy-and-hold strategy (even after adjusting for risk)?

(2) How does the frequency of trading impact the returns and risks of skilled and unskilled market timers?
(3) Given a trading frequency and an observed statistical relationship or correlation between a market timing signal and equity returns, what are the expected returns from following the market timing signal?
(4) What impact do market timers have on the mutual funds used as timing vehicles?

We address the viability of market timing using simulations based on actual equity market returns data. These simulations show how timing skill and timing frequency interact to affect expected returns, risk, and the likelihood of outperforming a buy-and-hold strategy. Unlike previous studies, which consider monthly timing to be the highest trading frequency, this paper explores the viability of *daily* market timing strategies. We show that daily market timing can be profitable for traders who can correctly time the market only 54% of the time. These daily timers need only possess a rather weak signal about subsequent equity market movements.

Sharpe (1975) provides the original analysis of returns to market timers. He demonstrates that market timers trading yearly would need a high degree of accuracy of 75% to match the expected returns of an investor who adopts an all-equity buy-and-hold strategy. We show that Sharpe's original analysis remains true: yearly market timers must possess a high degree of skill in order to expect to time the market profitably on an *ex ante* basis. Several other authors examine market-timing strategies that provide above average returns on a risk-adjusted basis. For example, Chua, Woodward and To (1987), Clarke, Fitzgerald, Berent and Statman (1990), Larsen and Wozniak (1995), Sy (1990), and Wagner, Shellans and Paul (1992) examine market timing strategies where timers trade at annual or monthly intervals. Under these frequencies of trading, timers with substantial skill can earn modest profits.

We doubt anyone would find it surprising that timing skill helps determine the potential gains or losses to following a market timing strategy. Indeed, it seems clear that a more skilled timer will earn a higher reward from market timing than will a less skilled trader. However, this is not the case. While Benning (1996) hints at this result in discussing how it is possible for less skilled traders to earn larger *realized* profits, we show that under certain circumstances (unrelated to market conditions) less skilled traders earn larger *expected* profits. The factor that reconciles this apparent puzzle is trading frequency. In particular, we show how timing skill and frequency affect the return-to-risk relationship compared with an investment along the Capital Market Line. Our paper demonstrates that the gains to market timing increase significantly when trading frequency increases. So, less skilled timers who trade more frequently are able to earn higher expected returns (with lower risks) than more skilled traders who trade less frequently.

Our overall conclusion is that traders who have the opportunity to costlessly trade daily and possess a modest skill or signal for predicting day-to-day movements in the equity markets will find market timing to be a profitable investment strategy.[3] On the issue of how the presence of market timers impacts the returns to the funds that they trade, we show that the skill with which a timer trades and the proportion of the fund's assets that they trade jointly determine the impact on the return to the fund. In particular, if a market timer uses an open-end mutual fund to engage in a profitable trading strategy, then this timer earns her gain at the expense of the mutual fund (i.e. its non-timing shareholders) through the dilution of positive returns. Therefore, funds that are used as timing vehicles by good market timers suffer decreased performance. This paper is organized as follows. Section 2 develops the paper's methodology. Section 3 presents our results and Section 4 concludes the paper.

2. RETURNS FROM MARKET TIMING

Most authors, beginning with Sharpe (1975), consider a model of two assets. The primary asset of interest is a risky portfolio, which we take to be a diversified portfolio of equities. The second asset, a money market account or cash, is simply a safe haven for the market timer when the equity market is not providing positive returns. The successful market timer predicts when the equity market wil1 generate a return below that of the money market. In these "down" markets, the timer holds cash, earning the money market return. In "up" markets the timer holds an equity portfolio or index, earning the higher return.

We adopt a methodology similar to that of Chua, Woodward and To (1987), who simulate yearly returns by generating a random distribution that matches the moments of the observed distribution of yearly returns. In contrast, we simulate the returns to market timing by using actual daily returns from the CRSP value-weighted index (with dividends) from July 3, 1962 through December 31, 1997. In order to assess the *expected* risks and rewards from market timing rather than simply measuring the *historical* risks and rewards, we take the following approach. We create 1,000 years of 252 daily equity index returns by drawing with replacement from the daily CRSP value-weighted index returns. This bootstrap method allows us to use the historical information about the distribution of daily returns, without restricting the distribution of monthly, quarterly, or yearly returns. We also create a corresponding series of money market returns, with the simplifying assumption of a daily-compounded annualized yield of 4%.[4]

From the daily returns, we generate portfolios based on a market timer's skill and trading frequency. In this paper, a timer's skill is captured in a single variable, s, where $0 \leq s \leq 1$. A skill level of one ($s = 1$) indicates a *perfectly good* market timer. That is, a timer who is always in the market (invested in equities) when it offers a return above the money markets, and in the money market when equities offer a lower return. A skill of zero ($s = 0$) represents a *perfectly bad* market timer. In this case, the market timer is always wrong. The perfectly bad market timer is invested in equities only when they offer a return below the money market. The market timer with no skill at all is indicated by a skill level of 50% ($s = 0.50$). This timer has an even odds chance at correctly predicting up or down markets. The intended interpretation of our skill measure is a probability of being correct about market movements. The perfectly good trader is always correct ($s = 1$), the perfectly bad trader is always incorrect ($s = 0$), and the unskilled trader ($s = 0.50$) may as well flip a fair coin to decide whether or not to hold equities. We allow s to vary continuously on the interval (0, 1), so that $s > 0.50$ indicates traders with some good (positive) market timing skill. Traders with $s < 0.50$ are traders who are, on average, wrong about market movements. We note here that this definition of skill treats type I and type II errors as equally important. That is, a trader can predict up or down markets with the same skill (or lack thereof).[5]

For each skill level, we also form portfolios based on a trader's frequency of trading (or timing). A daily timer decides each day whether to hold a position in the money market or the equity index. To simulate this procedure, each sample day we draw a number u from the random uniform distribution on the interval (0, 1). If $u \leq s$, then the trader holds the asset that offers the highest return on that day. Otherwise, the trader holds the asset that offers the lowest return on that day. For example, suppose that the skill is 60% ($s = 0.60$), we draw $u = 0.51$, and the day's market index return is 0.25%. The equity index return of 0.25% is higher than the money market return of approximately 0.0157%. Since $u \leq 0.60$, this trader will hold the equity index on this sample day. On average the trader with a skill of $s = 0.60$ will be "correct" (i.e. holding the asset with the highest return) on 60% of the days. Note that the perfectly good daily market timer is always earning the highest possible return between these two assets each day. We convert these daily returns into annual returns assuming daily compounding and calculate annual portfolio returns for each market timing skill level ($0 \leq s \leq 1$ in increments of 0.01). This procedure results in a sample of 1,000 annual returns for each skill level of daily timers.

We perform a similar procedure in forming portfolios for monthly, quarterly, and yearly market timers. For monthly timers, we first convert the daily equity index (money market) returns into monthly equity index (money market) returns. For each skill level in each month, we draw u and determine the holdings (equity index or money market) for that month. The monthly returns are then converted to annual returns, assuming monthly compounding. An analogous procedure is followed to obtain samples of 1,000 annual returns at the quarterly and yearly frequencies for various skill levels between 0 and 1. These 1,000 years comprise the distribution returns from market timing, given the *trading frequency* and *timing skill*. We compare the distribution of these 1,000 years of timing returns across skill levels, across trading frequencies, and with the distribution of 1,000 years of non-timed equity index returns.

3. TIMING SKILLS AND SIGNALS

3.1. Interpreting the Relationship between an Observed Signal and Timing Skill

In this section, we demonstrate the relationship between market timing skill and the statistical measure of serial correlation. Consider a market timer who relies on a signal x_t, received one period before time t. The timer uses this signal to make her decision regarding her investment holdings in time t. In other words, she trades at time $t-1$ in order to have her desired holdings at time t. Let a signal of $x_t=1$ indicate that the trader should be invested in the equity market index for the next period and $x_t=0$ indicate that the trader should be invested in the money market in period t. Let r_t be an indicator of the equity market index return at time t, where $r_t=1$ indicates when the equity index return exceeds the money market return, and $r_t=0$ indicates when the equity market index return is less than the money market return. In the special case where the distribution of excess returns is symmetric about zero (i.e. the probability that $r_t=1$ is 50%), the skill level described above is a simple function of the correlation between the signal x_t and the return indicator r_t. Defining $\rho_{x,r}$ be the correlation coefficient between x_t and r_t, results in a skill level of

$$s=\frac{\rho_{x,r}+1}{2}. \tag{1}$$

Alternatively, we can use Eq. (1) to solve for the correlation coefficient between the signal and the equity market indicator as a function of the skill level,

$$\rho_{x,r} = 2s - 1. \tag{2}$$

For example, a daily timer with a skill of 60% is following a timing signal that has a correlation with the equity market indicator of 0.20.[6] This equation is a close approximation in other cases when the probability that $r_t = 1$ is close to 50%. Since historically, daily returns are positive on about 52% of trading days, this approximation serves as a nice rule of thumb and creates an easy means of testing for statistical significance of one's market timing system. Notice that our treatment of skill level conforms to the usual notion of a probability measure ($0 \leq s \leq 1$), resulting in a properly specified correlation coefficient ($-1 \leq \pi \leq 1$).

3.2. Market Timing Returns by Trading Frequency and Timing Skill

Table 1 contains the results from the simulation of market timing returns. As expected, the average return increases monotonically in skill level within each timing frequency. For daily timers with perfectly bad skill, the average annual return is -46.24% with a standard deviation of only 4.22%. The unskilled daily timer ($s = 0.50$), has an average return of 5.68%, while the perfectly good daily timer has an average return of 118.24% per year. The table demonstrates the relationship between the trading frequency and the distribution of returns. For perfectly good timers, the rewards are clearly highest for those who trade daily, while the yearly timers have the most modest opportunities. Similarly, the losses for perfectly bad timers are also highest for daily timers. Most importantly, the breakeven skill levels differ across trading frequencies.

Figure 1 graphs the excess average return relative to the market index return as a function of skill level. Considering breakeven to be an average return at least as high as the market index's average return, an excess average return of zero represents the breakeven point. A key result is that breakeven skill decreases as trading frequency increases. From Fig. 1 and Table 1, we see that the breakeven skill for a daily trader is between 53% and 54%. This corresponds to a signal correlation of only approximately 0.08. For a monthly timer, the breakeven skill rises to over 62%. Quarterly and yearly timers have breakeven skill levels of approximately 68% and 83%, respectively. While the yearly timer's required skill of 83% is the same order of magnitude, it appears to be somewhat higher than Sharpe's (1975) finding of near 75%. We note that the market index's higher average return in the more recent years (since Sharpe's work) may account for this difference.

The columns reporting the "percentage of years beating the market index return" in Table 1 show breakeven skills are not driven by outliers in the distribution (which might cause the mean to be an upwardly biased estimate of

Table 1. Average Annual Returns from Daily, Monthly, Quarterly, and Yearly Market Timing.

This table reports annual returns from market timing strategies executed at different frequencies (daily, monthly, quarterly, and yearly). Skill level is measured as the proportion of time that the market timer correctly predicts the direction of the equity index, relative to an investment in the money market (assumed to give a 4% annualized return). We average across 1,000 years of returns, which are simulated by drawing (with replacement) 252 days per year from actual daily market index returns from July 3, 1962 through December 31, 1997. Standard deviations are reported below the means in parentheses. We use the CRSP Value-weighted Index as the equity index return. The percentage of years for which the market timing strategy return beats the market index return is also reported. Returns in boldface indicate that the timing strategy average return is above the average equity index return. The average equity index return is 13.04% with a standard deviation of 14.18%.

	Average Annual Percentage Returns (Standard Deviation)				Percentage of Years Beating the Market Return		
Skill	Daily	Monthly	Quarterly	Yearly	Daily	Monthly	Quarterly
0.00	− 46.24%	− 9.24%	− 2.70%	1.88%	0.0%	0.0%	0.0%
	(4.22)	(6.04)	(5.99)	(4.76)			
0.10	− 41.36	− 6.01	− 0.56	3.35	0.0	1.1	4.5
	(4.74)	(7.02)	(7.21)	(7.29)			
0.25	− 28.55	− 0.98	2.96	5.18	0.0	5.9	13.0
	(6.11)	(8.04)	(8.56)	(9.04)			
0.50	5.68	8.06	8.62	8.57	22.5	32.2	32.2
	(9.51)	(9.46)	(9.88)	(10.92)			
0.51	7.37	8.48	8.88	8.66	28.4	33.7	33.3
	(9.78)	(9.49)	(9.98)	(10.97)			
0.52	9.70	8.83	9.03	8.76	37.3	35.2	33.9
	(9.95)	(9.46)	(9.98)	(11.01)			
0.53	11.39	9.2	9.32	8.84	43.8	36.6	35.2
	(10.10)	(9.54)	(10.15)	(10.99)			
0.54	**13.29**	9.54	9.53	8.97	**52.4**	37.6	35.9
	(10.33)	(9.59)	(10.17)	(11.04)			
0.55	**15.22**	9.90	9.70	9.17	**61.0**	39.0	36.4
	(10.51)	(9.60)	(10.23)	(11.12)			
0.60	**26.02**	12.11	10.88	9.86	**89.8**	47.9	40.6
	(11.59)	(9.99)	(10.34)	(11.14)			
0.75	**58.33**	**18.40**	**14.48**	12.14	**100.0**	**74.1**	**55.8**
	(13.26)	(10.74)	(10.49)	(11.37)			
0.90	**91.44**	**25.02**	**18.08**	**13.84**	**100.0**	**93.8**	**74.4**
	(14.99)	(10.71)	(10.42)	(11.40)			
1.00	**118.24**	**29.23**	**20.55**	**15.17**	**100.0**	**99.7**	**85.6**
	(16.25)	(10.55)	(10.33)	(11.45)			

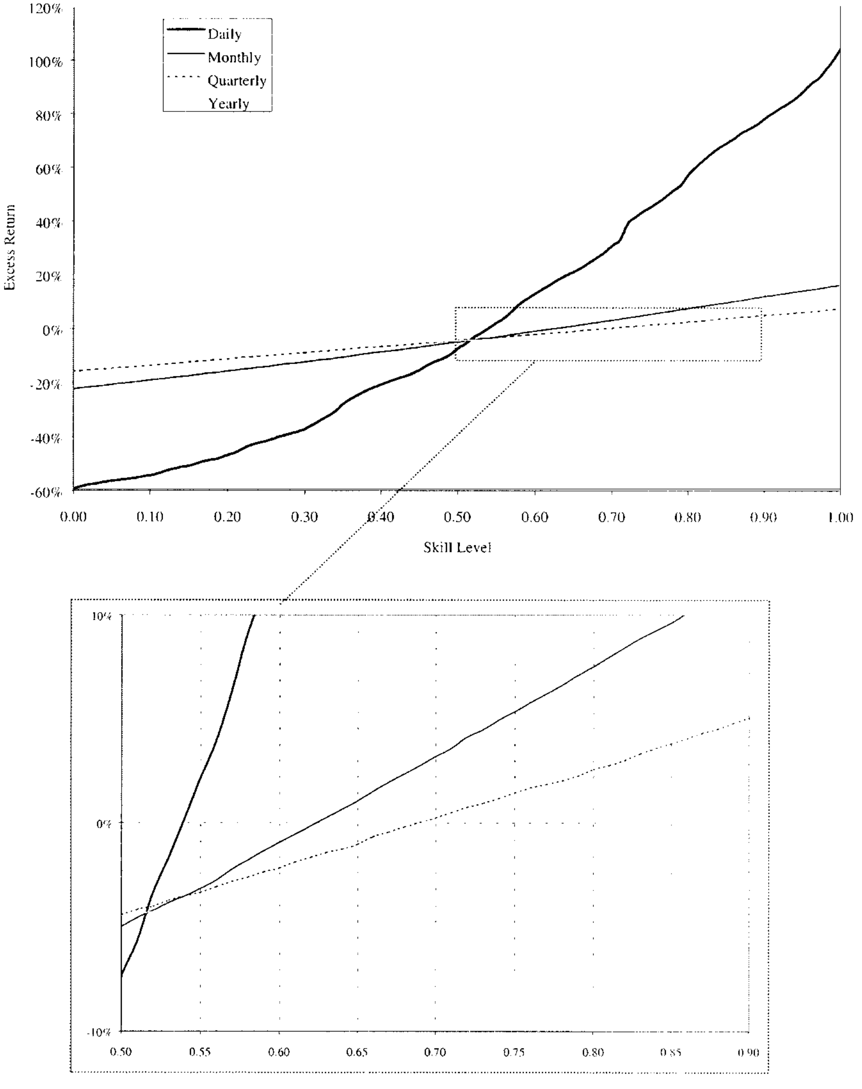


Fig. 1. Excess Market Timing Returns for Various Trading Frequencies and Skill Levels.

This figure plots the average excess returns (above returns from the equity index) from a market timing strategy with the given trading frequency. The skill level represents the probability that the market timer correctly identifies successful market timing opportunities.

the location of the distribution). These results can also be seen in Fig. 2, which shows the percentage of years that a market timing strategy beats the market index return as a function of skill level. Breakeven in this case is beating the market index return in 50% of the years. This figure shows how the frequency of trading affects the viability of market timing. For daily timers, there is almost no chance of beating the market index return unless the skill level is above 40%. However, with a skill level of 71% (a signal correlation of 0.42), the daily timer beats the market index return in every year. In contrast, a quarterly timer with a 71% skill level beats the market index in only about 50% of the years.[7]

One of the perceived risks associated with market timing is the increased risk of being out of up markets (i.e. invested in cash while equities earn positive returns). Among others, Schilling (1992, page 46) documents that “stocks haven’t appreciated in a smooth fashion, but in spurts, and the investor who is in and out of the market risks being out at times of great appreciation”. In Wall Street terminology, this is sometimes expressed as “market timers must be right twice, once when they exit the stock market and the second time when they reenter the market”. The penalty for being in and out of the market on a frequent basis when there is no skill is demonstrated in Table 1 (and Fig. 2) by the fact that percent of years that the market timer beats the equity index decreases as timing frequency increases. In effect, the penalty increases as the frequency increases since losses are compounded more frequently.

Sy (1990) points out that a decrease in portfolio volatility (relative to an all equity portfolio) associated with spending long periods invested in money market securities partially offsets this risk of decreased returns. For example, an unskilled market timer whose timing strategy results in spending 50% of the time in equities and 50% of the time in cash will have a return level similar to a buy-and-hold investor who attempts to hold a balanced portfolio of 50% equities and 50% cash. However, Clarke and de Silva (1998) note that the decrease in volatility is not this simple. They show that there is slightly higher volatility for a portfolio that combines cash and equities by holding them each in different states compared with a buy-and-hold portfolio that fixes the mix of cash and equities across all states. This risk-return tradeoff suggests that a market timer’s success or failure should be measured on a risk-adjusted basis. Since the Capital Market Line ranks as the prevalent risk-adjusted benchmark, we next compare the market timing portfolio’s return on a portfolio of the equity index and money market security with the same level of risk as our market timing portfolio.[8]

The virtues of successful market timing appear even stronger when considering the risk of the distribution of market timing returns. Figure 3 shows

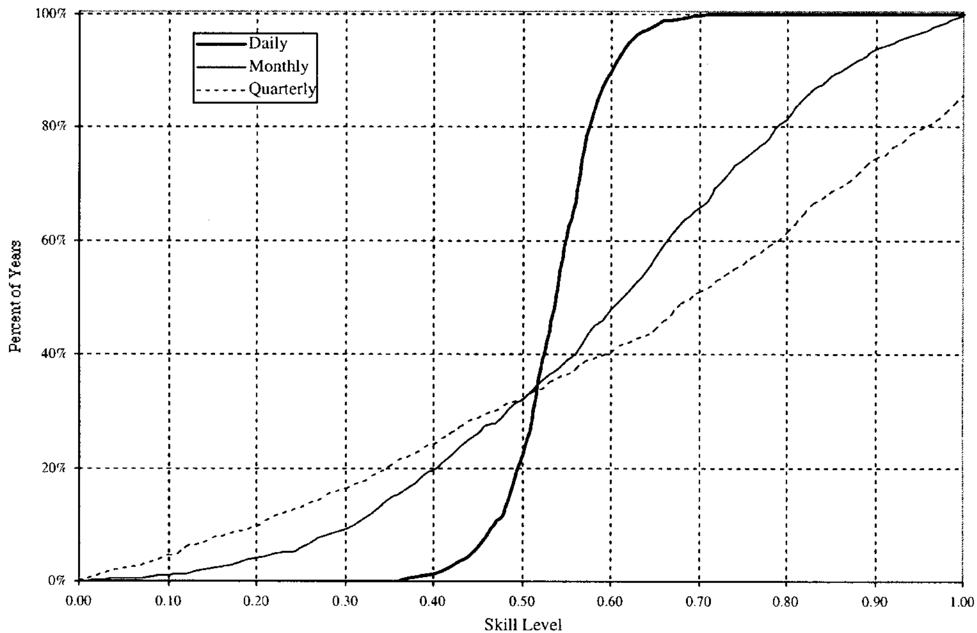


Fig. 2. The Percentage of Years in which the Market Timing Strategy Earn Higher Returns that the Return on the Market Index.

This figure plots the average percentage of years in which the market timing strategy of a given frequency beats the return on the equity index. The skill level represents the probability that the market timer correctly identifies successful market timing opportunities. Note that a perfectly good market timer (skill equal to 1.00) will not beat the equity index return in years in which equity returns are above the money market rate. This follows from the fact that the market timer is invested in equities in those years, earning the same return as the index. This occurs with nonzero, but decreasing, probability for quarterly, monthly and daily trading frequencies, respectively.

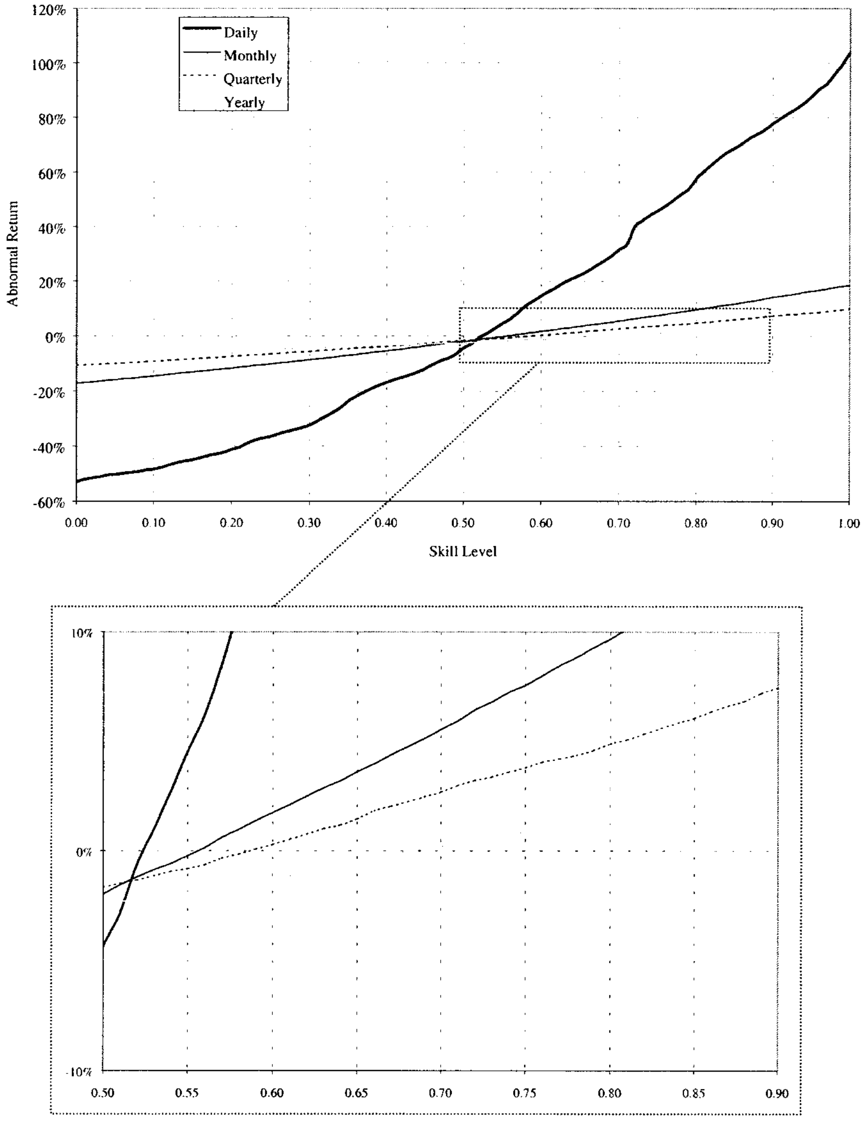


Fig. 3. Abnormal Market Timing Returns Relative to the Capital Market Line For Various Trading Frequencies and Skill Levels.

This figure plots the average abnormal returns (measured as the return from the trading strategy less the return given by the Capital Market Line) from a market timing strategy with the given trading frequency. The skill level represents the probability that the market timer correctly identifies successful market timing opportunities.

the market risk-adjusted abnormal return from market timing strategies over and above the return given by the Capital Market Line.[9] The breakeven skill level as measured against market risk-adjusted returns for daily market timing is below 53%. The breakeven skill for monthly, quarterly, and yearly trading frequencies are about 56%, 61%, and 71%, respectively. While the risk to return tradeoff lowers the breakeven skill level at all trading frequencies, the pattern across frequencies persists. Holding constant the skill level above the breakeven point and considering costless trading, a market timer earns higher risk-adjusted returns by trading more frequently.

These results are strikingly similar to those in Grinold (1989), which show that portfolio performance from an active trading strategy depends on a relationship between breadth and value added. Consistent with Grinold, we find that a trader can get by with less value added (skill) per trade when there is an increase in breadth (trading frequency). One potential puzzle arises with these results. As the market timer moves from yearly timing to daily timing, the required breakeven skill shrinks considerably. However, it is unclear whether predicting market moves over a short time period (e.g. daily) or a long time period (e.g. yearly) is more difficult. It could be that the difficulty in predicting market moves is greater for a daily timers compared with yearly timers.

For a given skill level, it should be clear that the expected profits would generally be decreased (and breakeven skill increased) in the presence of transaction costs, bid-ask spreads, and institutional restrictions on trading. With sufficiently high trading costs, buy-and-hold will always dominate market timing. Sharpe (1975) gives an informal presentation of market timing in the presence of trading costs, while Grant (1978) formalizes a set of optimal timing strategies under various conditions, including trading costs.

3.3. Implications of Market Timing for Portfolio Management

Thus far our discussion has focused on the returns to the market timer. It should be clear that the market timer's abnormal risk-adjusted returns do not occur in a vacuum. The previous analysis has assumed that for individual investors and some institutional traders, open-end mutual funds are likely to be the preferred market-timing vehicle.[10] Thus in this section, we examine how market timers would impact the reported returns of an open-end mutual fund.

Large daily inflows and outflows could impose costs on the mutual fund, which are borne by all shareholders, including those who do not trade. The mutual fund can only convert the trader's cash into a risky (productive) asset at an indirect cost in time (e.g. a delay or lag) and/or a direct cost in dollars (e.g. execution costs as discussed in Edelen (1999)).[11] We focus on the indirect costs, which we consider to be a "dilution effect". The dilution effect results when a trading shareholder obtains an immediate personal benefit of investment in the productive (risky) assets of the fund in exchange for investing her own cash. If fund traders are systematically good market timers, this indirect cost results in a negative impact on the NAV. Conversely, if traders on average exhibit poor timing ability, frequent trading of mutual funds has a positive impact on NAV. Essentially, any gain to the trading shareholder comes at the expense of the non-trading shareholders of the fund. This "zero-sum game" of trading mutual funds may result in peculiar observable phenomena. Specifically, mutual fund returns will reflect both the return to the fund's assets as well as an effect due to market timing.

Given that trading mutual funds is a zero-sum game, any excess return (loss) generated by a market timer becomes the expense (profit) to the remaining mutual fund investors. An example, shown in Fig. 4, illustrates this point. Consider a mutual fund with \$100,000 in assets. The fund holds 95% of the assets in risky securities and 5% in cash. The cash can be used to fund exchanges on behalf of trading shareholders. In this example, we will assume that all net cash flows are zero, except for the trades of a perfectly good market timer. Panel A shows the value of the fund's assets in the absence of the market timer. A key assumption is that the fund manager cannot convert cash into risky assets quickly – new cash can only be invested in the risky assets with several periods (days) lag. Without loss of generality, suppose that we have a perfect market timer who begins the example period by trading 1% of the fund's total assets. Panel B of Fig. 4 shows that the return on the fund's risky assets from period 0 to period 1 is positive, so the timer buys shares of the fund during period 0 (to be holding the fund during a positive return). At the end of the period 1, the timer "cashes in" her shares, resulting in a decreased holding of cash by the fund, shown in Panel C. Since the returns on the fund's risky assets are negative over the next two periods, the timer maintains her cash position. In period 3, the market timer uses all of her cash to purchase shares in the fund in order to capture the positive return to the fund's risky assets during the subsequent period. Notice that a higher proportion of the fund's assets experience a negative return while timer is holding a cash position. The NAV of the fund decreases while the timer holds cash, allowing the timer to buy more shares of the fund than she would have been holding in the absence of trading. As this example illustrates, the timer receives the full benefit of the risky assets' positive returns, despite the fact that the timer has only contributed unproductive assets (i.e. cash) to the fund. Thus, the timer's cash dilutes the positive returns, but all negative returns are undiluted since the timer has previously traded out of the fund. From Panel A we see that the mutual fund

Panel A: Fund's holdings in the absence of market timer

Time Period	End of Period			NAV	Per Period Fund Return	Cumulative Fund Return[1]
	Risky Asset	Cash	Shares Out			
0	$95,000.00	$5,000.00	1,000.00	$100.00		
1	$104,500.00	$5,000.00	1,000.00	$109.50	+9.50%	9.50%
2	$99,275.00	$5,000.00	1,000.00	$104.28	−4.77%	4.28%
3	$89,347.50	$5,000.00	1,000.00	$94.35	−9.52%	−5.65%
4	$98,282.25	$5,000.00	1,000.00	$103.28	+9.47%	3.28%

Panel B: Market timer's holdings

	Beginning of Period			Shares Traded	End of Period			Cumulative Timer Return
Time Period	Cash	Shares	Wealth		Cash	Shares	Wealth	
0	$1,000.00	0.00	$1,000.00	+10.00	$0.00	10.00	$1,000.00	
1	$0.00	10.00	$1,095.00	−10.00	$1,095.00	0.00	$1,095.00	9.50%
2	$1,095.00	0.00	$1,095.00	0.00	$1,095.00	0.00	$1,095.00	9.50%
3	$1,095.00	0.00	$1,095.00	+11.63	$0.00	11.63	$1,095.00	9.50%
4	$0.00	11.63	$1,198.70	−11.63	$1,198.70	0.00	$1,198.70	19.87%

Panel C: Fund's holding with the market timer

	Beginning of Period			End of Period			NAV	Cumulative Fund Return
Time Period	Risky Asset	Cash	Shares Out	Risky Asset	Cash	Shares Out		
0	$95,000.00	$4,000.00	990.00	$95,000.00	$5,000.00	1,000.00	100.00	
1	$104,500.00	$5,000.00	1,000.00	$104,500.00	$3,905.00	990.00	109.50	9.50%
2	$99,275.00	$3,905.00	990.00	$99,275.00	$3,905.00	990.00	104.22	4.22%
3	$89,347.50	$3,905.00	990.00	$89,347.50	$5,000.00	1,001.62	94.19	-5.81%
4	$98,282.25	$5,000.00	1,001.62	$98,282.25	$3,801.30	990.00	103.11	3.11%

[1] Cumulative return is the holding period return earned by purchasing the fund at Time 0 and holding to the end of the given period.

Fig. 4. Example of Mutual Fund Holdings in the Presence of a Perfect Market Timer.

would have returned 3.28% to its shareholders in the absence of a market timer. However, Panel C shows that the return to the mutual fund in the presence of a market timer is only 3.11%.

Mathematically, suppose that a mutual fund earns a return of R_t^{assets} on its combined assets (risky and cash) in the absence of any effect due to market timers. The return to the same fund in the presence of a market timer who trades τ^{timer} proportion of the fund's assets and earns a timing return of R_t^{timer} is given as

$$R_t^{fund} = R_t^{assets} - \tau^{timer}(R_t^{timer} - R_t^{assets}). \tag{3}$$

So, if a timer earns positive excess returns, then the fund's realized (or reported) return on NAV suffers. Alternatively, if the timer is, on average, wrong, then the fund's return is enhanced proportional to the percentage of the fund's assets that the timer trades. Applying Eq. (3) to the example from above, the impact is the timer's proportion of assets ($1,000 initial investment divided by the fund's non-timed assets of $99,000) multiplied by the timer's excess return above the fund's assets' return (19.87% minus 3.28%). This results in a negative impact of 17 basis points on the fund's return in the presence of the market timer (3.11%) compared with the fund's return without the market timer (3.28%).

In Table 2, we report the impact of market timers on mutual fund returns according to Eq. (3) using the previous results of the returns to daily market timers. We show the net effect (positive or negative) on the mutual fund's return for various skill levels and proportions of assets traded by a daily market timer. Since we assume that the market (CRSP value-weighted index) return is the mutual fund return in the absence of market timers, this table is simply illustrative of a potential "real-world" effect. The table shows that the presence of bad timers is a benefit to the mutual fund's return. Conversely, as the timer's skill increases, the effect of the fund's return becomes increasingly negative. The table also shows that the impact is directly related to the proportion of assets traded by the timer.[12]

The above results document the important potential impact on mutual fund returns by skilled market timers. The extant literature provides mixed evidence on mutual fund trader skill. Zheng (1999) finds evidence of "smart money" in mutual funds. That is, net cash flows appear correlated with subsequent mutual fund returns. On the other hand, Nesbitt (1995) suggests that the typical mutual fund investor buys high and sells low. From 1984–1994, he shows that, while the average mutual fund had a mean return of 10.96%, the average mutual fund investor, due to timing errors, had a mean return of only 9.88%. Given the zero-sum impact of market timers (assuming the mutual fund managers did not undertake major changes in their desired cash holding to fund redemptions) mutual funds have an incentive to attract these "bad" market timers. Indeed, if investors are bad market timers on average, funds should encourage frequent trading. Greene and Hodges (2000) report the a simple S&P trading rule that has a skill level of about 65% and find a 0.90% dilution impact for international funds with high fund flow levels. From our results in Table 2, a fund manager could interpret these flows as being consistent with from 2% to 5% of these funds assets are following a daily market timing rule.[13]

This implies that some of the positive returns associated with some funds may be associated with the bad timing ability of the fund investors. This positive impact may be reduced somewhat by the need to hold an increased level of cash or the transaction costs to liquidate equity positions on short

Table 2. Average Impact of Daily Market Timers on Mutual Fund Annual Returns.

This table reports the impact that daily market timers have on mutual fund average annual returns. Skill level is measured as the proportion of time that the market timer correctly predicts the direction of the equity index, relative to an investment in the money market (assumed to give a 4% annualized return). Returns are averaged across 1,000 years, which are simulated by drawing (with replacement) 252 days per year from actual daily market index returns from July 3, 1962 through December 31, 1997. We use the CRSP Value-weighted Index as the mutual fund return, absent market timing. The average mutual fund return is 13.04% in the absence of market timer trading.

		Proportion of beginning assets traded by the market timer					
Timer's Skill	Timer's return	0.0%	0.5%	1.0%	2.0%	5.0%	10.0%
0.00	− 46.24%	0.0%	0.30%	0.59%	1.19%	2.96%	5.93%
0.10	− 41.36%	0.0%	0.27%	0.54%	1.09%	2.72%	5.44%
0.25	− 28.55%	0.0%	0.21%	0.42%	0.83%	2.08%	4.16%
0.50	5.68%	0.0%	0.04%	0.07%	0.15%	0.375	0.74%
0.51	7.37%	0.0%	0.03%	0.06%	0.11%	0.28%	0.57%
0.52	9.70%	0.0%	0.02%	0.03%	0.07%	0.17%	0.33%
0.53	11.39%	0.0%	0.01%	0.02%	0.03%	0.08%	0.17%
0.54	13.29%	0.0%	0.00%	0.00%	0.00%	− 0.01%	− 0.02%
0.55	15.22%	0.0%	− 0.01%	− 0.02%	− 0.04%	− 0.11%	− 0.22%
0.60	26.02%	0.0%	− 0.06%	− 0.13%	− 0.26%	− 0.65%	− 1.30%
0.75	58.33%	0.0%	− 0.23%	− 0.45%	− 0.91%	− 2.26%	− 4.53%
0.90	91.44%	0.0%	− 0.39%	− 0.78%	− 1.57%	− 3.92%	− 7.84%
1.00	118.24%	0.0%	− 0.53%	1.08%	− 2.10%	− 2.26%	− 10.52%

notice in order to accommodate the market timers. We note that other fund types, such as pension funds, hedge funds, and closed-end funds should see no impact from market timers since they are less likely to be preferred market timing vehicles.

4. SUMMARY AND CONCLUSIONS

This paper utilizes simulations using actual daily returns from a market index to explore the how the interaction between market timing skill and trading frequency affects the return distribution from a market timing trading strategy. Consistent with previous studies, we find that a relatively high degree of skill is needed for a yearly market timing strategy to dominate a buy-and-hold indexing strategy. However, we demonstrate that traders need only a modest level of skill in order to beat the market on both a return basis and a market risk-adjusted return basis when employing a *daily* timing strategy. Specifically, in the absence of transaction costs, a trader who is correct about the daily movement in the stock market only 53% to 54% of the time would find it worthwhile to engage in market timing rather than follow a buy-and-hold strategy. In contrast, when trading at less frequent intervals (e.g. monthly or yearly), the skill of the timer must increase in order for a market timing strategy to be profitable.

For a daily timer relying on a signal of the next day's market index return, this signal must have a correlation coefficient of only about 0.08 with the market index indicator in order for the market timing strategy to break even (given a breakeven skill level of 54%). For yearly timers, this signal must increase to a rather precise correlation in excess of 0.60 between the signal and subsequent year's returns (given a breakeven skill level of 83%). This interpretation of the signal as an indicator variable is admittedly simplistic. We leave for future research a more comprehensive treatment of signal strength and the possibility of signals of extreme index movements.

Our results for market timers do not rest on increased risk or a leverage effect. Indeed, we implicitly assume that the market timer invests in the market index, rather than levering up her return (and risk) by identifying a more volatile portfolio. While the results rely on the absence of trading or opportunity costs, we justify this assumption by noting that many mutual fund companies allow unlimited (costless) transfers within their fund families. In particular, most 401(k) and 403(b) retirement plans explicitly allow unlimited transfers or exchanges. The ability to carry out a market timing strategy using tax-deferred retirement plans also minimizes any tax implications (see Hulbert (1992)). To the extent that a trader must pay for their signal, devote costly time to trading on the signal, or pay direct transactions costs, the breakeven skill level will undoubtedly rise. With these costs considered, the increase in the required breakeven skill level is likely to be positively related to trading frequency, leaving our results ambiguous in this case.

The results in this paper also relate to the recent phenomenon of the proliferation of day traders. We show that as the frequency of trading increases, the required skill to breakeven decreases. Day traders make timing decisions essentially every second. With their increased trading frequency, these traders would require an ability that is barely better than chance to predict market (or individual stock) price movements profitably, assuming sufficiently small per trade costs. An alternative interpretation is that traders who possess a large amount of skill can maximize the returns to their skill by trading more frequently. This result is not surprising with the realization that market timing payoffs are similar to the payoffs on a call option. More frequent trading (e.g. daily trading) is similar to a portfolio of options, since the successful market timer can capture the market index movements each day. In contrast, infrequent trading (e.g. yearly trading) is more similar to an option on a portfolio, since the annual market index return is simply a portfolio of the year's daily returns.

This paper offers new evidence on the issue of the propriety of engaging in market timing strategies. Traders who possess rather modest skills in predicting subsequent moves in the equity markets and trade frequently (and cheaply) will find market timing to be a profitable investment strategy. Indeed, with the ability of common individual mutual fund investors to engage in market timing strategies by transferring between cash and risky portfolios on a daily basis, their required skill level to do this profitably is surprisingly low. Finally, our results indicate that successful market timers who use open-end mutual funds are gaining at the expense of the fund's non-timing investors. We show that daily market timing by mutual fund investors can have a substantial impact on fund returns, possibly affecting the evaluation of a mutual fund manager's performance.

NOTES

1. Stale Prices refers to mutual fund Net Asset Value being calculated as of the close of the New York Stock Exchange, typically 4 p.m. Eastern Time. Mutual funds typically use the last market quote when calculating Net Asset Value. With International funds, small cap funds, and high yield bond funds, the mutual fund assets may not have traded for several hours and thus would not reflect market information revealed after the last trade.

2. One strategy alternates between international equity funds and money market accounts using the simple rule, "If the S&P 500 is up, hold open-end mutual funds with large foreign stock holdings, if the S&P 500 is down, hold money market mutual funds".

3. We justify the absence of trading costs in our analysis of market timers by noting that investors daily trading of open-end mutual funds is possible in a large subset of open end mutual funds. Both Chalmers, Edelen, and Kadlec (1999) and Greene and Hodges (2000) find that over half of the mutual funds in their data sets have neither restrictions nor fees associated with fund exchanges. Additionally, Greene and Hodges (2000), examine restrictions and fees contained in prospectuses of international equity funds, and generally find no difference in the level of fund flows of restricted and unrestricted funds. The suggests that while open end mutual funds may have rules in place to restrict active market timing, these flows have little impact on the levels of daily fund flows. They show that costless daily trading of open-end mutual funds is possible and occurs, as evidenced by large daily mutual fund flows.

4. A constant return of 4% on a money market portfolio is assumed for convenience. Our results remain qualitatively the same for different specifications of this rate. In fact, our returns appear slightly understated with a money market rate that co-varies with equity returns.

5. In general, the specification of market timing skill is quite complex. Our implicit (or simplifying) assumption is that the market timer receives a discrete signal of 1 (if market index returns are expected to exceed the money market rate) or 0 (if market index returns are expected to be below the money market rate). This paper does not consider more complex specifications of timing skill.

Alternative specifications include market timers who can only identify "up" markets or those who can only identify "down" markets, each with varying degrees of accuracy. Moreover, some traders may receive "up" and "down" indications that are supplemented with signal strength or precision of signal information. Yet another specification would be a timer who can only identify extreme market index movements (positive and/or negative). While these issues are quite interesting and potentially important, we abstract from them in this paper in order to focus on a transparent measure of skill.

6. The correlation coefficient between the equity market indicator and the signal is a function of the skill level and the probability distributions of the indicator and signal. More precisely, let p be the probability that the equity market indicator is one (i.e. $p = \Pr(r_t = 1)$). Let q be the probability that the signal is one (i.e. $q = \Pr(x_t = 1)$). If we define s as we have in this paper to be the probability of the signal accurately predicting an up market ($r_t = 1$) or a down market ($r_t = 0$) *with the same precision*, *the* q must be given by $q = (1-p)(1-s) + ps$. Finally, the correlation coefficient between the signal and equity market indicators is given by $\rho_{x,r} = \dfrac{ps - pq}{\sqrt{p(1-p)}\sqrt{q(1-q)}}$, yielding an exact relationship between the skill level and the signal-to-market indicator correlation.

This specification of skill level is similar in spirit to the more complex treatment of asset allocation skill level in Clarke, Fitzgerald, Berent, and Statman (1990).

7. By definition, the yearly timer can never beat the market, except when the market is down. For example, with a skill level of 0.70, the yearly timer will beat the market about 70% of years in which the equity index return is below the money market rate. Similarly, the yearly timer's return wil1 *equal* the equity market index's return in 70%

of the years in which the market index return is above the money market rate. This also occurs with quarterly timing in years in which the market index *exceeds* the money market rate in al1 quarters. In these years the market timer's returns cannot exceed the equity index return, but will equal it. For monthly and daily timers, this problem is possible, but much less likely to occur.

8. We note here that market timing strategies may produce return distributions such that standard deviation and beta are not sufficient descriptions of the distribution and risks. However, given the Capital Market Line's prevalent use in practice, we report risk-adjusted (in the CML sense) expected returns as a benchmark for comparison with other portfolio strategies.

9. Recall that the formula for the capital market line is $E(r_p) = r_f + \frac{E(r_M) - r_f}{\sigma_M} \times \sigma_p$. So, the expected return on any portfolio is the money market return plus a risk premium that is proportional to the portfolio's risk (standard deviation). The constant of proportionality is the market index risk premium divided by the market index risk (standard deviation). 10. This assumption is based on two key attributes of open-end mutual funds. The first is that mutual funds offer low- or zero-cost trading and easy access to fund redemptions and purchases. The second is that open-end mutual funds offer exchanges at a single net asset value (NAV) that is determined by the fair market value of the fund's holdings at the end of the day (Investment Company Act of 1940). Thus traders may purchase or sell large amounts of mutual fund shares without concern that an imbalance of supply and demand will adversely affect share price. Therefore, shareholders may have the luxury of trading mutual funds without paying an execution cost or having the risk of adverse price reactions.

11. In their analysis of derivatives use by mutual funds, Koski and Pontiff (1996) examine mutual fund flows in a manner somewhat related to our topic. They point out that cash inflows may decrease the risk exposure of the mutual fund. From this motivation, they explore the use of derivatives as a low-cost (time and dollar) way to address fund flows. While only about 20% of mutual funds use derivatives, they find some evidence that derivatives are used to offset the fund's risk changes due to cash inflows.

12. We intentionally ignore any disruption costs to the mutual fund of rebalancing the portfolio. These costs will deepen good market timers' negative impact. Since we focus on the dilution effect of market timers in this paper, we leave analysis of these costs to future research.

13. We are not suggesting that this is a correct estimate of the percentage of owners following a daily timing rule. It could be that 10–25% of assets following a weekly timing rule, which to the manager would appear as about 1/S this number following a daily rule.

ACKNOWLEDGMENTS

The authors wish to thank Robert Battalio, Peter Eisemann, Alfred Mettler, Husayn Sharur, John Thornton, John Yeoman, members of the Atlanta Society of Financial Analysts, and an anonymous referee for helpful comments. Any remaining errors are the sole responsibility of the authors.

REFERENCES

Benning, C. J. (1997). Prediction Skills of Real-World Market Timers. *Journal of Portfolio Management* (Winter), 55–65.

Bhargava, R., Bose, A., & Dubofsky, D. (1998). Exploiting international stock market correlations with open-end international mutual funds. *Journal of Business Finance & Accounting*, *25*, 765–773.

Bhargava, R., & Dubofsky, D. (2000). A note on fair value pricing of mutual funds. *Journal of Banking and Finance*, forthcoming.

Chua, J. H, Woodward, R. S., & To, W. C. (1987). Potential Gains from Stock Market Timing in Canada. *Financial Analysts Journal*, (September-October), 50–56.

Clarke, R. G., & de Silva, H. (1998). State-Dependent Asset Allocation. *Journal of Portfolio Management*, (Winter), 57–64.

Clarke, R. G., Fitzgerald, M. T., Berent, P., & Statman, M. Required Accuracy for Successful Asset Allocation. *Journal of Portfolio Management*, (Fall), 12–19.

Edelen, R., & Warner, J. (1999). The high-frequency relation between aggregate mutual fund flows and market returns. *Journal of Financial Economics*, forthcoming.

Grant, D. (1978). Market Timing and Portfolio Management. *Journal of Finance*, (September), 1119–1132.

Greene, J. T., & Hodges, C. W. (2000). The Dilution Impact of Daily Fund Flows on Open-end Mutual Funds. Georgia State University Working Paper.

Grinold, R. C. (1989). The Fundamental Law of Active Management. *Journal of Portfolio Management*, (Spring) 30–37.

Hulbert, M. (1992). Market Timing Strategies: Taxes Are a Drag. *Journal of Investing*, (August), 18–20.

Koski, J., & Pontiff, J. (1996). How are Derivatives Used? Evidence from the Mutual Fund Industry. University of Washington working paper.

Larson, G. A., & Wozniak, G. D. (1995). Market Timing Can Work in the Real World. *Journal of Portfolio Management*, (Spring), 74–81.

Nesbitt, S. L. (1995). Buy High, Sell Low, Timing Errors in Mutual Fund Allocations. *Journal of Portfolio Management*, (Fall), 57–60.

Shilling, A. G. (1992). Market Timing: Better than a Buy-and-Hold Strategy. *Financial Analysts Journal*, (March/April), 46–50.

Sharpe, W. F. (1975). Likely Gains from Market Timing. *Financial Analysts Journal*, (March/April), 60–69.

Sy, W. (1990). Market Timing: Is it a Folly? *Journal of Portfolio Management*, 11–16.

Wagner, J., Shellans, S., & Paul, R. (1992). Market Timing Works where it Matters Most . . . in the Real World. *Journal of Portfolio Management*, (Summer), 86–90.

Zheng, L. (1999). Is Money Smart? A Study of Mutual Fund Investors. *Journal of Finance*, forthcoming.

DYNAMIC HEDGE WITH FORECASTING: A MARTINGALE APPROACH

Chin-Shen Lee

ABSTRACT

This paper proposes a dynamic hedging rule in which hedge ratio is ex-ante updated upon the next period spot and futures prices forecasts. It is demonstrated that our hedging-with-forecasting rule could attain a multi-period perfect hedge status. Four commodities (Corn, Soybean, Wheat, and Cotton), four currencies (Pound, Yen, Mark, and Swiss Franc), and two stock indexes (S&P 500 and NYSE) data are employed in the empirical study. The empirical evidences strongly indicate that the proposed hedge-with-forecasting model outperforms the bi-variate GARCH(1,1) model. It is also shown that the proposed hedging-with-forecasting model is a more stable model over the bi-variate GARCH(1,1) model in hedging effectiveness. In addition, our model receives less impact of basis time-varying characteristic on accruing transaction cost of dynamic hedge.

1. INTRODUCTION

Most literatures about hedging are under the criteria of expected utility maximization and variance minimization. Although it is possible to derive a dynamic hedging model that is preference-free and without normality

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 205–229.
2001 by Elsevier Science Ltd.
ISBN: 0-7623-0798-6

assumption (Duffie & Jackson, 1990; Myers & Hanson, 1996), the assumption that futures price are an unbiased predictor of future spot price are imposed. Benninga, Eldor, and Zilcha (1983, 1984) bridged the gap of these two approaches and demonstrated that minimum-variance hedges (MVHs) would be consistent with expected utility maximization hedges under certain conditions including a martingale futures price. The concept of minimum-variance hedge ratio is an important result developed from the portfolio theory approach. Recently, more sophisticated econometric techniques are utilized to account for the characteristics of data in estimating MVH ratios, say for examples, autoregressive conditional heteroscedasticity (Cecchetti, Cumby & Figlewski, 1988; Baillie & Myers, 1991), conditional information (Myers & Thompson, 1989; Viswanath, 1993), convergence of spot and futures prices (Castelino, 1992), and heteroscedasticity (Mathews & Fackler, 1991). However, forecasting technique is rarely employed in hedging strategy development.

The purpose of this paper is to derive a dynamic multi-period hedging model with ex-ante updated hedge ratio using forecasting spot and futures prices of next period. A multi-period hedging strategy, in which the ex-ante hedge ratio is re-estimated on purpose of stabilizing the hedged portfolio values over periods,[1] referring to the next period spot and futures prices forecasts[2] are derived. It is shown that application of the proposed dynamic hedging rule will confine a hedged portfolio value into a martingale process without the presumption of martingale futures prices.[3] In other words, the path of the expected payoff for the hedged portfolio will remains constant over hedging periods. It appears that the basis risk is eliminated and a multi-period perfect hedge status is nearly achieved. Our stabilizing approach is so simple that it can be applied using any of time series data. From a different view, we provide a process to construct a hedged portfolio with martingale payoff. The deviation of the realized hedged portfolio payoff sequence from martingale depends on the ability of forecasting model employed. The discrepancy could be employed to measure the realized hedging performance. However, a selected forecasting model with systematic bias (over or under simultaneous) on spot and futures prices forecasts will not reduce the hedging performance seriously since the relative price difference of spot and futures is employed in our model. The co-movement phenomenon of spot and futures prices has positive effect on hedging performance improvement. The proposed model also benefit from less impact of basis time-varying characteristic on accruing transaction cost of dynamic hedge.

Four commodities (Corn, Soybean, Wheat, and Cotton), four currencies (Pound, Yen, Mark, and Swiss Franc), and two stock indexes (S&P500 and

NYSE) serve as empirical testing data. Three different hedging horizons, 3-month, 6-month, and 1-year, are performed in our empirical test. Section 2 sketches the dynamic hedge-with-forecasting rule. Section 3 gives detail description of empirical data sources. We evaluate the hedging performances of the hedge-with-forecasting model in Section 4. Summary and conclusions are drawn in the last section.

2. HEDGING WITH FORECASTING

In order to develop a dynamic hedge with forecasting, we start with one-period model and generalize the results into multi-period.

2.1. One-period Model

Basis is defined as the difference between spot and futures prices, $B_t = F_t - S_t$. The expected basis change in the period $[t-1, t]$ is expressed as:

$$E(\Delta B_t) = E[(F_t - S_t) - (F_{t-1} - S_{t-1})].$$

Consider a hedged portfolio consists of both spot holdings ($X_{s,t}$) and futures positions ($X_{f,t}$) at time t. Let R_t represent the return of the hedged portfolio during time interval $[t-1, t]$. It follows that the one-period return is a linear combination of the expected price change of spot and futures:

$$E(R_t) = X_{s,t} E(S_t - S_{t-1}) + X_{f,t} E(F_t - F_{t-1}) \tag{1}$$

where S_t and F_t represent the spot price and futures price at time t, respectively. Equation (1) can be rewritten in terms of basis change. Using $h_{t-1} = -X_{f,t}/X_{s,t}$ to represent the proportion of the spot position that is hedged at time t, it follows that

$$\begin{aligned} E(R_t) &= X_{s,t}[E(S_t - S_{t-1}) - h_{t-1} E(F_t - F_{t-1})] \\ &= X_{s,t}[(1 - h_{t-1}) E(\Delta S_t) - h_{t-1} E(\Delta B_t)]. \end{aligned} \tag{2}$$

It is clear that the expected return of the hedged portfolio is composed of the expected basis change and the expected gain or loss from the unhedged spot position.

If the expected change in the basis is zero, the expected gain or loss is reduced as $h_{t-1} \to 1$, according to the interpretation of Ederington (1979). The statement holds for the traditional one-on-one perfect hedge. If the basis is constant over the periods (i.e. spot and futures prices move together perfectly), the spot position could be fully covered by the futures position applying a one-on-one hedge in the one-period hedging. A perfect hedge is seldom attainable because of the realized basis fluctuation. Under minimum-variance framework, a perfect hedge is an extreme case that achieves riskless status. By using the concept of synthetic security, we introduce a new skill to adjust the hedge ratio for period t (h_{t-1}) to manipulate the payoff form of a hedged portfolio. This arrangement would approach the perfect hedge situation in a more realistic manner.

For the one-period static model as described in Eq. (2), h_{t-1} is the selected hedge ratio which is initialed at time t-1 and lifted at time t. Hence, the selected hedge ratio is incorporated with the futures contract to form synthetic futures contract denoted as $F_t^h = h_{t-1} F_t$ and $F_{t-1}^h = h_{t-1} F_{t-1}$. Notice that the size of the synthetic futures contract should be smaller than the original one since the selected hedge ratio is always less than unity. According to the synthetic futures contract, the expected return of a hedged portfolio, as in Eq. (1), could be rewritten as:

$$\begin{aligned} E(R_t) &= X_{s,t} E(S_t - S_{t-1}) + X_{f,t} E(F_t - F_{t-1}) \\ &= X_{s,t} \{E(S_t - S_{t-1}) - h_{t-1} E(F_t - F_{t-1})\} \\ &= X_{s,t} \{E(S_t - S_{t-1}) - E(F_t^h - F_{t-1}^h)\} \\ &= X_{s,t} E(S_t - F_t^h) - (S_{t-1} - F_{t-1}^h)] \\ &= X_{s,t} E(\Delta B_t^h) \end{aligned} \tag{3}$$

where B_t^h is defined as the difference between the synthetic futures and spot prices for time t, i.e. $B_t^h = S_t - F_t^h = S_t - h_{t-1} F_t$; $B_{t-1}^h = S_{t-1} - F_{t-1}^h = S_{t-1} - h_{t-1} F_{t-1}$ and $\Delta B_t^h = B_t^h - B_{t-1}^h$. B_t^h is named the "position basis" to distinguish it from the commonly used basis B_t since it is the difference between the spot position value and hedging futures position value. The change of position basis reflects a hedged portfolio value change over the hedging period. Consequently, the hedge ratio corresponds to the synthetic futures is unity and the expected payoff form of a hedged portfolio is analogous to that of a naive hedge. In addition, the risk of a hedge is solely derived from the uncertainty of the position basis change during the hedging period.

2.2. Multi-period Model

Intuitively, the residual risk of a naive hedge is induced by the relative instability of a spot-futures price relationship over the hedging horizon. It is essentially the basis risk. In a multi-period hedging situation, expected return of a hedged portfolio for each period is simplified to changes of corresponding expected position basis specified in Eq. (3). If we adjust the size of the synthetic futures contract to offset the changes in the spot-futures price relationship for each sub-period, the uncertainty of position basis change is reduced. On the extreme, a perfect hedge might be attained when the position basis approaches flat over several hedging periods.

Consider a hedger who intends to hedge over [0, T]. The hedging horizon is divided into arbitrary N sub-periods. From Eq. (3), the expected return of a hedged portfolio in each sub-period is

$$E(R_t) = X_{s,t} E(\Delta B_t^h) \quad \text{for } t = 1, 2, \ldots, N.$$

with the size of the spot position $X_{s,t}$ fixed over the hedging horizon. The payoff path of a hedged portfolio depends upon the realized position basis in each period. Traditionally, the MVH ratio is selected to reduce the variation of payoff over the hedging horizon. A multi-period perfect hedge is achieved if the hedging rule can be applied so that a sum for the prevailing payoffs along the path equal to zero, i.e. $\sum_{t=1}^{N} R_t = 0$.

Lemma

A dynamic hedging rule creates a sequence of real random variables $\{B_t^h\}$, which is a martingale, such that the total expected payoff of the hedging rule is zero. That is, $E(\sum_{t=1}^{N} R_t) = 0$.

Proof:
Applying proposition 1.2.4 (Lamberton & Lapeyre, 1996)[4] to the dynamic hedging problem, a strategy is designed to render the sequence of associated position basis ($\{B_t^h\}$) as a martingale process, with the corresponding predictable sequence (H_t) defined as $H_t = 1$, for $t = 1, 2, \ldots, N$. According to this setting, the resulted total expected return of the dynamic hedging rule is zero. That is,

$$E\left(\sum_{t=1}^{N} R_t\right) = \mathrm{E}\left(\sum_{t=1}^{N} 1\Delta B_t^h\right) = 0 \qquad (4)^5$$

The forecasting model, such as the Kalman filter model, is introduced into the dynamic hedging rule to create a sequence of position basis satisfying the martingale properties. Based upon the definition of martingale, the sequence of position basis random variables is as follows:

$$E(B_{t+1}^h \mid \Omega_t) = B_t^h, \text{ for t} = 0, 1, \ldots, N-1 \qquad (5)$$

where Ω_t is the information set consisting of past and current prices at time t. Equation (5) could be rewritten as:

$$E(S_{t+1} - h_t F_{t+1} \mid \Omega_t) = S_t - h_{t-1} F_t$$

where $\hat{S}_{t+1}$ and $\hat{F}_{t+1}$ are denoted as the forecasts of one-period-ahead expected spot and futures prices, respectively, i.e. $\hat{S}_{t+1} = \mathrm{E}(\mathrm{S}_{t+1} \mid \Omega_t)$ and $\hat{F}_{t+1} = E(F_{t+1} \mid \Omega_t)$. Therefore, the estimated hedge ratio is:

$$\hat{\mathrm{h}}_{\mathrm{t}} = \frac{\hat{\mathrm{S}}_{\mathrm{t}+1} - S_t + h_{t-1} F_t}{\hat{F}_{t+1}} \; \forall t = 0, 1, \ldots, N-1. \tag{6}$$

This dynamic hedging strategy of adjusting the hedge ratio over periods, as depicted in Eq. (6), creates a martingale sequence of position basis random variables. The resulted total expected return of a hedged portfolio over the hedging horizon is zero as illustrated in Eq. (4). Actually, we provide a process to construct hedged portfolio with martingale payoff. The initial input of Eq. (6), h_0, is estimated by the OLS regression and the recursive relationship of consecutive hedge ratios in Eq. (6) is employed to determine hedge ratio path over the whole hedging horizon.

The hedging performance depends on the sequence of realized position basis deviate from martingale property. Obviously, it is the discrepancy determined by prediction ability of selected forecasting model. However, a selected forecasting model with systematic bias (over or under simultaneous) on spot and futures prices forecasts will not damage hedging performance seriously since position basis reflects the relative price difference of spot and futures. The natural co-movement phenomenon of spot and futures prices could be observed to have positive effect on hedging performance improvement.

2.3. The Kalman Filter Approach

The one-period-ahead spot and futures, $\hat{S}_{t+1}$ and $\hat{F}_{t+1}$, needed to be estimated. To solve the problem, it is assumed that common consensus of investors in the expectation of future prices, which can be described as the following state space time series system exists. It is also assumed that common expectation reflects all available ex-ante information and projects into the following first equation, or the state transition equation:

$$\begin{cases} Z_{t+1} = F Z_t + v_{t+1} \\ Y_t = H' Z_t + w_t \end{cases}$$

$$E(v_t v_\tau) = \begin{cases} Q & \textit{for } t = \tau \\ 0 & \textit{otherwise} \end{cases}$$

$$E(w_t w_\tau) = \begin{cases} R & for\ t = \tau \\ 0 & otherwise \end{cases}$$

where the 2×1 vector $Y_t = (S_t,\ F_t)'$ contains the variables to be forecasted. Under these assumptions, it is implied that all valuable ex-ante information are available in the forecasts from the state space system. The Kalman filter model is employed to generate the forecasts of $Y_t = (S_t,\ F_t)'$ in this work.

The basic Kalman filter model could be expressed as the following innovation representation,

$$\begin{cases} \hat{Z}_{t+1|t} = F\hat{Z}_t + K_t(Y_t - H'\hat{Z}_{t|t-1}) \\ Y_t = H'\hat{Z}_{t|t-1} + e_t \end{cases} \tag{7}$$

with F, H, Q, R matrices as the unknown parameters in the forecasting model. Parameter estimation using the maximum likelihood method is discussed in detail in Appendix. The state vector forecasts, $\hat{Z}_{t+1|t}$, could be acquired through the recursive transition mechanism provided by the Kalman gain matrix K_t. The next period forecasts, $\hat{Y}_{t+1|t} = (\hat{S}_{t+1},\ \hat{F}_{t+1})$, will be obtained through $H'\hat{Z}_{t+1|t}$. Given an initial value of h_{t-1}, the next period hedge ratio $\hat{h}_t$ can be acquired using Eq. (6) if $\hat{S}_{t+1}$ and $\hat{F}_{t+1}$ are estimated. This estimated hedge ratio would fulfill our objective of martingale process of position basis.

3. DATA AND METHODOLOGY

Four commodities (Corn, Soybean, Wheat, and Cotton), four currencies (Pound, Yen, Mark, and Swiss Franc), and two stock indexes (S&P500 and NYSE) are selected to conduct the hedging analysis. Data covers the period spanning from December 1, 1987 through May 31, 1996. Daily closing prices for spot and nearby futures are obtained from the Futures Industry Institute (FII) Data Center. To avoid large price disturbances for maturing contracts, all futures are rolled over to the next nearby contracts one week prior to maturity. Data from the first 2 years are employed for model building, while simulation of hedging is performed using the remaining data of 6 years, from May 1990 to June 1996. Three different hedging horizons, 3-month, 6-month, and 1-year, are used.[6] There are 24 periods in a 3-month hedge, 12 periods in a 6-month hedge and 6 periods in a 1-year hedge during the 6-year out-of-sample period. Hedge ratios are adjusted weekly and the hedging performance of our hedge-with-forecasting model is compared to that of the bi-variate GARCH(1,1) model proposed by Myers (1991) and Baillie and Myers (1991).[7]

Table 1. The Percentage of Risk Reduction* for Corn Market.

Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model	Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model
1-year			**3-month**		
1	68.7249	98.7182	**1**	88.3746	97.1996
2	74.7531	95.1098	**2**	77.4039	79.2052
3	61.7941	98.5324	**3**	59.2477	89.4830
4	81.5567	99.6921	**4**	59.6416	92.6960
5	69.6273	98.1014	**5**	94.2898	94.7043
6	85.9723	99.8669	**6**	47.2792	65.5611
6-month			7	89.6302	95.8195
1	80.7555	98.8660	**8**	88.8326	92.4022
2	35.4527	97.1955	**9**	84.0892	98.2722
3	77.7557	88.1704	**10**	75.9783	96.4722
4	76.0053	95.0449	**11**	59.9062	74.1126
5	80.1312	98.9259	**12**	58.7723	81.1260
6	34.8357	95.5823	**13**	85.3287	97.4398
7	78.9691	99.1089	**14**	84.7447	99.1530
8	88.6176	98.7330	**15**	90.1069	94.4318
9	74.2740	97.6337	**16**	95.0713	98.0681
10	54.9280	98.3309	**17**	89.9786	97.5515
11	87.0246	99.4845	**18**	44.7600	84.4809
12	86.7339	99.6665	**19**	55.0324	84.6798
			20	72.4465	97.5454
			21	82.0222	94.4381
			22	92.5590	99.5309
			23	94.7660	98.2631
			24	85.5605	99.2318

* The percentage risk reduction is denoted as $[\text{var}(\Delta U) - \text{var}(\Delta P)]/\text{var}(\Delta U)$, where $\text{var}(\Delta U)$ is variance of value change without the hedge and is variance of value change with the hedge. As usual, the percentage risk reduction is a measure of hedging effectiveness (Ederington, 1979).

3.1. Bi-variate GARCH(1, 1) Model

The GARCH specification requires the modeling of the first two conditional moments of the bivariate distributions of S_t and F_t. For parsimony and without loss of generality, the GARCH(1,1) model is adopted to account for the time-varying variances and covariances, as applied in Kroner and Sultan (1991, 1993) and Park and Switzer (1995).

To proceed the GARCH(1,1) analysis, the following conditional mean model is constructed

Table 2. The Percentage of Risk Reduction for Soybean Market.

Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model	Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model
1-year			**3-month**		
1	5.8315	94.7489	**1**	36.8748	82.7749
2	21.2645	94.4664	**2**	12.5710	92.6439
3	25.1318	98.2900	**3**	14.9588	91.1541
4	46.1635	97.4712	**4**	10.5095	98.1781
5	18.5017	99.0151	**5**	39.3834	94.9328
6	23.4869	99.6525	**6**	–29.3905	84.9564
6-month			7	46.8164	96.9329
1	13.0568	89.9397	**8**	6.6914	96.8504
2	–11.0715	95.0953	**9**	41.7551	96.9549
3	23.1839	91.5716	**10**	32.9987	96.9172
4	5.5266	98.4688	**11**	7.8152	87.5914
5	30.7868	98.3453	**12**	4.3180	91.4667
6	3.8704	96.1017	**13**	40.3533	97.4568
7	47.7917	96.8928	**14**	59.9579	94.2202
8	45.3188	92.6005	**15**	42.9742	80.4498
9	22.4135	99.0826	**16**	47.0126	95.5252
10	2.0582	96.4763	**17**	16.3981	98.7063
11	–8.7598	97.4872	**18**	53.1702	95.3500
12	44.2281	99.6586	**19**	–19.1995	87.2500
			20	16.8100	94.0049
			21	–27.1403	88.5693
			22	35.5074	97.0303
			23	21.4690	98.9309
			24	50.4719	99.5157

$$\Delta S_t = \alpha + u_t$$

$$\Delta F_t = v_t,$$

with Δ as the first difference operator that implicitly assumes futures price is martingale. Let ε_t represent $(u_t, v_t)'$, the prediction errors have a time-varying covariance matrix

$$\sum_t = E(\varepsilon_t' \varepsilon_t \mid \Omega_{t-1})$$

Table 3. The Percentage of Risk Reduction for Wheat Market.

Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model	Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model
1-year			**3-month**		
1	54.3928	96.0258	**1**	64.0878	95.7710
2	68.5181	99.0667	**2**	40.5123	57.7106
3	22.3553	88.5667	**3**	67.2221	73.7364
4	60.2249	98.7312	**4**	69.1112	74.0072
5	64.8248	98.1167	**5**	78.0929	86.5475
6	54.8337	97.0229	**6**	41.8292	96.1256
6-month			7	84.7256	95.3795
1	53.5232	96.3546	**8**	56.9536	92.6437
2	69.8402	94.9659	**9**	67.8534	98.3356
3	63.8952	96.0446	**10**	56.5224	96.6397
4	72.0889	95.4019	**11**	55.1418	87.0137
5	63.5538	97.6951	**12**	7.9439	80.0402
6	10.6460	81.3502	**13**	58.1056	84.8221
7	55.5236	95.5813	**14**	51.5133	97.4796
8	67.1006	98.0439	**15**	65.3152	94.8894
9	67.3567	98.6751	**16**	68.9293	87.7194
10	62.1253	94.9959	**17**	68.0100	94.5360
11	75.0695	98.6717	**18**	69.6876	83.6640
12	43.5544	91.4187	**19**	60.4625	90.9101
			20	64.3967	65.4206
			21	77.3848	95.8051
			22	67.0126	95.6611
			23	71.2757	68.3166
			24	38.2767	91.1443

Following Bollerslev (1986) and Baillie and Myers (1991), $\sum_t$ can be specified as

$$vech\left(\sum_t\right) = C + Avech\left(\varepsilon_{t-1}\varepsilon'_{t-1}\right) + Bvech\left(\sum_{t-1}\right),$$

with C as a (2×1) vector of parameters, A as a (3×3) matrix of parameters and *vech* as the column stacking operator that stacks the lower triangular portion of a symmetric matrix. This bivariate GARCH(1,1) model allows autocorrelation in the squared prediction errors (i.e. volatility) to be modeled flexibly.

Assuming that the conditional distributions of the prediction errors are normal, the log-likelihood function for a sample of T observations on spot and futures prices is

$$L(\theta) = -T\log 2\Pi - 0.5\sum_{t=1}^{T}\log\left|\sum_t(\theta)\right| - 0.5\sum_{t=1}^{T}\varepsilon_t(\theta)'\sum_t^{-1}(\theta)\varepsilon_t(\theta)$$

where $\theta = \{\alpha, C, A, B\}$ is the set of all conditional mean and variance parameters. The estimate of $\sum_t$ can be obtained from the algorithm of maximum likelihood estimate of θ developed by Berndt, Hall, Hall and

Table 4. The Percentage of Risk Reduction for Cotton Market.

Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model	Hedge Period	Bi-variate GARCH(1,1) Model	Hedging-with-Forecasting Model
1-year			**3-month**		
1	63.8035	99.0716	**1**	36.0639	92.8026
2	60.0821	99.5667	**2**	64.9844	80.2726
3	53.9918	96.6983	**3**	82.0254	98.7509
4	65.3949	99.6125	**4**	84.8713	98.5971
5	84.0480	99.6440	**5**	52.3108	96.9208
6	61.8115	97.2312	**6**	37.9181	98.5471
6-month			7	84.8592	94.5384
1	45.6277	98.4188	**8**	80.9257	92.5029
2	84.1859	99.4796	**9**	58.5884	89.7344
3	47.2456	99.3297	**10**	30.7135	93.1485
4	76.6711	94.5884	**11**	83.9483	94.5063
5	39.0093	96.9279	**12**	71.0024	73.1634
6	78.5580	95.9931	**13**	30.5832	76.1111
7	50.6656	81.9255	**14**	69.6928	81.7004
8	76.3473	99.3252	**15**	89.1691	99.6917
9	77.3824	97.9178	**16**	70.4628	96.5527
10	87.5972	98.9734	**17**	75.8164	98.4899
11	63.2883	96.6870	**18**	79.6188	94.7060
12	64.1010	95.6743	**19**	87.4453	98.3878
			20	90.1099	89.2448
			21	55.5603	95.3845
			22	75.0188	91.0743
			23	94.3828	96.9422
			24	49.4382	91.0712

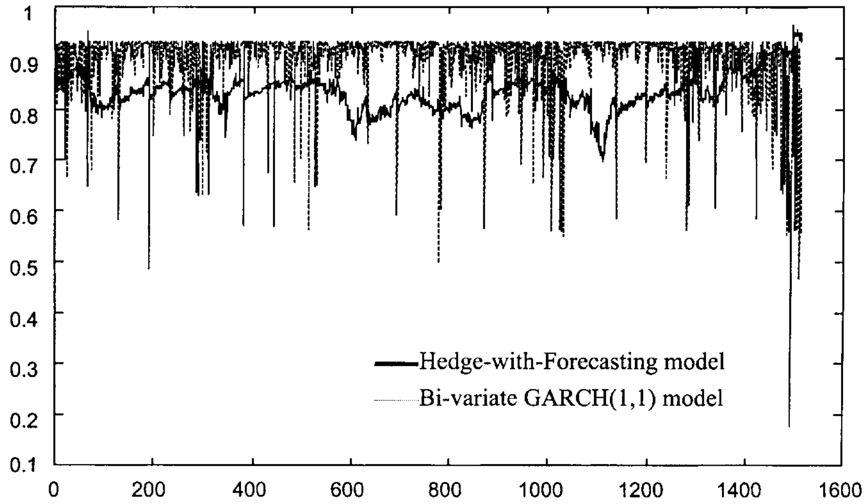


Fig. 1. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Corn Market.

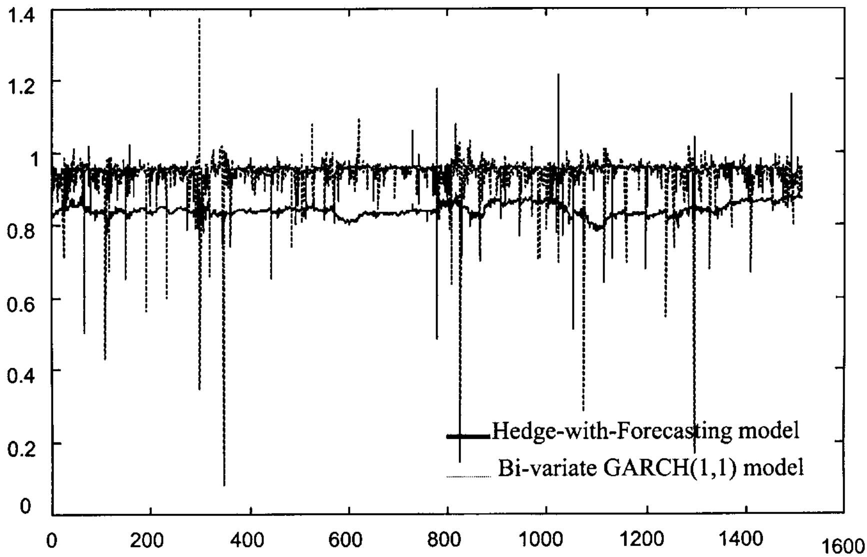


Fig. 2. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Soybean Market.

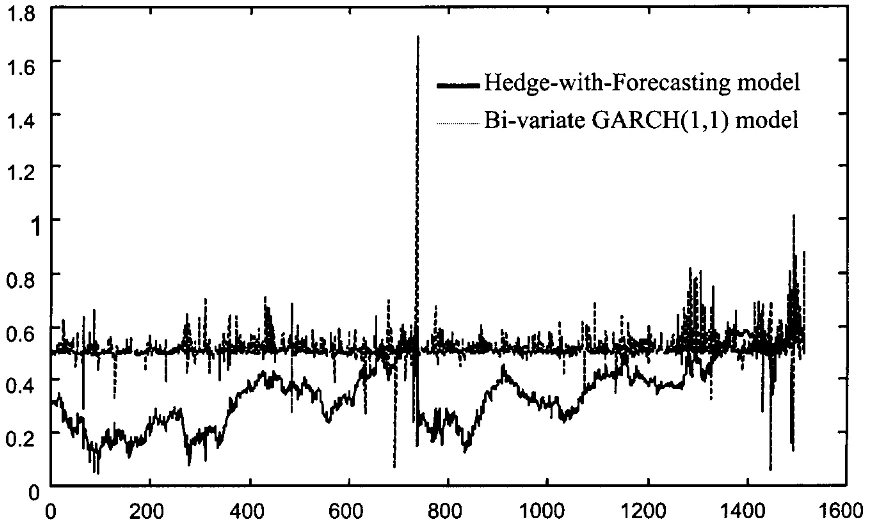


Fig. 3. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Soybean Market.

Hausman (1974). The time-varying optimal hedge ratio estimates computed from GARCH model is

$$\hat{h}_{t-1} = X_{s,t-1} \frac{\hat{\sigma}^{21}}{\hat{\sigma}^{22}}$$

with martingale assumption on the process of futures prices.

4. HEDGING PERFORMANCE

Table 1 reports the period-by-period hedging results for the Corn market using the bi-variate GARCH(1,1) model and our approach. Results suggest that the hedge-with-forecasting strategy significantly outperforms the bi-variate GARCH(1,1) model. For the 1-year hedging period, the bi-variate GARCH(1,1) model reduces approximately 60% to 85% of volatility of uncovered spot positions. The hedge-with-forecasting technique surprisingly reduces 95% to 99% of spot volatility. Since up to 95% of volatility is hedged away, the return series of the hedged portfolio is practically flat. Indeed, the objective of our model is to achieve a very stable return for the hedged portfolio. The evidence suggests that our hedge-with-forecasting hedging model can create a sequence of realized position basis which nearly adheres to martingale properties. The 6-month hedging period is nearly as good as the 1-year hedge period. For every period, the risk reduction of the bi-variate GARCH(1,1) model is less than that of our model, and the new strategy eliminates more than 95% of spot risk in every but one period.

The performance of the new hedging approach for the 3-month hedge period is less impressive than longer period hedges. It shows that the co-movement phenomenon of spot and futures prices is more significant in long-term period and/or the forecasting ability of the Kalman filter model increased with larger data set. Nevertheless, it still outperforms the bi-variate GARCH(1,1) model in every hedging period. In the worst case, our hedging strategy reduces 65% of cash volatility, while the bi-variate GARCH(1,1) model leaves more than a 40% risk exposure to hedgers in 5 cases.

Hedging performances for the Soybean, Wheat, and Cotton markets reported in Tables 2, 3, and 4, respectively, are analogous to that of the Corn market. Our hedging strategy outperforms the bi-variate GARCH(1,1) model in every case. Risk reduction made by the hedge-with-forecasting approach is often above

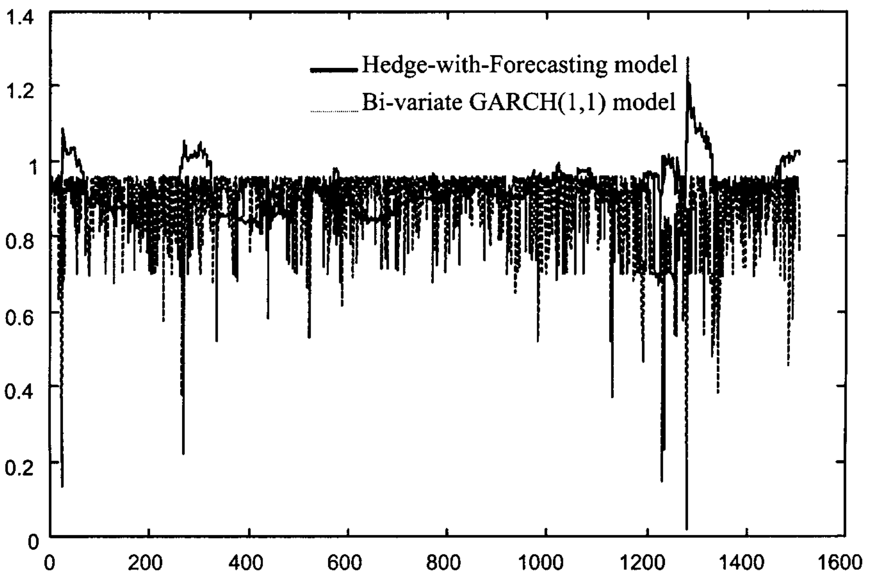


Fig. 4. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Cotton Market.

Table 5. Summary of Statistics for Percentage Risk Reduction Using Bi-variate GARCH(1,1) and Hedge-with-Forecasting Models for Commodity Hedge.

Hedging-Model	Statistics	Percentage Risk Reduction 1-year	6-month	3-month
		Corn		
Bi-variate GARCH(1,1) Model	Average	73.74	71.29	77.33
	Maximum	85.97	88.62	95.07
	Minimum	61.79	34.84	44.76
	Standard Deviation	8.91	18.99	16.01
Hedge-with-Forecasting Model	Average	98.34	97.23	91.74
	Maximum	99.87	99.67	99.53
	Minimum	95.11	88.17	65.56
	Standard Deviation	1.72	3.21	9.00
		Soybean		
Bi-variate GARCH(1,1) Model	Average	23.40	18.20	23.46
	Maximum	46.16	47.79	59.96
	Minimum	5.83	–11.07	–29.39
	Standard Deviation	13.10	20.73	24.88
Hedge-with-Forecasting Model	Average	97.27	95.98	93.27
	Maximum	99.65	99.66	99.52
	Minimum	94.47	89.94	80.45
	Standard Deviation	2.19	3.11	5.38
		Wheat		
Bi-variate GARCH(1,1) Model	Average	54.19	58.69	60.43
	Maximum	68.52	75.07	84.73
	Minimum	22.36	10.65	7.94
	Standard Deviation	16.55	17.46	16.05
Hedge-with-Forecasting Model	Average	96.26	94.93	86.85
	Maximum	99.07	98.68	98.34
	Minimum	88.57	81.35	57.71
	Standard Deviation	3.93	4.73	11.38
		Cotton		
Bi-variate GARCH(1,1) Model	Average	64.86	65.89	68.15
	Maximum	84.05	87.60	94.38
	Minimum	53.99	39.01	30.58
	Standard Deviation	10.20	16.64	19.80
Hedge-with-Forecasting Model	Average	98.64	96.27	92.20
	Maximum	99.64	99.48	99.69
	Minimum	96.70	81.93	73.16
	Standard Deviation	1.32	4.80	7.35

Table 6. Summary of Statistics for Percentage Risk Reduction Using Bi-variate GARCH(1,1) and Hedge-with-Forecasting Models for Currency Hedge.

Hedging-Model	Statistics	Percentage Risk Reduction 1-year	6-month	3-month
		Sterling Pound		
Bi-variate GARCH(1,1) Model	Average	67.91	67.03	65.80
	Maximum	80.45	80.54	86.09
	Minimum	59.06	47.55	22.86
	Standard Deviation	7.92	11.32	18.90
Hedge-with-Forecasting Model	Average	97.64	96.40	94.04
	Maximum	99.86	99.77	99.66
	Minimum	93.04	90.70	81.91
	Standard Deviation	2.53	2.92	4.71
		Japanese Yen		
Bi-variate GARCH(1,1) Model	Average	68.71	68.28	70.28
	Maximum	77.76	79.93	85.44
	Minimum	57.67	30.86	17.57
	Standard Deviation	8.03	14.08	15.09
Hedge-with-Forecasting Model	Average	99.17	97.61	94.38
	Maximum	99.67	99.75	99.58
	Minimum	97.96	90.52	84.86
	Standard Deviation	0.69	2.68	3.93
		Deutsche Mark		
Bi-variate GARCH(1,1) Model	Average	71.06	70.03	69.19
	Maximum	75.19	76.51	89.22
	Minimum	68.00	52.23	41.51
	Standard Deviation	2.60	7.53	12.88
Hedge-with-Forecasting Model	Average	99.07	98.30	96.54
	Maximum	99.50	99.60	99.50
	Minimum	98.31	96.49	91.34
	Standard Deviation	0.43	0.89	2.04
		Swiss Franc		
Bi-variate GARCH(1,1) Model	Average	69.50	69.15	68.94
	Maximum	76.18	80.31	86.03
	Minimum	52.54	47.25	22.87
	Standard Deviation	9.03	9.98	14.59
Hedge-with-Forecasting Model	Average	98.79	98.22	95.11
	Maximum	99.44	99.27	99.39
	Minimum	97.73	95.61	75.01
	Standard Deviation	0.62	0.98	5.22

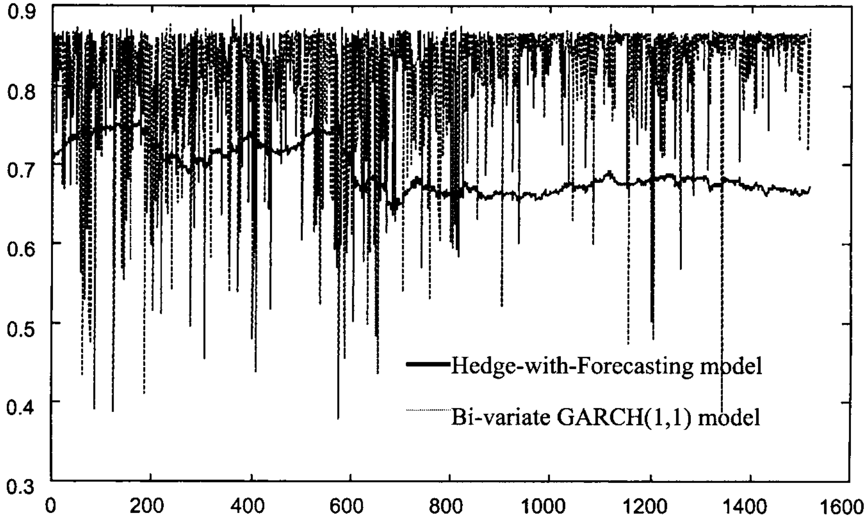


Fig. 5. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Pound Market.

90%, leaving little return volatility of the hedged portfolio. This performance reinforces the superiority of our model.

Table 5 provides a summary of statistics for the percentage risk reduction on the four commodity markets applied by the two models. It is shown that the average risk reduction is always larger for the hedge-with-forecasting approach under every hedging horizon. Moreover, the smaller standard deviation of hedging performance indicates that our model is more stable than the bi-variate GARCH(1,1) model. We also find that the average performance of the hedge-with-forecasting model improves as the hedge period lengthens. For example, in the Corn market, a 91.74% risk is eliminated in the 3-month hedge, while 98% is eliminated in the 1-year hedge. The above evidences demonstrate that our hedge-with-forecasting model is more suitable than bi-variate GARCH(1,1) model to incorporate the co-movement phenomenon of spot and futures prices to improve the hedging performance. In addition, the hedging performance of our model is increased with larger data set.

Figures 1 to 4 illustrate the paths of the adjusted hedge ratios over the 6-year out-of-sample period for the four commodity markets as suggested by the two models. The series of bi-variate GARCH(1,1) generated hedge ratios show large fluctuation overtime, which imply the need for frequent and large scale of

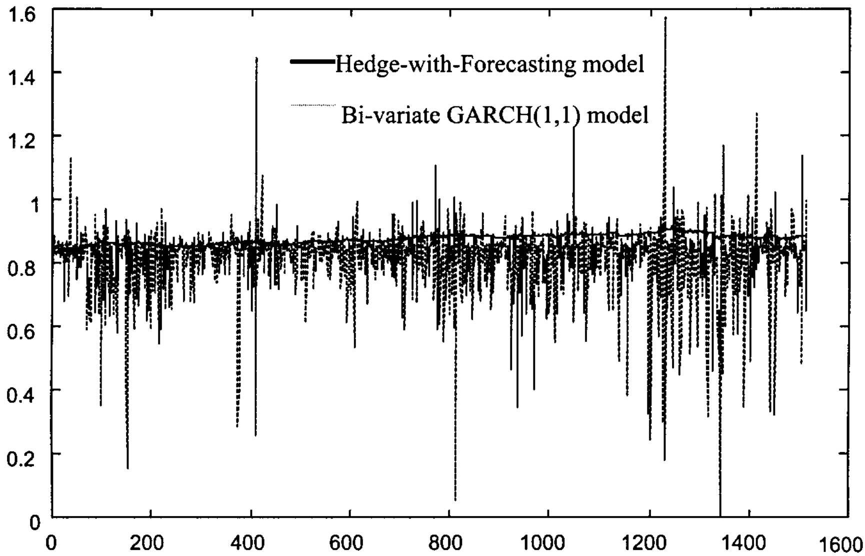


Fig. 6. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Yen Market.

portfolio rebalance. According to the bi-variate GARCH(1,1) results, hedgers must frequently buy and sell relatively large amounts of futures contracts in order to follow the hedge ratios suggested by the bi-variate GARCH(1,1) model. The resulting nontrivial transaction costs impose a serious disadvantage for implementing the bi-variate GARCH(1,1) model in practice. Many authors have questioned the cost-effectiveness of the bi-variate GARCH(1,1) model over the simple OLS static hedge. The hedge-with-forecasting model, on the other hand, suggests very little period-by-period changes of hedge ratio, therefore, requires only minor adjustments of a hedged portfolio. The stability of hedge ratios implies that significant transaction costs are saved by implementing our hedge strategy. The consecutive hedge ratio is the same while spot price change equals to futures price change during the period.[8] Therefore, both basis convergence and basis time-varying characteristics increase the transaction cost of hedging models. In comparison with bi-variate GARCH(1,1) model, our model receives less impact of basis time-varying characteristics on accruing transaction cost of dynamic hedge ratio adjustment.

Hedging performance using currency futures on the Pound, Yen, Mark, and Swiss Franc are also tested, and the summary of statistics are reported in

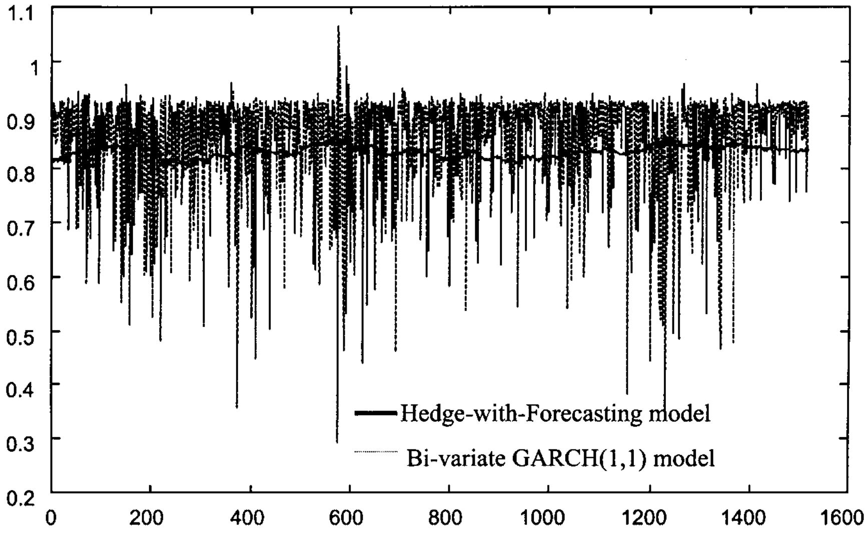


Fig. 7. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Mark Market.

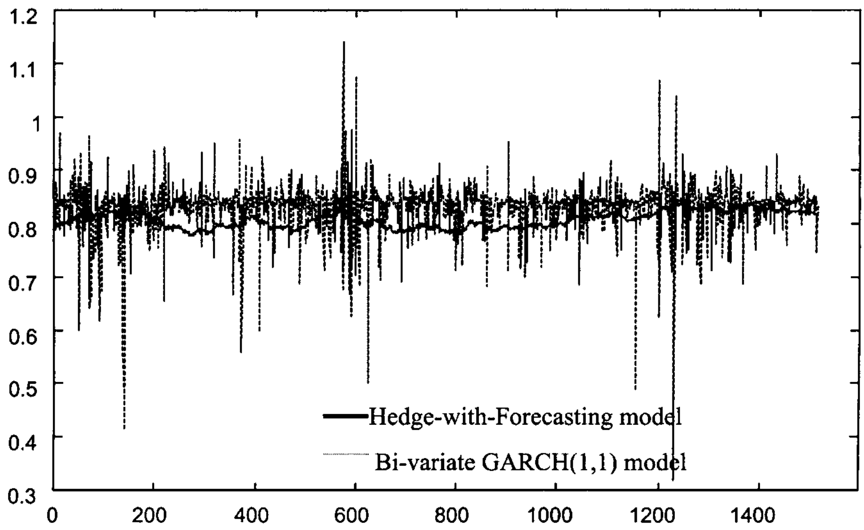


Fig. 8. Hedge Ratio Paths of Hedge-with-Forecasting and Bi-variate GARCH(1,1) Models for Swiss Franc Market.

Table 6. Results closely resemble the agriculture markets. The hedge-with-forecasting approach consistently exhibits over 94% of risk reduction, which is a large improvement when compared to the average 60% to 70% risk reduction by the bi-variate GARCH(1,1) model. Plots of the hedge ratios in Figs 5 to 8 also demonstrate that only minor adjustments are needed to achieve a substantial risk reduction.

Table 7 summarizes the hedging performance on the S&P500 and NYSE stock indexes. The bi-variate GARCH(1,1) model produces better and more consistent hedging results than in other markets. It is able to hedge over 90% of spot volatility, and the variation of performance is smaller than previous results. However, the proposed hedge-with-forecasting model still outperforms the bi-variate GARCH(1,1) model in every incidence. For the 1-year hedging of the S&P500 index, the hedge-with-forecasting approach hedges away more

Table 7. Summary of Statistics Percentage Risk Reduction Using Bi-variate GARCH(1,1) and Hedge-with-Forecasting Models for Index Hedge.

Hedging-Model	Statistics	Percentage Risk Reduction		
		1-year	6-month	3-month
		S&P 500		
Bi-variate GARCH(1,1) Model	Average	90.76	90.60	91.42
	Maximum	92.10	93.21	94.66
	Minimum	89.47	86.60	84.69
	Standard Deviation	0.84	1.81	2.25
Hedge-with-Forecasting Model	Average	99.55	98.66	97.08
	Maximum	99.87	99.83	99.67
	Minimum	99.07	96.75	91.27
	Standard Deviation	0.28	0.98	2.33
		NYSE Index		
Bi-variate GARCH(1,1) Model	Average	89.67	89.79	90.48
	Maximum	92.38	92.58	93.88
	Minimum	84.62	85.13	81.58
	Standard Deviation	2.73	2.69	3.11
Hedge-with-Forecasting Model	Average	99.50	98.64	96.81
	Maximum	99.83	99.83	99.66
	Minimum	98.67	96.93	86.94
	Standard Deviation	0.44	1.05	2.89

than 99% of spot risk, while the bi-variate GARCH(1,1) leaves approximately 10% of uncovered risk. For shorter period hedges such as the 3-month hedge, an average of greater than 97% of volatility is eliminated by the hedge-with-forecasting model. Again the dramatic risk reduction provides evidence for the effectiveness of hedging using the hedge-with-forecasting model. It is a common observation that basis time-varying variation is less significant to finical futures since the seasonal demands and supplies of agriculture commodity often augment basis variation. Therefore, both models provide better hedging performance than previous results.

5. CONCLUSION

A hedge-with-forecasting technique which utilizes spot and futures prices forecasts to improve the risk reduction ability of hedge is proposed in this paper. Forecasts of next-period spot and futures prices by the Kalman filter are employed to derive information for basis change. The hedged portfolio returns are stabilized by adjusting the hedge ratio to achieve a martingale value process of hedged portfolio.

The agriculture, currencies, and stock indexes markets are tested using both the proposed hedge-with-forecasting model and bi-variate GARCH(1,1) model. Empirical evidences suggest significant improvement of hedging effectiveness relative to the bi-variate GARCH(1,1) results. Also, the smaller variance of the percentage risk reduction suggests a more reliable performance attained by the hedge-with-forecasting than by the bi-variate GARCH(1,1) model. Finally, unlike the bi-variate GARCH(1,1) model whose hedge ratio is volatile over time, the hedge-with-forecasting approach produces a very smooth time path of hedge ratio. This means that only a small amount of futures trading is required to adjust the dynamic hedge ratio, and the reduction of transaction costs is just as significant as the reduction of spot risk. These empirical evidences demonstrate that our model can extract available information from the price forecasts of spot and futures to improve hedging performance. The co-movement phenomenon of spot and futures prices has positive effect on hedging performance improvement in our model. In contrast to bi-variate GARCH(1,1) model, our model receives less impact of basis time-varying characteristic on accruing transaction cost of dynamic hedge.

NOTES

1. To stabilize the hedged portfolio values over periods is an alternative way to achieve the status of minimum-variance of hedged portfolio value. To the extreme, a perfect hedge fixed the value of hedged portfolio over periods.

2. The Kalman Filter technique is employed in this paper.

3. Futures price is a martingale, in other words, futures price is an unbiased predictor of future spot price. This assumption has been made in numerous articles about hedge. However, Kolb (1997) remarks that there are debates on the martingale property of futures price and it is inclusive to date.

4. Proposition 1.2.4 (Martingale Transform Lemma) indicates that an adapted sequence of real random variables (M_n) is a martingale if and only if for any predictable sequence (H_n), we have: $E\left(\sum_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{H_n}\Delta M_n\right)=0$.

5. This is an immediately result from Proposition 1.2.4 (Lamberton & Lapeyre, 1996), that is the predictable sequence (H_n) defined as $H_t = 1$, for $t = 1, 2, \ldots, N$., and the random variables (M_n) defined as (B_t^h).

6. The hedging period may not match the life of futures contract. In our empirical hedging test, the hedge ratio is adjusted weekly over the contract life according the dynamic hedging rule derived in Section 2.

7. Both articles provide evidences that time-varying optimal hedge ratio estimates computed from the GARCH model perform marginal better than the constant estimates obtained using conventional regression techniques. Henceforth, we take the GARCH model as the benchmark for performance.

8. From Eq. (6), $h_t = h_{t-1}$ if $S_{t+1} - S_t = F_{t+1} - F_t$.

REFERENCES

Baillie, R. T., & Myers, R. J. (1991). Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge. *Journal of Applied Econometrics*, *6*, 109–124.

Benninga, S., Eldor, R., & Zilcha, I. (1983). Optimal Hedging in the Futures Market under Price Uncertainty. *Economic Letters*, *13*, 141–145.

Benninga, S., Eldor, R., & Zilcha, I. (1984). The Optimal Hedge Ratio in Unbiased Futures Markets. *The Journal of Futures Markets*, *4*, 155–159.

Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, J. A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economics and Social Measurement*, *4*, 653–665.

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

Burmeister, E., & Wall, K. D. (1986). Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to the German Hyperinflation. *Journal of Econometrics*, *20*, 255–284.

Castelino, M. G. (1992). Hedge Effectiveness: Basis Risk and Minimum-Variance Hedging. *American Journal of Agricultural Economic*, *12*, 187–201.

Cecchetti, S. G., Cumby, R. E., & Figlewski, S. (1988). Estimation of the Optimal Futures Hedge. *Review of Economics and Statistics*, *70*, 623–630.

Duffie, D., & Jackson, M. O. (1990). Optimal Hedging and Equilibrium in a Dynamic Futures Market. *Journal of Economic Dynamics and Control*, *14*, 21–33.

Ederington, L. (1979). The Hedging Performance of the New Futures Markets. *Journal of Finance*, 157–170.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Kolb, R. W. (1997). *Understanding Futures Markets*. Blackwell Publishers Inc.

Kroner, K., & Sultan, J. (1993). Time-varying Distributions and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial & Quantitative Analysis*, *28*, 535–551.
Lamberton, D., & Lapeyre, B. (1996). *Introduction to Stochastic Calculus: Applied to Finance*. NJ: Chapman & Hall.
Mathews, K. H., & Fackler, P. L. (1991). Hedge Ratios in the Presence of Heteroskedasticity. *Paper presented at AAEA annual meeting*, August, 4–7.
Myers, R. J. (1991). Estimating Time-Varying Optimal Hedge Ratios on Futures Markets. *Journal of Futures Markets*, *11*(1), 39–53.
Myers, R. J., & Thompson, S. R. (1989). Generalized Optimal Hedge Ratio Estimation. *American Journal of Agricultural Economics*, *71*, 858–868.
Myers, R. J., & Hanson, S. D. (1996). Optimal Dynamic Hedging in Unbiased Futures Markets. *American Journal of Agricultural Economics*, *78*, 13–20.
Park, T. H., & Switzer, L. N. (1995). Bivariate GARCH Estimation of the Optimal Hedge Ratios for Stock Index Futures: A Note. *Journal of Futures Markets*, *15*(1), 61–67.
Viswanath, P. V. (1993). Efficient Use of Information, Convergence Adjustments, and Regression Estimates of Hedge Ratios. *The Journal of Futures Markets*, *13*, 43–53.

APPENDIX

The Kalman filter model is an efficient forecasting model employed in the state space time series study. The state space representation of the Kalman filter model adopted in this research is as follows:

$$\begin{cases} Z_{t+1} = FZ_t + v_{t+1} \\ Y_t = H'Z_t + w_t \end{cases}$$

$$E(v_t v_\tau) = \begin{cases} \mathrm{Q} & \text{for } \mathrm{t} = \tau \\ 0 & \textit{otherwise} \end{cases}$$

$$E(w_t w_\tau) = \begin{cases} \mathrm{R} & \text{for } \mathrm{t} = \tau \\ 0 & \textit{otherwise} \end{cases}$$

where $Z_t(4 \times 1)$ denotes the state vector with the parameters of the VARMA(1,1) of S_t and F_t as its elements, $Y_t(2 \times 1) = (F_t,\ F_t)'$ denotes the observed vector of spot and future prices, $F(4 \times 4)$ is the transition matrix of, and $H(2 \times 4)$ is the parameter matrix of the observation equation. Y_{t+1} is the intended next period price vector. We use an innovation model to represent the above system as follows:

$$\begin{cases} \hat{Z}_{t+1|t} = F\hat{Z}_{t|t-1} + K_t(Y_t - H'\hat{Z}_{t|t-1}) \\ Y_t = H'\hat{Z}_{t|t-1} + e_t \end{cases}$$

Let Y_t be the matrix for Y_t through Y_t, i.e. $Y_t = (Y'_t, Y'_{t-1}, \ldots, Y'_1)$ the forecast for the conditional mean of state vector is represented as

$$\hat{Z}_{t+1|t} \equiv \hat{E}(Z_{t+1}|Y_t).$$

The forecasts of state vectors of each consecutive periods, say $\hat{Z}_{1|0}$, $\hat{Z}_{2|1}$, $\hat{Z}_{3|2}, \ldots, \hat{Z}_{t|t-1}$, can be obtained through the recursive algorithm of the transition equations. Consequently, Y_{t+1} could be forcasted effectively. The forcast error matrix $P_{t+1|t}$ (4×4) for Z_{t+1} could be defined as:

$$P_{t+1|t} \equiv E[(Z_{t+1} - \hat{Z}_{t+1|t})(Z_{t+1} - \hat{Z}_{t+1|t})'].$$

We introduce the maximum likelihood estimation for the unknown parameters F, H, Q, R in the Kalman filter model below. According to the theory of Kalman filter forecasting model (Hamilton (1994)), $\hat{Z}_{t|t-1}$ and $\hat{Y}_{t|t-1}$ forecasts are optimal in any functions of $Y_{t-1} = (Y'_{t-1}, Y'_{t-2}, \ldots, Y'_1)'$, if the initial state Z_1 and innovations $\{w_t, v_t\}_{t=1}^{T}$ obey multivariate Gaussian distribution. Meanwhile, if Z_1 and $\{w_t, v_t\}_{t=1}^{T}$ obey multivariate Gaussian distribution, Y_t is also Gaussian under T_{t-1} with mean and $H'Z_{t|t-1}$ variance $H'P_{t|t-1}H + R$, i.e.

$$Y_t \mid \mathrm{Y}_{t-1} \sim N[(H'\hat{Z}_{t|t-1}),(H'P_{t|t-1}H + R)],$$

where $P_{t|t-1} \equiv E[(z_t - \hat{Z}_{t|t-1})(Z_t - \hat{Z}_{t|t-1})']$ is the mean square error matrix of forecasts for $\hat{Z}_{t|t-1}$. The probability distribution of $Y_t \mid Y_{t-1}$ is

$$f_{Y_t|Y_{t-1}}(Y_t \mid Y_{t-1})$$

$$= (2\pi)\frac{-n}{2} \left| H'P_{t|t-1}H + R \right|^{\frac{-1}{2}}$$

$$\times \exp\left[\frac{-1}{2}(Y_t - H'\hat{Z}_{t|t-1})'(H'P_{t|t-1}H + R)^{-1}(Y_t - H'\hat{Z}_{t|t-1}) \right]$$

for $t = 1,2, \ldots, T$.

The sample loglikelihood function could be easily constructed through above probability density function as

$$\sum_{t=1}^{T} \log f_{Y_t|(X_t,\, Y_{t-1})}(Y_t \mid (X_t, Y_{t-1})$$

The maximum likelihood estimation of F, H, Q, R in the Kalman filter forecasting model could be acquired through maximizing above loglikelihood function. The maximization algorithm can be performed by numerical optimization procedure provided by Burmeister and Wall (1982). The MATLAB scientific calculation package is used to solve our optimization problem.

The Kalman filter forecasts for $Y_{t+1} = (S_{t+1}, F_{t+1})'$ is obtained through the following procedure:

Step 1

Calculate the initial values for $Z_{1|0} = E(Z_1)$ and $P_{1|0} = E[(Z_1 - E(Z_1))(Z_1 - E(Z_1))']$.

According to Hamilton (1994), the following initial values could be used

$$\hat{Z}_{1|0} = 0,$$

$$vec(P_{1|0}) = [I - (F \otimes F)^{-1} \cdot vec(Q).$$

Step 2

Start from $\hat{Z}_{1|0}$, the intended forecasts of $Z_{t+1|t}$ could be acquired by way of the following recursive relationship

$$Z_{t+1|t} = FZ_{t|t-1} + K_t(Y_t - H'Z_{t|t-1})$$

$$P_{t|t-1} = F[P_{t-1|t-2} - P_{t-1|t-2}H(H'P_{t-1|t-2}H + R)^{-1}H'P_{t-1|t-2}]F' + Q$$

$$= (F - K_tH')P_{t-1|t-2}(F' - HK_t') + K_tRK_t' + Q$$

where $K_t = FP_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}$ is the Kalman gain. The Kalman gain K_t is the upmost important character for it provides the dynamic correction mechanism for the state vector.

Step 3

Using $\hat{Z}_{t+1|t}$ the forecasts of $Y_{t+1} = (S_{t+1}, F_{t+1})'$ could be easily obtained

$$\hat{Y}_{t+1}|_t \equiv \hat{E}(Y_{t+1} \mid Y_t) = H'\hat{Z}_{t+1|t}.$$

The mean square error of forecasts for $\hat{Y}_{t+1}|t$ is

$$MSE = E[(Y_{t+1} - \hat{Y}_{t+1|t})(Y_{t+1} - \hat{Y}_{t+1|t})'] = H'P_{t+1|t}H + R.$$

This Page Intentionally Left Blank

MEASURING THE INTEREST RATE RISK OF BONDS WITH EMBEDDED OPTIONS

Steven V. Mann and Pradipkumar Ramanlal

ABSTRACT

We present a method to measure the duration and convexity of bonds with embedded options that accounts for representative contractual features and realistic movements in the yield curve. The method unifies two aspects of the literature that have evolved almost independently, namely, specification of the interest-rate process and characterization of how the yield curve changes shape. We take the view that these two aspects are intimately linked and must be considered in tandem when evaluating interest rate exposure. Finally, we show that significant errors in risk assessment can result if either the bond's optionality or anticipated changes in the yield curve's shape are ignored.

1. INTRODUCTION

Modified duration and convexity measure the sensitivity of a bond's price to a given change in yield. Traditional formulations of these measures are derived under the assumption that the yield curve is flat and moves in parallel shifts. In other words, all bond yields are the same regardless of maturity and these yields are perfectly correlated. These duration/convexity measures also presuppose that a shift in the yield curve leaves the bond's cash flows unchanged (i.e. the bond does not have an embedded option). The purpose of

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 231–254.

ISBN: 0-7623-0798-6

this study is to show how duration/convexity can be measured for bonds with embedded options while allowing for realistic changes in the yield curve's shape. Option-embedded bonds with realistic contract provisions are considered.

Measuring the interest rate risk of a bond with an embedded option is complicated for two reasons. First, the bond's cash flows depend on the path that future interest rates take. Second, interest rates in turn follow a complex process which manifests in the different ways the yield curve changes shape. Two strands of research have developed concurrently that speak to these issues. One strand characterizes the interest rate process while the other focuses on how the yield curve changes shape. In developing duration and convexity measures for bonds with embedded options, we merge these two strands of research.

Models of the interest rate process fall into two broad categories: equilibrium models and no-arbitrage models. Equilibrium models (see Vasicek, 1977; Rendleman & Bartter, 1980; Brennan & Schwartz, 1982; Cox, Ingersoll & Ross, 1985; Longstaff & Schwartz, 1992) use information about the economy (e.g. the short-term interest rate, the long-term interest rate, production, inflation rate, etc.) to explain the yield curve's behavior. The intent of these models is to determine the fair value of Treasury bonds given a set of economic state variables. Conversely, no-arbitrage models (see Ho & Lee, 1986; Hull & White, 1990, 1993, 1994; Black, Derman & Toy, 1990; Black & Karasinski, 1991; Heath, Jarrow & Morton, 1992; Amin & Morton, 1994) are useful in pricing bonds with embedded options. These models take Treasury bond prices as given (without questioning whether or not they are correctly priced) and price the embedded option relative to the observed bond prices. These models are calibrated to replicate the observed U.S. Treasury spot curve. More complicated versions of these models are calibrated to the volatility curve and cap curve in addition to the spot curve.

Another strand of literature has focused on how the yield curve changes shape. As noted above, traditional duration/convexity measures assume the yield curve is flat and yield changes result in parallel shifts. Several recent studies address this inadequacy and develop interest rate risk measures that allow for more realistic changes in the yield curve's shape. To accomplish this task, it is necessary to characterize interest rate movements by factors and consider correlations among these factors if they are not orthogonal. Several studies proceed along these lines, for example: Bierwag, Kaufman and Latta (1987) consider a two-factor model for yield curve shifts; Ho (1992) identifies key rates that characterize the yield curve's shape; Willner (1996) suggests that taking the yield curve's level, slope and curvature as the appropriate factors;

and Mann and Ramanlal (1998) derive explicit expressions for duration and convexity of straight bonds consistent with observed changes of the yield curve's shape (i.e. level, slope and curvature). In each case, factor sensitivities (i.e. the impact on the bond's price given changes in specific factors that characterize the yield curve's shape) can be estimated and interpreted as partial durations used to measure interest rate risk exposure (see also, Reitano, 1990).

While these two strands of research have evolved, for the most part, independently, we argue that they are fundamentally linked. This linkage is inevitable because an acceptable characterization of the interest rate process must be one where the implied changes in the yield curve's shape is consistent with those actually observed. For equilibrium models, it is necessary to show that the model reproduces not only the oft-observed yield curve shapes but also likely changes in shape as well. Conversely, for no-arbitrage models, the model's calibration must incorporate not only the observed yield curve and its volatility structure, but also how they change. Thus, when developing duration/convexity measures for bonds with embedded options, the challenge is to specify an interest-rate process that is consistent with both the observed yield curve and the way it changes shape and then measure the interest-rate risk that arises from these likely changes in shape. The goal of our analysis is to accomplish this task.

Our research combines the work of Hull and White (1993) on interest rate processes and that of Mann and Ramanlal (1998) on how to capture changes in the yield curve's shape. We develop a method for estimating the duration/convexity for bonds with embedded options that allows for realistic changes in the yield curve's shape. In addition, we document the impact of that the bond's contractual features (e.g. time to first call) and parallel versus non-parallel shifts in the yield curve have on these measures of interest rate risk.

The remainder of the paper is structured as follows. The yield curve data and sample period employed in our analysis is described in Section 1. Section 2 provides the model for the interest rate process as well as likely changes in the yield curve shape. Both are calibrated to market data. The impact of a callable bond's interest rate risk (measured by effective duration/convexity) to changes in the call's structure and to non-parallel shifts in the yield curve is documented in Section 3. Conclusions follow in Section 4.

2. YIELD CURVE DATA

We utilize yield curve data obtained from Coleman, Fisher, and Ibbotson's *Historical U.S. Treasury Yield Curves* (1995 edition) where spot rates are estimated using Treasury bond prices updated on a monthly basis during the sample period January 1985 to December 1994. For each month during the sample period, spot rates for the maturities (in years) 1, $1^1/_2$, 2, 3, 4, 5, 7, 10, 15, 20 and 30 are available.[1] We choose January 1985 as the beginning point of our sample period because it coincides with the initiation of the *Separate Trading of Registered Interest and Principal of Securities* (STRIPS) program.[2] The process of coupon stripping and reconstituting bonds in the STRIPS market has had a substantial impact on the U.S. Treasury market (see Fabozzi, 2000). Estimated spot rates for maturities less than one year are not utilized in our analysis due to their substantially higher volatility.

3. MODEL

The model has two parts: specification of the interest-rate process and characterization of changes in the spot curve's shape. As noted previously, these components are dependent in the sense that the interest-rate process implies movements in the spot curve that must be consistent with those observed. We consider each part in turn.

The Interest Rate Process

For the interest-rate process, we adopt a simpler version of Black and Karasinski's no-arbitrage model (1991). The model is simpler so as avoid the Hull and White (1994) criticism of non-stationary volatilities. Specifically, Black and Karasinski propose a one-factor model, which is calibrated to replicate exactly the spot curve, the volatility curve and the differential cap curve.[3] Hull and White claim that the Black and Karasinski model is over-parameterized which can lead to a nonstationary volatility structure. For example, if the current volatility curve is incorporated exactly into the model, it presupposes how the volatility will evolve in the future. Actual future volatility is likely to differ. As a result, calculated values of interest-rate contingent claims may be unreliable (see Caverhill, 1995; Hull & White, 1995). Thus, we consider the following model for the short rate r:

$$dLog(r) = [\theta(t) - aLog(r)]dt + \sigma dz \tag{1}$$

The model assumes interest rates follow a mean-reverting, lognormal process. The function $\theta(t)$ is chosen so the model replicates the spot curve exactly. The remaining parameters, a and σ, are constants that specify the speed at which the short rate reverts to its mean and volatility of the short rate, respectively. Assuming that a single factor, dz, drives the interest rate process implies that all spot rates are perfectly correlated. In Black and Karasinski (1991), the parameters a and σ are functions of time – just like $\theta(t)$ – chosen so that the model replicates the volatility curve and the differential cap curve exactly as well. Summarizing to this point, our model is a special case of Black and Karasinski since we replicate only the spot curve (via our choice of $\theta(t)$) and treat a and σ as constants rather than as functions of time.

This simple model has several positive attributes. Mean reversion ensures that the interest rate process is stationary and lognormality precludes negative interest rates. While the square root process would also ensures positive rates (see, Cox, Ingersoll & Ross, 1985), it only approximates the yield curve and does not fit every point.[4] Since our model replicates only the spot curve, it is not subject to overparameterization and thus avoids Hull and White's (1994) criticism of Black and Karasinski's (1991) model. One potential limitation of our model is that a single factor drives the interest rate process. Accordingly, as noted, all spot rates are perfectly correlated. Nevertheless, Mann and Ramanlal (1998) show that a single factor (the short interest rate and its predicted effect on the spot curve's slope and curvature) provides an accurate description of how the spot curve changes shape. As a result, we opt for the simplicity of a single factor model and avoid considering multi-factor models like those proposed by Brennan and Schwartz (1982) and Longstaff and Schwartz (1992).

The Spot Curve's Shape

In order to model how the spot curve changes shape, we draw on results from Mann and Ramanlal (1998). They consider several competing parameterizations of the spot curve and distinguish the critical factors that influence its shape. Then, they identify the specific parameterization and the corresponding factors that most closely predict actual changes in the spot curve's shape. Their findings are summarized in the following set of equations:

$$r_\tau = \alpha + \beta\tau + \gamma\tau^2 + e_\tau \tag{2a}$$

$$dr_\tau = \left[1 + \left(\frac{\delta\beta}{\delta\alpha}\right)\tau + \left(\frac{\delta\gamma}{\delta\alpha}\right)\tau^2\right]d\alpha + n_\tau \tag{2b}$$

In (2a), r_τ is the spot rate for a given term to maturity τ. The spot curve is parameterized by the current short rate α and the spot curve's slope and curvature, β and γ, respectively. In other words, the interest-rate process is characterized by three factors α, β, and γ. The residual term e_τ permits actual spot rates to deviate from their expected values (given by the quadratic expression $\alpha + \beta\tau + \gamma\tau^2$) by a random component that varies with the term to maturity τ. They assume that the short rate influences the spot curve's slope and curvature, i.e. $\beta = \beta(\alpha)$ and $\gamma = \gamma(\alpha)$. Accordingly, the change in the spot rate, dr_τ, for a given change in the short rate, $d\alpha$, is given by (2b), which is a single-factor representation of how the yield curve changes shape. The noise term n_τ is a random shock to the expected change in spot rates.

Litterman and Scheinkman (1991) motivate our choice of factors. They find that most of the variation in the spot curve's shape can be explained by three factors. Moreover, these factors correspond closely to the short rate (i.e. level) and the spot curve's slope and curvature. Jones (1991) shows that movements in the short rate (i.e. changes in level) are correlated with changes in slope/curvature. For example, when the spot curve shifts upward, it typically becomes flatter and less curved. Conversely, when the spot curve shifts downward, it typically becomes steeper and more curved. By incorporating the impact of short rate changes on the slope and curvature, the spot curve's parameterization effectively reduces to a single-factor specification that is most convenient given our choice of the one-factor model for the interest rate process in (1).

Compatibility of the Interest Rate Process and Changes in the Spot Curve's Shape

The next step toward pricing bonds with embedded options and ultimately measuring their interest rate risk is to calibrate (1) and (2) in a mutually compatible way. Equation (1) specifies the interest rate process while Eq. (2) characterizes changes in the spot curve's shape. The following set of Eq. (1) parameters must be estimated: $\theta(t)$ (function chosen to replicate the spot curve), a (degree of mean reversion), and σ (volatility of the short rate). Correspondingly, Eq. (2) parameters must also be estimated namely $\delta\beta/\delta\alpha$ (how a shift in the short rate impacts the spot curve's slope) and $\delta\gamma/\delta\alpha$ (how a shift in the short rate impacts the spot curve's curvature). These two sets of parameters are not independent. Specifically, the interest-rate process in (1) must be calibrated so that the implied movements in the spot curve are consistent with those specified in (2).

Why is this mutual consistency important? For the moment, let us first assume that the interest rate process follows (1) but ignore (2). We can then value a security using an interest rate tree that is generated for some appropriate choice of $\theta(t)$, a and σ. Given this valuation procedure, it is possible to determine how the security price changes given a shift in the spot curve. However, the shift chosen is arbitrary and will likely be inconsistent with either the spot curve's actual behavior or the behavior implied by the assumed interest-rate process. Indeed, several alternatives commonly used are susceptible to this inconsistency. For example, one approach is to assume the entire spot curve experiences a parallel shift, which is the assumption used when calculating modified duration and convexity for option-free bonds. Another approach is to cut the spot curve into sections and then calculate the sensitivity of the security's price with respect to changes in each section in turn, keeping the remainder of the spot curve unchanged (see, e.g. Ho, 1992; Willner, 1996).

Empirically, we know that the spot curve does not move in parallel shifts. Furthermore, partial durations with respect to specific factors (i.e. sections of the spot curve) albeit an improvement over traditional duration measures ignore correlations among the factors. Interest rate risk exposure must be calculated based on likely changes in the entire spot curve's shape, i.e. as specified in Eq. (2).

Measuring the interest rate risk of bonds with embedded options adds yet another layer of complexity. For example, suppose that a callable bond is priced using an interest rate tree assuming a particular option-adjusted-spread (OAS). Within the OAS framework, effective duration and convexity measures can be computed in a manner described by Kalotay, Williams and Fabozzi (1993) by increasing/decreasing interest rates at each node on the tree by a small amount while the OAS is held constant. The problem with this approach is two-fold. First, shifting each node of the tree by a fixed amount produces a new tree that is inconsistent with the distributional assumptions employed in generating the original tree (Babbel & Stavros, 1992). Second, assuming the interest-rate process follows a random walk is incompatible with the assumption that changes in the spot curve's shape follow a predictable pattern. In other words, whatever pattern we assume for how the spot curve changes shape when calculating the security's duration and convexity; this pattern must be consistent with the underlying interest-rate process that is employed to generate the tree when evaluating the security's value.

For both specifications (1) and (2), the short interest rate is exogenous and all spot rates are perfectly correlated. Notice that α in Eq. (2) is the current short rate, which corresponds to the initial value of r in Eq. (1). For each specification, the (instantaneous) standard deviation of the spot rate with maturity τ is related to the short rate's standard deviation and the model's parameters. For specification (1), this is given in Hull (1993):

$$\sigma(x_\tau) = \sigma(x_0)\frac{(1 - e^{-a\tau})}{a\tau}, \tag{3}$$

where $x_\tau = \text{Log } r_\tau$ (logarithm of the sport rate with term to maturity τ) and $\sigma(x_\tau)$ is the standard deviation of x_τ. Note that $\sigma(x_0) = \sigma$ which is the standard deviation of the logarithm of the short rate. Alternatively, for specification (2),

$$\sigma^2(y_t) = \sigma^2(y_0)\left(1 + \left(\frac{\delta\beta}{\delta\alpha}\right)\tau + \left(\frac{\delta\gamma}{\delta\alpha}\right)\tau^2\right)^2 + \sigma_\eta^2, \tag{4}$$

where $y_\tau = dr_\tau$ and $\sigma(yr_\tau)$, $\sigma(yr_0)$ are the standard deviations of the spot rates and short interest rates, respectively. The term $\sigma(n)$ is the standard deviation of the random shock n_τ in (2b). Notice that (4) is written in terms of nominal rates, whereas log rates are used to specify (3). This difference is due, in part, to the independent evolution of the related strands of literature and the fact that (3) and (4) are parsimonious specifications as opposed to being derived from some general equilibrium model.

Before implementing the model, we must estimate the parameters $\sigma(x_0)$ and a in (3) and $\sigma(r_0)$, $\delta\beta/\delta\alpha$, and $\delta\gamma/\delta\alpha$ in (4). As noted, the remaining function $\theta(t)$ in (1) is chosen so that the interest rate process replicates a given spot rate curve in order to satisfy the no-arbitrage condition. We estimate these parameters using the U.S. Treasury term structure data. The following procedure is adopted. First, we use the cubic-spline method to interpolate each of the 120 monthly spot curves in our ten-year sample period. Recall, for each spot curve, rates for only selected terms to maturity are available (specifically, 1, $1^1/_2$, 2, 3, 4, 5, 7, 10, 15, 20 and 30 years). Interpolation yields spot rates for all terms from 1 to 30 years. Second, for each term to maturity, we use the 120 spot rates to calculate the standard deviations $\sigma(x_\tau)$ and $\sigma(r_\tau)$. The latter term requires changes in spot rates and these are calculated over 6-month holding periods. Third, using a non-linear curve-fitting algorithm, we fit these standard deviations to the functional forms in (3) and (4), respectively. The resulting estimates are: $\sigma(x_0) = 0.2650$ and $a = 0.09173$ in (3) and $\sigma(r_0) = 0.0045$, $\delta\beta/\delta\alpha = -0.03903$ and $\delta\gamma/\delta\alpha = 0.0006922$, and $\sigma(n) = 0.0060$ in (4).

To assess the fit of our model, we plot the estimated standard deviations and the correspondingly fitted functionals for different maturities in Fig. 1. Panels A and B display results corresponding to (3) and (4), respectively. In each panel, Line 1 represents the estimated values and Line 2 represents the fitted form. Given the limited number of parameters and the strict functional forms, we deem the fit to be reasonably good while acknowledging the complexity of the underlying processes. The results can be summarized as follows:

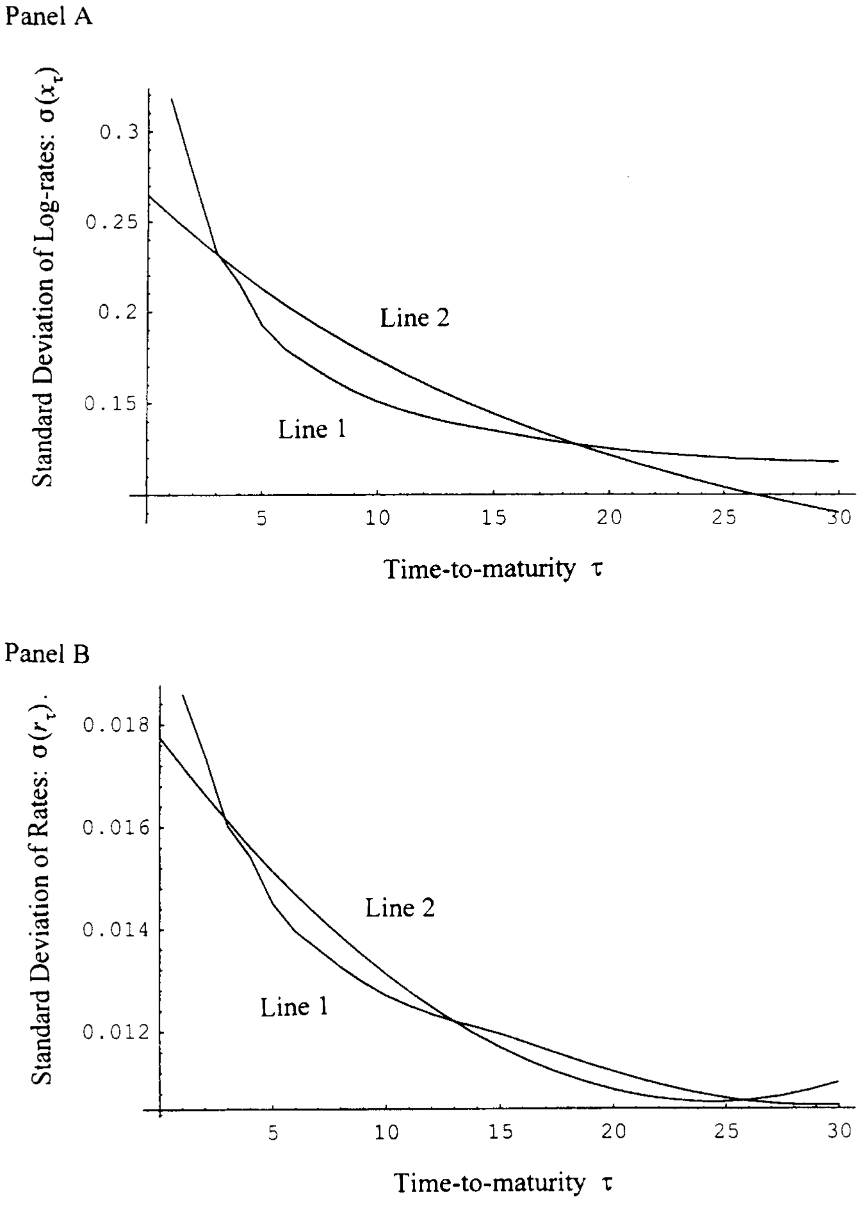


Fig. 1. Standard deviations of spot rates for different terms to maturity. Panel A (Panel B) displays the standard deviations of log (nominal) rates. In both panels, Line 1 displays the estimated standard deviations and Line 2 shows the fitted functional forms in Eqs (3) and (4).

Interest-rate process (Equation 1)	$a = 0.09173$ $\sigma(x_0) = 0.2650$
Spot-curve movements (Equation 2)	$\partial\beta/\partial\alpha = -0.03903$ $\partial\gamma/\partial\alpha = 0.0006922$

As a further check, we have also conducted a detailed analysis of the stability of parameter estimates of the model specified in Eqs (1) and (2). These parameter estimates specify: (a) the interest rate process, and (b) spot curve movements. Both are necessary to estimate interest rate risk of option-embedded options. Our sample period for yield curve data is the 120-month interval from January 1985 to December 1994. We analyze the parameter estimates in sub-periods several ways. First, we divide the 120-month interval into **four** non-overlapping 30-month periods (1–30, 31–60, 61–90, 91–120). Then, we divide the interval to **three** over-lapping 60-month periods (1–60, 31–90, 61–120). Next, into **two** overlapping 90-month periods (1–90, 31–120). Finally, we have the **single** 120-month interval (1–120). Thus, we have a total of 10 different sample periods. For each of these ten periods, the parameters estimates for a and $\sigma(x_0)$ in Eq. (3) are presented in Table 1. In all ten subperiods, the parameter estimates have the correct sign which suggests the model is well-specified.

The parameters estimates for $\delta\beta/\delta\alpha$, $\delta\gamma/\delta\alpha$, $\sigma(y_0)$, and $\sigma(n)$ in Eq. (4) are presented in Table 2 for each of these ten subperiods. Once again, the parameters estimates have the correct sign (i.e. the model is well specified) in all periods except the first (Months 1–30). On closer examination, we believe

Table 1. Sensitivity of Parameter Estimates for *a* and $\sigma(x_0)$ for Various Subperiods.

Sample Period	a	$\sigma(x_0)$
Months 1–30	0.0097	0.1727
Months 31–60	0.0022	0.0659
Months 61–90	0.0246	0.2045
Months 91–120	0.1544	0.2273
Months 1–60	0.0092	0.1322
Months 31–90	0.1272	0.1580
Months 61–120	0.1910	0.3082
Months 1–90	0.0403	0.1546
Months 31–120	0.1761	0.3164
Months 1–120	**0.0917**	**0.2650**

Table 2. Sensitivity of Parameter Estimates for $\delta\beta/\delta\alpha$, $\delta\gamma/\delta\alpha$, $\sigma(y_0)$, and $\sigma(n)$ for Various Subperiods.

Sample Period	$\delta\beta/\delta\alpha$	$\delta\gamma/\delta\alpha$	$\sigma(y_0)$	$\sigma(n)$
Months 1–30	0.7149	− 0.00212	0.0011	0.0102
Months 31–60	− 0.0580	0.00148	0.0065	0.0030
Months 61–90	− 0.0672	0.00111	0.0029	0.0028
Months 91–120	− 0.0010	0.00020	0.0013	0.0077
Months 1–60	− 0.0168	0.00028	0.0032	0.0075
Months 31–90	− 0.6403	0.00155	0.0072	0.0020
Months 61–120	− 0.0579	0.00090	0.0063	0.0039
Months 1–90	− 0.0346	0.00075	0.0039	0.0060
Months 31–120	− 0.0586	0.00116	0.0066	0.0035
Months 1–120	− 0.0390	0.00069	0.0045	0.0059

this results may be attributable to the fact that this period experienced particularly low volatility at the short end of the spot curve and the volatility curve was uncharacteristically flat.

Taken together, the results presented in Tables 1 and 2 suggest that the magnitudes of parameter estimates vary some from the those obtained when we estimate over the entire sample period (Months 1–120) which is the bottom row of each Table. However, this is not surprising given the observation of Hull and White (1994) on changing volatilities. The remedy in practice is to price securities using forward-looking volatility estimates obtained by interest-rate derivatives.

Implementing the Model

There are several available methods to implement stochastic interest-rate models. These methods differ along a number of dimensions. For example, models can be either binomial (see, e.g. Ho & Lee, 1986) or trinomial (see, e.g. Hull & White, 1993) depending on the number of possible branches at each node. Models may also be contrasted in terms of all possible interest rate paths versus only those paths that are most likely to occur (see, e.g. Ho, 1992). Still others propose varying the time step to gain numerical efficiency (Black, Derman & Toy, 1990) while others advance efficient algorithms like forward induction methods (Jamshidian, 1991) and efficient iterative procedures (Bjerksund & Stensland, 1996). In most applications, accuracy and speed are both important considerations. This is particularly true for this study because we are attempting to evaluate not only security prices but also the first/second derivatives of price with respect to changing interest rates. Accurate prices are essential. However, more importantly, spurious errors owing to the tree's discreteness must be minimal. With these factors in mind, we employ the methodology developed by Hull and White (1994) which is particularly suited to our purpose. Our no-arbitrage interest-rate tree (corresponding to Eq. (1)) is constructed according to the blueprint provided in Section III of their paper to which we refer the reader.

Pricing bonds with embedded options on an interest rate tree is straightforward. The terminal cash flows are specified at the tree's end nodes, which correspond to the security's maturity date. These cash flows are discounted back one node at a time using the forward interest rate at that node. At each preceding node, as we discount cash flows, if the bond's value exceeds the contractually stipulated call price plus accrued interest, the security will be called and accordingly the bond's value at that node is set equal to the call price plus accrued interest. This adjustment is made before further discounting down the tree. In addition, for a coupon bond, at each node corresponding to the coupon date, the coupon amount is added to the bond's value at that node. The revised value is then tested against the call provision and updated accordingly as previously described. Further discounting then proceeds until we arrive at the initial node of the tree, at which point we have the security's current price P.

To calculate the effective duration and convexity of a bond with an embedded option, we adopt the methodology proposed in Kalotay, Williams and Fabozzi (1993). Specifically, we upshift and downshift the spot curve by a prespecified number of basis points. We then reconstruct the interest rate tree consistent with the shifted curve and reevaluate the bond's price as above. The prices corresponding to the upshift and downshift are denoted by P^+ and P^-, respectively. Given the initial spot curve, the shift at each point on the curve varies in magnitude. For a specific shift in the short rate $\Delta\alpha$, this magnitude is given by Eq. (2b). Following this method ensures that the shift we consider is consistent with the distributional properties of the underlying interest-rate process.

Given the values for P, P^+ and P^-, effective duration and convexity can be calculated using the following expressions:

$$\mathrm{Dura}_{\mathrm{effective}} = -\frac{1}{P}\frac{dP}{d\alpha} = -\frac{(P^+ - P^-)}{2P\Delta\alpha} \tag{5a}$$

$$\mathrm{Conv}_{\mathrm{effective}} = \frac{1}{P}\frac{d^2P}{d\alpha^2} = \frac{(P^+ - 2P + P^-)}{P(\Delta\alpha)^2} \tag{5b}$$

It is natural to calculate duration and convexity with respect to changing short rates, as opposed to shifting segments of the spot curve, because the short rate is the only exogenous factor in the model. All other rates are endogenous.

4. RESULTS ON INTEREST RATE EXPOSURE

In the analysis that follows, we determine the impact of security-specific factors and model-specific attributes on interest-rate exposure. Specifically, we analyze the impact of (1) contractual features like call deferrals and coupon rates and (2) assumptions about changes in the term structure of interest rates on effective duration and convexity. Our sample period of interest rates is from January 1985 to December 1994. For simplicity, we present our analysis using term structure data from the twenty-fifth month in our sample period (i.e. January 1987) to calibrate the tree. We have repeated the analysis for several other term structures in our sample period to insure that the results are not sensitive to any specific period. We calculate the effective duration and convexity using (5a) and (5b), respectively, for $\Delta\alpha$ of plus/minus 20 basis points. The interest-rate tree is constructed with 750 time steps per calendar year for a total of 7500 steps for the 10-year maturity bonds that we consider. We do this to ensure the integrity of the numerical results.

The analysis is performed on bonds with 10 years until maturity that are callable at par. We examine the impact on interest-rate exposure of differing contractual features. To accomplish this task, we vary the amount of time until the first call date from 2 to 10 years, in increments of 1 year, for a total of 9 cases.[5] In addition, we consider coupon rates between 5 and 10% in increments of $^1/_2$% for a total of 11 cases. Since the yield curve to which we calibrate the model has interest rates around 7%, this range of coupon rates allows us to examine the impact of varying degrees of moneyness on effective duration/convexity. In summary, we examine a total of 99 different cases according to varying coupon rates and call dates.

For each of these 99 cases, we also examine the effect of either ignoring the call feature or assuming the term structure only incurs parallel shifts has on a bond's duration and convexity. It is often convenient to make these assumptions because it allows the usage of the traditional formulas for these measures corresponding to straight bonds. However, if they are judged to have a significant impact on a callable bond's measured interest rate risk, then retaining these two assumptions can be costly.

We undertake a three-step approach to analyze the marginal impact of each of these assumptions. First, we ignore the call feature altogether and assume the spot curve moves only in parallel shifts. Valuation is accomplished by simply discounting cash flows and using the well-known formulas for duration and convexity. Second, we account for the call feature and consider various times to the first call date, but continue to assume the spot curve moves in parallel shifts. In other words, we value bonds by discounting along the interest-rate tree generated by Eq. (1) but in Eq. (2) we assume that $\delta\beta/\delta\alpha = 0$ (i.e. changes in the level of the short rate have no impact on spot curve's slope) and $\delta\gamma/\delta\alpha = 0$ (i.e. changes in the level of the short rate have no impact on the spot curve's curvature). In the third step, we account for both the call feature and nonparallel shifts. Specifically, we value bonds by discounting along the tree based on (1) and in (2) we assume $\delta\beta/\delta\alpha = -0.03903$ and $\delta\gamma/\delta\alpha = 0.0006922$. By comparing the results in step two with those in step one, we document the effect of the call feature on measured interest rate risk. The marginal effect of non-parallel shifts becomes apparent by comparing results in step three to those in step two.

Effective Duration

Our estimates for effective duration are displayed in Fig. 2 (Panels A and B) and Fig. 3 (Panels A and B). In both sets of figures, the horizontal axes represent the 99 cases described above with different times to the first call date and various coupon rates. Specifically, one horizontal axis represents the 9 different call-deferred periods from 2 to 10 years (1 year increments) while the other horizontal axis represents the 11 different coupons rates that range from 5 to 10% ($^1/_2$% increments). The following notation is used:

$\text{Dura}_{\text{noncallable-parallel}}$: Effective duration when the call feature is ignored and the spot curve is assumed to move in parallel shifts.

$\text{Dura}_{\text{callable-parallel}}$: Effective duration accounting for the bond's call feature but assuming the spot curve moves in parallel shifts.

$\text{Dura}_{\text{callable-nonparallel}}$: Effective duration accounting for both the bond's call feature and nonparallel shifts in the spot curve

$\text{Dura}_{\text{noncallable-parallel}}$, the effective duration for the base case, ignoring the call feature and assuming parallel shifts, is given by the traditional duration formula for straight bonds. $\text{Dura}_{\text{callable-parallel}}$, the effective duration of callable bonds with parallel shifts, is displayed in Panel A of Figure 2. For convenience, Panel A also displays duration measures for the base case, which simply assumes that the call-deferred period equals 10 years given that we examine bonds with 10 years to maturity. This base case is denoted by Line x.

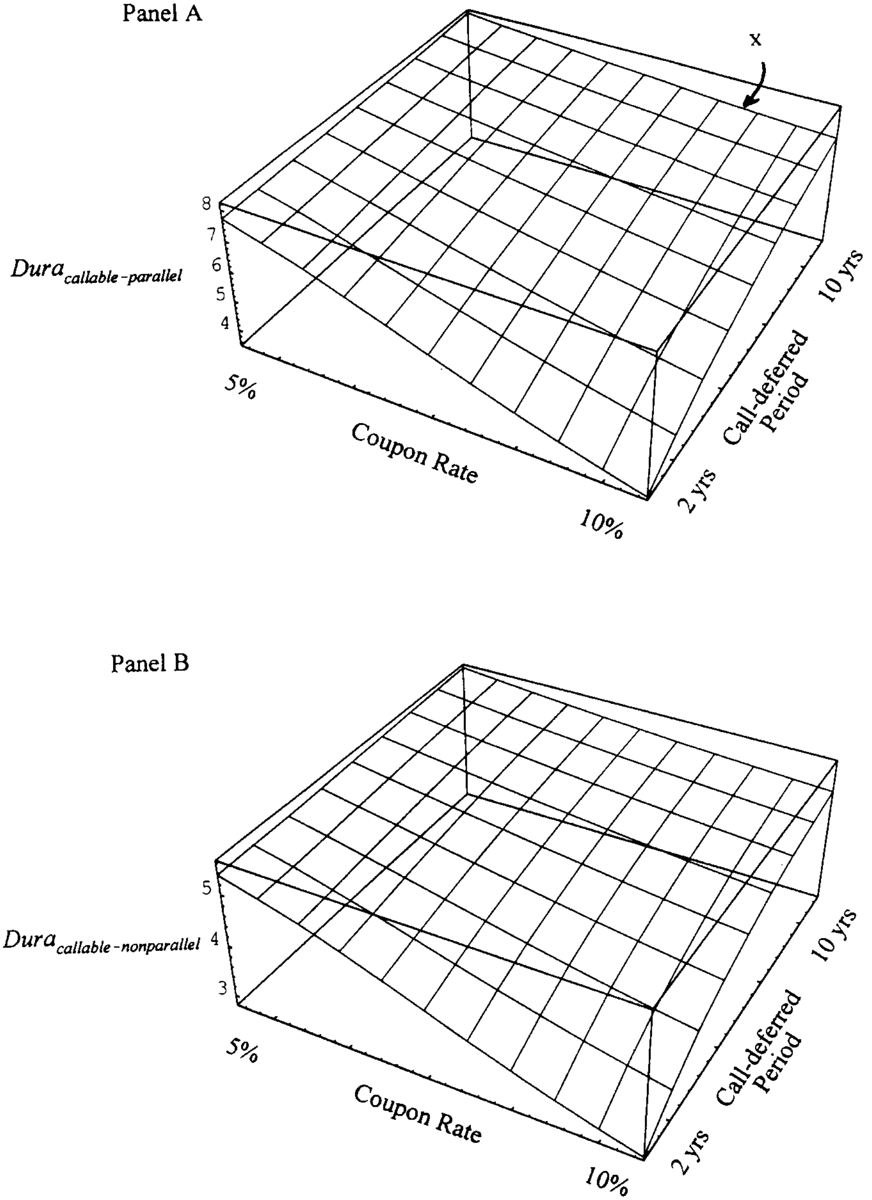
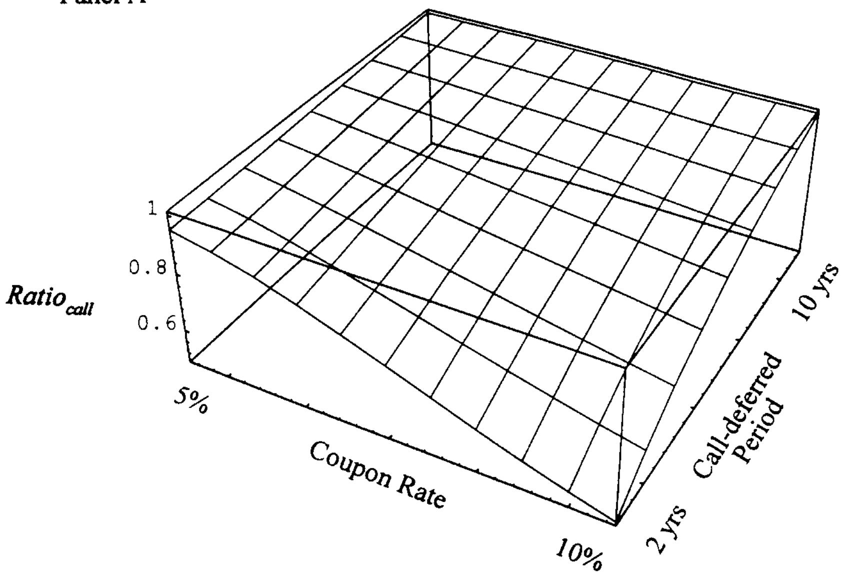


Fig. 2. Effective durations for different coupon rates and call-deferred periods for 10-year bonds. Panel A (Panel B) displays durations of call-deferred bonds for parallel (nonparallel) shifts in the spot curve.

Panel A



Panel B

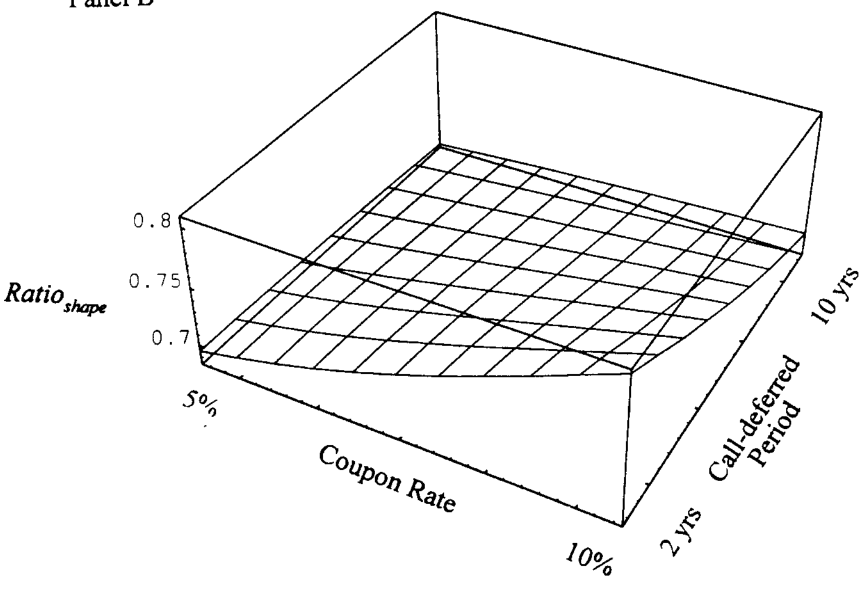


Fig. 3. Ratio of effective durations. Panel A displays the ratio of durations of callable vs. noncallable bonds for parallel shift in the spot curve. Panel B displays for callable bonds the ratio of durations for nonparallel vs. parallel shifts in the spot curve.

Let's examine the information presented in Panel A. In the base case, we are considering how effective duration changes as we change the coupon rates when the spot curve experiences parallel shifts. The effective duration of this 10-year bullet at the lowest coupon rate (5%) is approximately 8 (i.e. the top left rear corner of the graph). As we move to the right along Line x, we obtain the familiar result that when the coupon rate rises the effective duration falls. If we hold the coupon rate constant and move along the lines that run from the back of the graph to the front, we see that as the call-deferred period decreases the effective duration also decreases. Moreover, this reduction in duration is much more dramatic for high-coupon bonds (10%) than it is for low coupon bonds (5%). The reason is simple. High coupon bonds (in particular, bonds with high coupon rates relative to spot rates) will almost certainly be called as their first call date arrives because they are likely to be in-the-money. As a result, the effective duration of high-coupon bond callable in two years approaches the duration of a two-year non-callable bond. Conversely, the decrease in the call-deferred period for low-coupon bonds (in particular, bonds with low coupon rates relative to spot rates) lowers the effective duration only slightly since these bonds are unlikely to be called given they will be out-of-the-money.

Given the results in Panel A, we can draw the following conclusions about the effective duration of callable vs. non-callable bonds. It is well-known that, other things equal, the effective duration of a non-callable decreases as the bond moves toward maturity. The effective duration of a callable bond decreases at a faster rate than a non-callable bond as it moves toward maturity, other things equal. The faster decrease is due to the combination of two effects: the reduction in the time to maturity and the reduction in time until the first call date. The latter factor will dominate as the coupon rate rises.

Panel B of Fig. 2 displays similar results for callable bonds allowing for realistic changes in the spot curve's shape (slope and curvature) for a given shift in the short interest rate. The most telling result can be gleaned by examining the numbers that label the vertical axis. The effective duration for all 99 cases ranges from approximately 5.5 to approximately 3.0 whereas in Panel A, they range from 8.0 to 4.0. Accordingly, the effective duration of a callable bond is measurably lower when we consider realistic changes in the spot curve's shape compared to traditional duration measures that assume parallel movements in the spot curve. The intuition is straightforward. For an upward (downward) shift in the short interest rate, the net effect of a flattening (steepening) slope and a decreased (increased) curvature is to moderate the impact of the shift on longer rates and correspondingly on the bond's value. As a result, a callable bond's sensitivity to a spot curve shift is effectively lower.

We will return to this issue in more detail below. Note, as well, the surface in Panel B has the same basic shape as Panel A – slopes downward as one moves from the back of the figure to the front and from the left of the figure to the right.

Figure 3, Panel A displays the ratio of the effective durations obtained by accounting for call features versus ignoring them, assuming a parallel shift in the spot curve in both cases. The ratio is denoted as $\text{Ratio}_{\text{call}}$ and is computed as follows:

$$\text{Ratio}_{\text{call}} = \frac{\text{Dura}_{\text{callable-parallel}}}{\text{Dura}_{\text{noncallable-parallel}}} \tag{6}$$

This ratio conveys the marginal impact of the call feature on duration. As before, it is calculated for 99 previously described cases. The intuition gleaned from Panel A in Fig. 2 is confirmed. Ignoring the call feature overstates effective duration, and this overstatement is more pronounced as coupon rates increase and as call-deferred periods shorten. Indeed, when the coupon rate is 10% and the call-deferred period is 2 years, ignoring the call feature in our framework can introduce a 100% error in the calculated value of duration.

The marginal impact of realistic changes in the spot curve's shape on duration is depicted in Panel B of Fig. 3. This panel shows the ratio of effective durations assuming non-parallel and parallel shifts, accounting for the call feature in both cases. The ratio is denoted by $\text{Ratio}_{\text{shape}}$ and is computed as follows:

$$\text{Ratio}_{\text{shape}} = \frac{\text{Dura}_{\text{callable-nonparallel}}}{\text{Dura}_{\text{callable-parallel}}} \tag{7}$$

As before, these ratios are calculated for all 99 cases of coupon rates and call-deferred periods. The surface in Panel B slopes upward from the back (long call-deferred period) to the front (short call-deferred period) and slopes up from left (low coupon rates) to the right (high coupon rates). Accordingly, we can draw the following inference. When accounting for realistic changes in the yield curve's shape, the percentage reduction in effective duration is higher for longer call deferrals and lower coupon rates. These are also bonds that have the highest effective duration.

The intuition for this result is as follows. As is well-known, when the spot curve only shifts in a parallel manner, longer duration bonds are more responsive to a given change in rates. However, when both slope and curvature change as predicted, the spot rates obviously do not change by the same number of basis points for all maturities. For an upward shift in yield, the flattening of the curve moderates the impact of changing short rates on long rates, while the decreased curvature exacerbates it. The impact of the change in slope dominates, as expected, given the relative sizes of the estimated parameters in Eq. (2). For a downward shift, the net impact of the steepening slope and increased curvature is again to moderate the impact of changing short rates on long rates. Thus for both upward and downward shifts in the short rate, the net moderating effect of changes in slope and curvature is greatest for highest duration bonds – lowering their effective duration by more proportionally. This effect is consistent with what we observe empirically, namely, the spot-rate volatility decreases as the term to maturity increases.

Effective Convexity

Estimates for effective convexity are displayed in Fig. 4. As in the previous figures, one horizontal axis represents the 9 different call-deferred periods ranging from 2 to 10 years in 1-year increments. The other horizontal axis represents the 11 different coupon rates ranging from 5 to 10% in $^1/_2$% increments. Convexity measures are displayed along the vertical axis. The following notation is used:

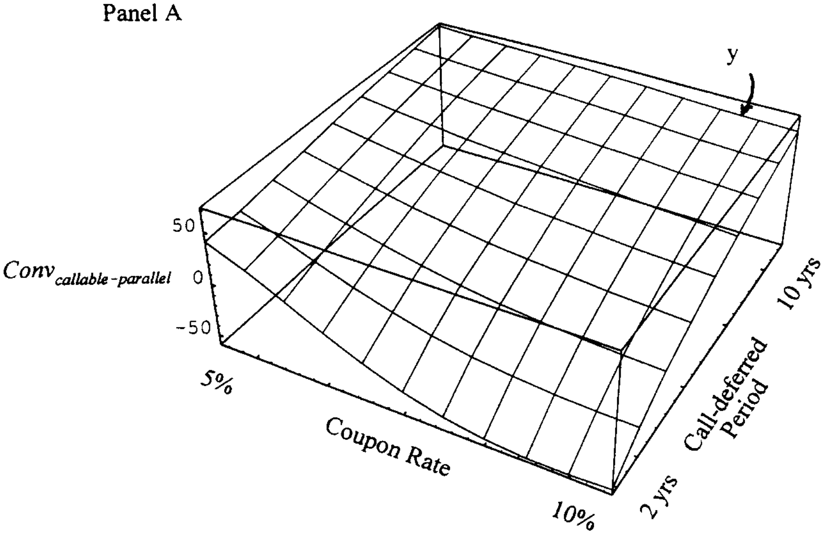
$\mathrm{Conv}_{\text{noncallable-parallel}}$: Effective convexity when the call feature is ignored and the spot curve is assumed to move in parallel shifts.

$\mathrm{Conv}_{\text{callable-parallel}}$: Effective convexity, accounting for the bond's call feature, but assuming the spot curve moves in parallel shifts.

$\mathrm{Conv}_{\text{callable-nonparallel}}$: Effective convexity accounting for both the bond's call feature and nonparallel shifts in the spot curve.

$\mathrm{Conv}_{\text{noncallable-parallel}}$, the effective convexity for the base case ignoring the call feature and assuming parallel shifts, is given by the traditional convexity formula for straight bonds. $\mathrm{Conv}_{\text{callable-parallel}}$, the effective convexity of callable bonds with parallel shifts, is displayed in Panel A of Fig. 4. For convenience, Panel A also displays convexity measures for the base case, which simply assumes that the call-deferred period equals 10 years, given that we examine bonds with 10 years to maturity. This base case is denoted by the Line y.

Let us examine the information presented in Panel A. In order to understand these results, we need to consider the plight of the callable bond owner. The bondholder has, in effect, two positions – a long position in a noncallable bond and a short position in a call option. As a result, the price of a callable bond is simply the price of the noncallable bond less the price of the call option. The



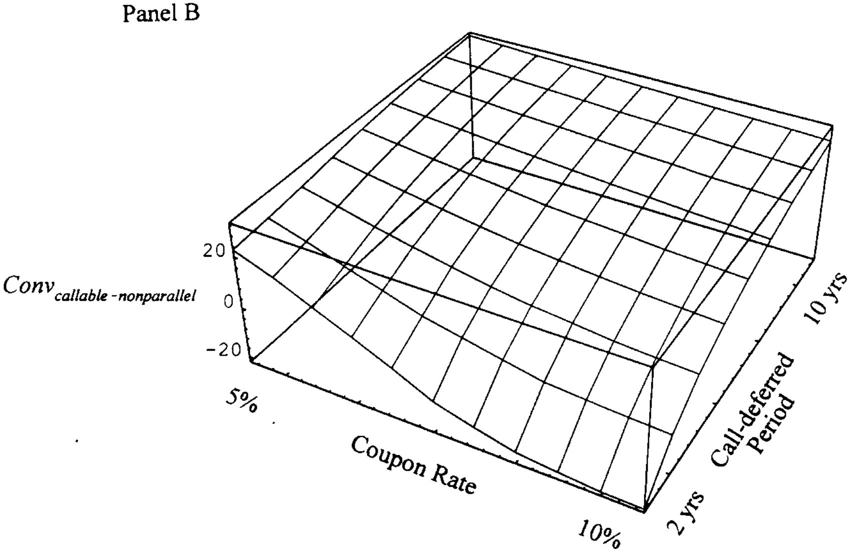


Fig. 4. Effective convexities for different coupon rates and call-deferred periods for 10-year bonds. Panel A (Panel B) shows convexity of call-deferred bonds for parallel (nonparallel) shifts in the spot curve.

option component of the callable bond alters its convexity, i.e. its price versus yield relationship. In general, a callable bond has a lower convexity than its noncallable counterpart.

As noted, Line y in Panel A of Fig. 4 represents the convexity of a ten-year noncallable bond for various coupon rates. Line y slopes downward to the right because convexity is higher when coupon rates are lower, all else equal. As we shorten the call-deferred period and/or increase the coupon rates, the convexity decreases and eventually turns negative (front right corner of the figure). The intuition is straightforward once again. The higher the coupon rate and the shorter the time until the first call date, the more valuable the embedded call option (which the callable bondholder has written) and the less valuable the callable bond is relative to its noncallable counterpart. These factors impede the callable bond's price rise as yields fall causing convexity to become negative. Simply put, the decreased convexity of the callable bond is driven by the increased value of the embedded call option whose value increases even further as yields fall.

Panel B of Figure 4 displays the convexities for callable bonds and allows for realistic changes in the spot curve's shape (slope and curvature) for a given shift in the short interest rate. Note the scale on the vertical axis, the convexities range from approximately 20 to –20. The convexity results are quite similar to those for duration. When we account for realistic changes in the yield curve shape, the convexities for callable bonds are uniformly lower in absolute value (i.e. closer to zero). The intuition is similar as well. For an upward (downward) shift in the short rate, the net effect of a flattening (steepening) slope and decreased (increased) curvature is to moderate the impact of the shift on long rates. Accordingly, a callable bond's sensitivity to a yield curve shift is effectively lower.

It is apparent from Panels A and B of Figure 4 that ignoring either the bond's contractual call features or the likely ways that the yield curve changes shape can introduce significant errors in the calculations for convexity.

Implications for Measuring Interest Rate Risk

Accounting for the call feature and realistic changes in the yield curve's shape provides a more accurate measure of a callable bond's interest rate risk. Consider a portfolio manager who wants to hedge a portfolio of callable bonds. The manager will take an opposite position in cash-market securities or in a derivative instrument, so that any loss in the position to be hedged is offset by a gain in the hedging vehicle. To accomplish this task, the portfolio manager should match the duration and convexity (accounting for the call feature and realistic changes in the yield curve's shape) of the position to be hedged with the hedging vehicle. For any change in interest rates, the change in the value of the hedged position will more closely mirror a change in the hedging vehicle if the duration and convexity measures we propose are employed.

Accounting for the call feature and realistic changes in the yield curve's shape also has implications for assessing the cost of convexity. To see this, consider two bonds, Bond A and Bond B, that have the same duration (calculated in the traditional manner) but different convexities. Suppose Bond A has the higher convexity. Other things equal, Bond A will (almost always) have a higher price and a lower yield. The difference in yields between these two bonds is the amount that investors are willing to forego in exchange for the prospect of enhancing portfolio performance as a result of the higher convexity. This difference in yields is often referred to as the cost or value of convexity. The same scenario can be applied to noncallable and callable bonds. Part of the difference in yield between otherwise identical callable and noncallable bonds is due to the higher convexity of the noncallable bond. While the benefits of greater positive convexity are apparent when the yield curve is flat and shifts in parallel, Ilmanen (1995) demonstrates that the benefits of higher convexity depend on how the yield curve changes shape when yields change. Thus, in order to properly assess the value of convexity it seems logical to employ duration/convexity measures that account for realistic changes in the yield curve's shape when measuring interest rate risk for a callable bond.

5. CONCLUSION

We introduce duration and convexity measures for bonds with embedded options that explicitly account for realistic changes in the yield curve's shape. A key contribution of our analysis is the specification of an interest rate process that is consistent with the observed yield curve and how its shape changes, and a prescription for measuring duration and convexity in this environment. In doing so, we merge two important strands of fixed-income research. We illustrate the features of duration/convexity measures using data from U.S. Treasury yield curves.

Our numerical results document the marginal impact of changes in the bond's call feature (e.g. length of call-deferred period) and parallel vs. non-parallel shifts in the yield curve on interest rate risk. We find that the effective duration of a callable bond is measurably less when we consider realistic changes in the spot curve's shape than traditional duration measures that assume parallel movements in the spot curve. Ignoring the call feature can also introduce significant errors when measuring interest-rate risk. Moreover, the convexity differentials between noncallable and callable bonds are much smaller than traditional convexity measures would lead us to believe.

NOTES

1. Details of the procedure for estimating spot rates are provided in Coleman, Fisher and Ibbotson.
2. The STRIPS program was announced by the U.S. Treasury in August 1984 and began in February 1985.
3. A differential cap pays the difference between a short interest rate and the strike rate, if positive, multiplied by some notational principal on a particular date. A cap is a package of differential caps. The differential cap curve depicts the relationship between the prices of at-the-money caps and maturity.
4. While this model is used widely on Wall Street, the fact that the square root model only approximates the yield curve is a source of concern among derivatives traders. One could easily argue that a minor discrepancy in the yield curve fit could be magnified through the leverage inherent in virtually any derivatives position into a severe problem.
5. A bond with 10 years to maturity which is call-deferred for 10 years is simply non-callable or "bullet" bond.

REFERENCES

Amin, K. I., & Morton, A. J. (1994). Implied Volatility Functions in Arbitrage-free Term Structure Models. *Journal of Financial Economics*, *35*, 141–180.

Babbel, D. F., & Zenios, S. A. (1992). Pitfalls in the Analysis of Option-adjusted Spreads. *Financial Analysts Journal*, *48*, 65–69 (July-August).

Bierwag, G. O., Kaufman, G. G., & Latta, C. M. (1987). Bond Portfolio Immunization: Tests of Maturity, One- and Two-factor Duration Matching Strategies. *Financial Review*, *22*, 203–219 (May).

Bjerksund, P., & Stensland, G. (1996). Implementation of the Black-Derman-Toy Interest Rate Model. *Journal of Fixed Income*, *6*, 67–75 (September).

Black, F., Derman, E., & Toy, W. (1990). A One-factor Model of Interest Rates and its Application to Treasury Bond Options. *Financial Analysts Journal*, *46*, 33–39 (January-February).

Black, F., & Karasinski, P. (1991). Bond and Option Pricing When Short Rates Are Lognormal. *Financial Analysts Journal*, *47*, 52–59 (July-August).

Brennan, M. J., & Schwartz, E. S. (1982). An Equilibrium Model of Bond Pricing and a Test of Market Efficiency. *Journal of Financial and Quantitative Analysis*, *17*, 301–329 (September).

Carverhill, A. P. (1995). A Note on the Models of Hull and White for Pricing Options on the Term Structure. *Journal of Fixed Income*, *5*, 89–96 (September).

Coleman, T., Fisher, L., & Ibbotson, R. (1995). *U.S. Treasury Yield Curves 1926–1994*. New York: Moody's Investor Service.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, *53*, 385–407.

Fabozzi, F. J. (2000). *Bond Markets, Analysis and Strategies*, 4th edition. Upper Saddle River: Prentice Hall.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, *60*, 77–105.
Ho, T. S. Y. (1992). Key Rate Durations: Measures of Interest Rate Risk. *Journal of Fixed Income*, *2*, 29–44 (September).
Ho, T. S. Y., & Lee, S. B. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance*, *41*, 1011–1029 (December).
Hull, J. (1993). *Options, Futures and Other Derivative Securities*. Prentice Hall.
Hull, J., & White, A. (1990). Pricing Interest-rate Derivative Securities. *Review of Financial Studies*, *4*, 573–592.
Hull, J., & White, A. (1993). One-factor Interest-rate Models and the Valuation of Interest-rate Derivative Securities. *Journal of Financial and Quantitative Analysis*, *28*, 235–254.
Hull, J., & White, A. (1994). Numerical Procedures for Implementing Term Structure Models I: Single-factor Model. *The Journal of Derivatives*, *2*, 7–16 (Fall).
Hull, J., & White, A. (1995). A Note on the Models of Hull and White for Pricing Options on the Term Structure: Response. *Journal of Fixed Income*, *5*, 97–102 (September).
Ilmanen, A. (1995). Convexity Bias and the Yield Curve. *Understanding the Yield Curve*, Part 5. New York: Salomon Brothers.
Jamshidian, F. (1991). Forward Induction and Construction of Yield Curve Diffusion Models. *Journal of Fixed Income*, *1*, 62–74 (June).
Jones, F. J. (1991). Yield Curve Strategies. *Journal of Fixed Income*, *1*, 43–51 (September).
Kalotay, A. J., Williams, G. O., & Fabozzi, F. J. (1993). A Model for Valuing Bonds and Embedded Options. *Financial Analysts Journal*, *49*, 35–46 (May–June).
Litterman, R., & Scheinkman, J. (1991). Common Factors Affecting Bond Return. *Journal of Fixed Income*, *1*, 54–61 (June).
Longstaff, F. A., & Schwartz, E. S. (1992). Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model. *Journal of Finance*, *47*, 1259–1282.
Mann, S. V., & Ramanlal, P. (1997). Relative Performance of Yield Curve Strategies. *Journal of Portfolio Management*, *23*, 64–70 (Summer).
Mann, S. V., & Ramanlal, P. (1998). Duration and Convexity Measures When the Yield Curve Changes Shape. *Journal of Financial Engineering*, *7*, 35–58.
Reitano, R. R. (1990). Non-parallel Yield Curve Shifts and Durational Leverage. *Journal of Portfolio Management*, *16*, 62–67 (Summer).
Rendleman, R., & Bartter, B. (1980). The Pricing of Options on Debt Securities. *Journal of Financial and Quantitative Analysis*, *15*, 11–24.
Vasicek, O. A. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, *5*, 177–188.
Willner, R. (1996). A New Tool for Portfolio Managers: Level, Slope and Curvature Durations. *Journal of Fixed Income*, *6*, 48–59 (June).

TWO-FACTOR JUMP-DIFFUSION INTEREST RATE PROCESS: AN EMPIRICAL EXAMINATION IN TAIWAN MONEY MARKET

Shih-Kuo Yeh and Bing-Huei Lin

ABSTRACT

In this paper, we investigate a jump-diffusion process, which is a mixture of an O-U process with mean-reverting characteristics used by Vasicek (1977) and a compound Poisson jump process, for the term structure of interest rates. We develop a methodology for estimating both the one-factor and two-factor jump-diffusion term structure of interest rates models and complete an empirical study for Taiwan money market interest rates.

In the empirical study, we use weekly interest rates on the 10-day, 30-day, 90-day, and the 180-day commercial papers to estimate parameters in the one-factor and the two-factor jump-diffusion models. The results show that the jump-diffusion model is significant, either for the one-factor model or the two-factor model, with the two-factor model fitting better. In explanations, the first factor is more associated with shorter-term interest rates, and the second factor is associated with both short-term and longer-term interest rates.

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 255–282.

ISBN: 0-7623-0798-6

1. INTRODUCTION

Conventionally, financial variables such as stock prices, foreign exchange rates, and interest rates are assumed to follow a diffusion process with continuous time paths when pricing financial assets. Despite their attractive statistical properties and computation convenience by which they are unanimously accepted for theoretical derivation, more and more empirical evidence has shown that pure diffusion models are not appropriate for these financial variables. For example, Jarrow and Rosenfeld (1984), Ball and Torous (1985), Akgiray and Booth (1986), Jorion (1988), and Lin and Yeh (1997) all found evidence indicating the presence of jumps in the stock price process. Tucker and Pond (1988), Akgiray and Booth (1988), and Park, Ann and Fujihara (1993) studied foreign exchange markets and concluded that the jump-diffusion process is more appropriate for foreign exchange rates. In pricing and hedging with financial derivatives, jump-diffusion models are particularly important, since ignoring jumps in financial prices will cause pricing and hedging risks.

For interest rates, jump-diffusion processes are particularly meaningful since the interest rate is an important economic variable which is, to some extent, controlled by the government as an instrument for its financial policy. Hamilton (1988) investigated U.S. interest rates and found changes in regime for the interest rate process. Das (1994) found movements in interest rates display both continuous and discontinuous jump behavior. Presumably, jumps in interest rates are caused by several market phenomena, such as money market interventions by the Fed, news surprises, and shocks in the foreign exchange markets, and so on.

Classical term structure of interest rate models, such as the Vasicek (1977) model, the Cox, Ingersoll and Ross (CIR, 1985) model, the Brennan and Schwartz (1978) model, and other extended models all assume that processes of state variables (such as the short-term interest rate, or the long-term interest rate, or others) which drive interest rate fluctuations follow various diffusion processes. Their assumptions are inconsistent with the *a prior* belief and empirical evidence regarding the existence of discontinuous jumps in interest rates. At a cost of additional complexity, Ahn and Thompson (1988) extended the CIR model by adding a jump component to the square root interest rate process. Using a linearization technique, they obtained closed-form approximations for discount bond prices. Similarly, Baz and Das (1996) extended the Vasicek model by adding a jump component to the O-U interest rate process, and obtained closed-form approximate solutions for bond prices by the same linearization technique. They also showed that the approximate formula is quite accurate.

Although theoretical derivations for the jump-diffusion term structure models have been developed, the associated empirical work has not been done. A formal model of the term structure of interest rates is necessary for the valuation of bonds and various interest rate options. More importantly, parameter values or estimates are required for the implementation of a specific model. To price interest rate options, with closed-form solutions or by numerical methods, one must have values of the parameters in the stochastic processes that determine interest rate dynamics. Hence parameter estimation is a very first step in the application and analysis of interest rate option pricing models.

In this study, we investigated a jump-diffusion process, which is a mixture of an O-U process with mean-reverting characteristics used by Vasicek (1977) and a compound Poisson jump process, for interest rates. Closed-form approximate solutions for discount bond prices were derived by Baz and Das (1996). Essentially the approximate model is a one-factor term structure model. It has the disadvantage that all bond returns are perfectly correlated, and it may not be adequate to characterize the term structure of interest rates and its changing shape over time. However, the model at least can incorporate jump risks into the term structure model, making the model more complete relative to the pure diffusion Vasicek model. In addition, just as the simple diffusion Vasicek model, the short-term interest rate can move to negative values under the extended jump-diffusion Vasicek model. Realizing these limitations, this study did not attempt to propose a "perfect" model for pricing term structure of interest rates, instead its purpose was to show whether a jump-diffusion model is statistically significant, and appropriate for describing interest rate dynamics.

In this study, we also extended the Baz and Das (1996) one-factor jump-diffusion model to a multi-factor jump-diffusion model, and developed a methodology for estimating the extended Vasicek jump-diffusion term structure of interest rates model and completed an empirical study for Taiwan money market interest rates. The state variables (such as the instantaneous short-term interest rate, and other factors) that drives the term structure of interest rates dynamics was not observable, and the observed bond prices are functions of the state variable. Thus we needed to use the change of variable technique to obtain the likelihood function in terms of the observed bond prices, in order to conduct a maximum likelihood estimate. The estimation procedure of this study is similar to Chen and Scott (1993).

In the empirical study, we use weekly interest rates on the 10-day, 30-day, 90-day, and the 180-day commercial papers to estimate parameters in the one-factor and the two-factor jump-diffusion models. The sample period is from

July 23, 1983 to September 27, 1997. The results show that the jump-diffusion model is significant, either for the one-factor model or the two-factor model, with the two-factor model fitting better. This is as expected, since one-factor models do not fit the versatile term structure of interest rates very well. For the two-factor model, compared to the second factor, the first factor exhibits characteristics of stronger mean-reversion, higher volatility, and more frequent jumps in the process. In explanations, the first factor is more associated with shorter-term interest rates, and the second factor is associated with both short-term and longer-term interest rates.

Since the assumption of an appropriate stochastic process for the interest rate and the estimation of its associated parameters are of critical importance when pricing and hedging with term structure of interest rates and interest rate derivatives, the results and the methodology for estimating parameters in the jump-diffusion process have important implications for the area of financial engineering.

The rest of this paper is organized as follows: Section 2 specifies the Vasicek and the jump-diffusion term structure of interest rates models. Section 3 presents the empirical methodology used in this study. Section 4 specifies the data, and analyzes the results of parameters estimation and term structure fitting. Section 5 is the summary of the study. Some proofs are in the Appendices.

2. THE JUMP-DIFFUSION INTEREST RATE MODEL

One of the classical term structure of interest rate models is the Vasicek (1977) model. In the Vasicek model, the instantaneous short-term interest rate r is defined by the following diffusion process called the Ornstein-Uhlenbeck (O-U) process

$$dr(t) = \alpha(\beta - r(t))dt + \sigma dW(t), \tag{1}$$

where α is the mean reversion coefficient, β is the long-term mean of the short-term interest rate, t denotes time path, and σ is the instantaneous volatility of the short-term interest rate. $dW(t)$ is the increment of a standard Wiener process. Let the random variable $r_t \equiv [r(t) \mid r(0) = r_0]$ denote the level of short-term interest rate at time t, conditional on the level of short-term interest rate at the initial time 0, $r(0) = r_0$. Under the O-U process, r_t follows a normal distribution with mean and variance as follows:

$$E(r_t) = e^{-\alpha t} r_0 + \beta(1 - e^{-\alpha t}); \tag{2}$$

$$Var(r_t) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}). \tag{3}$$

The price of a zero-coupon bond at time t, maturing at time T, $P(r,t,T)$ can be determined by a function of the short-term interest rate r and time to maturity $T-t$. That is

$$P(r,t,T) = \exp[-A(t,T)r + B(t,T)], \tag{4}$$

where

$$A(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha};$$

$$B(t,T) = \left(\beta - \frac{\xi\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right)(A(t,T) - (T-t)) - \frac{\sigma^2 A^2(t,T)}{4\alpha^3},$$

and ξ is the market price of interest rate risk, which is assumed to be constant.

To incorporate the discontinuous jump behavior in interest rate dynamics, following Baz and Das (1996), the short-term interest rate r is defined by the extended Vasicek jump-diffusion process

$$dr(t) = \alpha(\beta - r(t))dt + \sigma\, dW(t) + JdN(t), \tag{5}$$

where α is the mean reversion coefficient, β is the long-term mean of the short-term interest rate, t denotes time path, and σ is the instantaneous volatility of the short-term interest rate associated with the diffusion component. $dW(t)$ is the increment of a standard Wiener process, $N(t)$ represents a Poisson process with intensity rate λ. The probability that only one jump happens during the instantaneous period $[t,t+dt]$ is λdt. If there is one jump happening during the period $[t,t+dt]$ then $dN(t)=1$, and $dN(t)=0$ represents no jump happening during that period. J denotes the magnitude of a jump, which is assumed to be a normal variable with mean equal to θ and standard deviation equal to δ. Moreover, $dW(t)$ is assumed to be independent with $dN(t)$, which means that the diffusion component and the jump component of the short-term interest rate are independent of each other. Under the process specified in Eq. (5), as shown in the Appendices, r_t is defined as

$$r_t = e^{-\alpha t}\Big(r_0 + \int_0^t e^{au}\alpha\beta du + \int_0^t e^{\alpha u}\sigma\, dW(u) + \sum_{j=1}^{N(t)} e^{\alpha T_j} J_j\Big), \tag{6}$$

where T_j is the time that the j-th jump happens and $0 < T_1 < T_2 < \cdots < T_{N(t)} < t$, $N(t)$ represents the number of jumps happening during the period between time 0 and time t. It can be shown in the Appendices that the probability density function for r_t is:

$$f(r_t) = \sum_{n=0}^{\infty} \left[\frac{e^{-\lambda t}(\lambda t)^n}{n!} \right] \cdot \int_0^t \int_0^t \cdots \int_0^t \bar{\omega}(r_t; M, S) \cdot \frac{1}{t^n} d\tau_1 d\tau_2 \cdots d\tau_n, \quad (7)$$

where $\omega(r_t; M, S)$ denotes a normal density function with mean M and standard deviation S, and

$$M = e^{-\alpha t} r_0 + \beta(1 - e^{-\alpha t}) + \theta e^{-\alpha t} \sum_{j=1}^{n} e^{\alpha \tau_j};$$

$$S = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) + \delta^2 e^{-2\alpha t} \sum_{j=1}^{n} e^{2\alpha \tau_j}.$$

Also as shown in the Appendices, the mean and variance of the conditional random variable r_t can be obtained as

$$E(r_t) = e^{-\alpha t} r_0 + (\beta + \frac{\lambda\theta}{\alpha})(1 - e^{-\alpha t}); \quad (8)$$

$$Var(r_t) = \frac{\sigma^2 + \lambda(\theta^2 + \delta^2)}{2\alpha}(1 - e^{-2\alpha t}). \quad (9)$$

Assume that the market price of interest rate diffusion risk is constant and equal to ξ, and the jump risk is diversifiable.[1] Under the jump-diffusion interest rate process specified in Eq. (5), according to Baz and Das (1996), using a linearization technique, the price of a zero-coupon bond at time t, which matures at time T, $P(r,t,T)$ can be approximately determined by a function of the short-term interest rate r and time to maturity $T - t$. That is

$$P(r,t,T) = \exp[-A(t,T)r + B(t,T)], \quad (10)$$

where

$$A(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha};$$

$$B(t,T) = \frac{-Ee^{-2\alpha(T-t)}}{4\alpha^3} + \frac{(\alpha D + E)e^{-\alpha(T-t)}}{\alpha^3} + \frac{(2\alpha D + E)(T-t)}{2\alpha^3} - C;$$

$$C = \frac{D}{\alpha^2} + \frac{3E}{4\alpha^3};$$

$$D = \xi\sigma - \alpha\beta - \theta\lambda;$$

$$E = \sigma^2 + (\delta^2 + \theta^2)\lambda.$$

To ensure that bond prices converge to zero for an arbitrarily large maturity, the additional condition $2\alpha D + E < 0$ is necessary. Under the model in Eq. (10), the short-term interest rate r has a linear relationship with the logarithm of discount bond prices, that is

$$r = \frac{-\log P(r,t,T) + B(t,T)}{A(t,T)}. \tag{11}$$

And the yield to maturity of a zero-coupon bond expiring $T - t$ periods hence is given by

$$R(r,t,T) = \frac{A(t,T)r - B(t,T)}{T - t}. \tag{12}$$

The entire term structure of interest rates then can be defined.

Essentially the approximate model is a one-factor term structure model, it has the disadvantage that all bond returns are perfectly correlated, and it may not be adequate to characterize the term structure of interest rates and its changing shape over time. However, the model at least can incorporate jump risks into the term structure model, making the model more complete relative to the pure diffusion Vasicek model. In addition, just as the simple diffusion Vasicek model, the short-term interest rate can move to negative values under the extended jump-diffusion Vasicek model.

The one-factor model can be easily extended to two-factor model with the assumption that the two factors are mutually orthogonal. Assume that the dynamics of two factors y_1 and y_2, that drive the instantaneous short-term interest rate r, follow the following processes

$$dy_1(t) = \alpha_1(\beta_1 - r_1(t))dt + \sigma_1 dW_1(t) + J_1 dN_1(t); \tag{13}$$

$$dy_2(t) = \alpha_2(\beta_2 - r_2(t))dt + \sigma_2 dW_2(t) + J_2 dN_2(t); \tag{14}$$

where $r(t) = y_1(t) + y_2(t)$, $dy_1(t)$ and $dy_2(t)$ are independent with each other. Under the processes, y_1 and y_2 follow a probability distribution as in Eq. (7). Similar to Eq. (10), the bond price under the two-factor model is

$$P(r,t,T) = \exp[-A_1(t,T)y_1 - A_2(t,T)y_2 + B_1(t,T) + B_2(t,T)], \tag{15}$$

where

$$A_i(t,T) = \frac{1 - e_i^{-\alpha_i(T-t)}}{\alpha_i};$$

$$B_i(t,T) = \frac{-E_i e^{-2\alpha_i(T-t)}}{4\alpha_i^3} + \frac{(\alpha_i D_i + E_i)e^{-\alpha_i(T-t)}}{\alpha_i^3} + \frac{(2\alpha_i D_i + E_i)(T-t)}{2\alpha_i^3} - C_i;$$

$$C_i = \frac{D_i}{\alpha_i^2} + \frac{3E_i}{4\alpha_i^3};$$
$$D_i = \xi_i \sigma_i - \alpha_i \beta_i - \theta_i \lambda_i;$$
$$E_i = \sigma_i^2 + (\delta_i^2 + \theta_i^2)\lambda_i;$$
$$i = 1, 2.$$

And the yield to maturity of a zero-coupon bond expiring $T-t$ periods hence is given by

$$R(r,t,T) = \frac{A_1(t,T)y_1 + A_2(t,T)y_2 - B_1(t,T) - B_2(t,T)}{T-t}. \tag{16}$$

And the entire term structure of interest rates can be defined.

3. THE EMPIRICAL METHODOLOGY

In this section, we develop an empirical methodology for the jump-diffusion model. For the one-factor model, assume that there are $T+1$ observations for the state variable (the instantaneous short-term interest rate): $r(0), r(1), r(2), \cdots, r(T)$. Since the jump-diffusion process specified in Equation (5) is Markovian, the conditional likelihood function for the sample is

$$L(r_1, r_2, \cdots, r_T; \Theta) = f[r(1) \mid r(0)] \cdot f[r(2) \mid r(1)] \cdot \cdots \cdot f[r(T) \mid r(T-1)], \tag{17}$$

where $\Theta = (\alpha, \beta, \sigma, \theta, \delta, \lambda)$, which denotes the parameter set to be estimated in the model. And the log-likelihood function for the sample of observations is

$$\log L(r;\Theta) \equiv \log L(r_1, r_2, \cdots, r_T; \Theta) = \sum_{i=1}^{T} \log f[r(i) | r(i-1)]. \tag{18}$$

Since the state variable in the model is the instantaneous short-term interest rate which is unobservable, to develop the maximum likelihood estimator for the parameters of the processes that derive interest rate changes, we develop a likelihood function for the observed bond price as functions of the unobservable state variables. In the jump-diffusion model, according to Eq. (11), the logarithm of the price of a discount bond is a linear function of the state variable, and the change of variable technique can be used to obtain the joint density functions and the log-likelihood function for a sample of observations on discount bond price.

In our estimation, we use the interest rates on four Taiwan commercial papers (10-day, 30-day, 90-day, and 180-day) to estimate the jump-diffusion model. Following Chen and Scott (1993), we add measurement errors as additional random variables in the estimation, in order to perform a change of variables from the unobservable state variables to the observed bond rates. The 10-day commercial paper rate is used and modeled without error because it is one of the most actively traded Treasury bonds, and it is frequently used as an indicator of short-term interest rates. The system of equations for the model estimation is

$$\begin{aligned}
\ln P(t, T_1) &= -A(t, T_1)r(t) + B(t, T_1);\\
\ln P(t, T_2) &= -A(t, T_2)r(t) + B(t, T_2) + e_{1t};\\
\ln P(t, T_3) &= -A(t, T_3)r(t) + B(t, T_3) + e_{2t};\\
\ln P(t, T_4) &= -A(t, T_4)r(t) + B(t, T_4) + e_{3t},
\end{aligned} \tag{19}$$

where $P(t, T)$ is the price of the discount bond with time to maturity equal to $T_i - t$, and $T_1 - t$ is equal to 10 days. e_{1t}, e_{2t}, and e_{3t} are measurement errors. In the estimation, we allow serial correlation and contemporaneous correlation between the measurement errors. The serial correlation is modeled as a first-order autoregressive process

$$e_{jt} = \rho_j e_{j,t-1} + \varepsilon_{jt};\ j = 1, 2, 3, \tag{20}$$

where the innovation of the measurement error is assumed to be normally distributed, that is $\varepsilon_{jt} \sim N(0, \sigma^2(\varepsilon_j))$. Thus measurement errors are assumed to have a joint normal distribution. The log-likelihood function of the estimation then has the following form

$$\begin{aligned}
&\hat{\ell}(\hat{r}, \Theta) = \log \mathrm{L}(\hat{r}_1, \hat{r}_2, \cdots, \hat{r}_T; \Theta)\\
&- T \ln |J| - \frac{3T}{2}\log(2\pi) - \frac{T}{2}\log|\Omega| - \frac{1}{2}\sum \varepsilon_t' \Omega^{-1} \varepsilon_t,
\end{aligned} \tag{21}$$

where $\hat{r}$ is the substitute for the unobservable state variable (the instantaneous short-term interest rate). $\hat{r}$ is estimated, according to Eq. (11), by inverting the observed 10-day commercial paper rate, which is modeled without measurement errors. $\varepsilon_t' = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})$, and Ω is the covariance matrix for ε_t, which is assumed to be a diagonal matrix with elements $\sigma(\varepsilon_1)$, $\sigma(\varepsilon_2)$, and $\sigma(\varepsilon_3)$ along the diagonal. The elements of the matrix J are functions of $A(t, T_i)$, the coefficients in the linear transformation from r to $\log P(t, T_i)$, and the Jacobian of the transformation is $|J^{-1}|$. The likelihood function for the measurement error is conditional on an initial value for e_0.

In calculating the likelihood function for r_t, since the density function in Eq. (7) involves multiple integrals, it is very difficult to proceed. To make the estimation possible, we substitute the true density function by an approximate function.[2] That is

$$f(r_t) = \sum_{n=0}^{\infty} \left[\frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!} \right] \cdot \bar{\omega}(r_t; M', S'), \tag{22}$$

where

$$M' = e^{-\alpha t} r_0 + \beta(1 - e^{-\alpha t}) + \frac{n}{\alpha t}\theta(1 - e^{-\alpha t});$$

$$S' = \frac{\sigma^2}{2\alpha}(1 - e - 2^{\alpha t}) + \frac{n}{2\alpha t}\delta^2(1 - e^{-2\alpha t}).$$

With this approximate density function, when $\alpha t \rightarrow 0$, it converges to the true density function in Eq. (7). Thus either when α or t is small, the approximate density function will provide an appropriate substitute for the true density function.

For the two-factor model, assume that there are $T+1$ observations for the state variables: $y_1(0), y_1(1), y_1(2), \cdots, y_1(T)$ and $y_2(0), y_2(1), y_2(2), \cdots, y_2(T)$. Similar to Eq. (18), the log-likelihood function for the sample of the state variable is

$$\log L(y_i;\Theta_i) \equiv \log L(y_{i1}, y_{i2}, \cdots, y_{iT};\Theta_i) = \sum_{j=1}^{T} \log f[y_i(j) \mid y_i(j-1)], \tag{23}$$

where $\Theta_i = (\alpha_i, \beta_i, \sigma_i, \theta_i, \delta_i, \lambda_i,)$, which denotes the parameter set to be estimated in the model; $i = 1, 2$.

Similar to the case of one-factor model in Eq. (19), we set the system of equations for the 2-factor model estimation is

$$\ln P(t, T_1) = -A_1(t, T_1)r_1(t) - A_2(t, T_1)r_2(t) + B_1(t, T_1) + B_2(t, T_1);$$

$$\ln P(t, T_2) = -A_1(t, T_2)r_1(t) - A_2(t, T_2)r_2(t) + B_1(t, T_2) + B_2(t, T_2) + e_{1t};$$

$$\ln P(t, T_3) = -A_1(t, T_3)r_1(t) - A_2(t, T_3)r_2(t) + B_1(t, T_3) + B_2(t, T_3) - (e_{1t} + e_{2t});$$

$$\ln P(t, T_4) = -A_1(t, T_4)r_1(t) - A_2(t, T_4)r_2(t) + B_1(t, T_4) + B_2(t, T_4) + e_{2t}; \tag{24}$$

where $P(t, T_i)$ is the price of the discount bond with time to maturity equal to $T_i - t$, and $T_1 - t$ is equal to 3 months. e_{1t} and e_{2t} are measurement errors. The unobservable state variables in this two-factor model are computed by inverting $\ln P(t, T_1)$ and $\ln P(t, T_2) + \ln P(t, T_3) + \ln P(t, T_4)$. The measurement errors are also assumed to follow the same distribution as in the one-factor model. The log-likelihood function of the estimation then has the following form

$$\hat{\ell}(\hat{y}_1, \hat{y}_2; \Theta) = \sum_{t=1}^{2} \log L(\hat{y}_{i1}, \hat{y}_{i2}, \cdots, \hat{y}_{iT}; \Theta_i)$$

$$- T \ln | J | - T \log(2\pi) - \frac{T}{2} \log | \Omega | - \frac{1}{2} \sum_{j=1}^{T} \varepsilon_t' \Omega^{-1} \varepsilon_t, \qquad (25)$$

where $\Theta = (\Theta_1, \Theta_2$, which is the parameter set to be estimated in the two-factor model. The estimation procedure is similar to the case of one-factor model as described above.

4. THE DATA AND EMPIRICAL ANALYSIS

In estimating the term structure of interest rates, we used interest rates on 4 Commercial papers: the 10-day, 30-day, 90-day, and 180-day Commercial papers to estimate the term structure of interest rates. Weekly data from July 23, 1983 to September 27, 1997 with a total sample of 741 observations were used. The summary statistics of the sampled interest rates for parameters estimation is in Table 1. During the period from 1983 to 1997, for example, the average weekly 10-day interest rate was about 6.46%, with the maximum and minimum interest rate equal to 16.70% and 2.20% respectively. On average, the longer the time to maturity the higher the interest rate is, revealing that an upward sloping term structure is a normal case. The non-zero skewness and excess kurtosis show that interest rates are not normally distributed. To assess the reliability of the model estimation, we divided the whole sample period into two sub-periods: from July 23, 1983 to December 29, 1990, and from January 5, 1991 to September 27, 1997. According to Table 1, interest rates in the second sub-period are on average higher than those in the first sub-period. While interest rates in the first sub-period are in general less volatile, and exhibit less skewness and leptokurtosis in distribution than those in the second sub-period.

Figure 1 plots the weekly 10-day and 180-day interest rates from July 23, 1983 to September 27, 1997. Normally, the 180-day interest rate is higher than the 10-day interest rate, which reflects an upward sloping term structure. Occasionally, the 180-day interest rate is lower than the 10-day interest rate, reflecting a negatively sloping term structure. For example, during the periods from 1984 to 1986, and in 1997, a negatively sloping term structure of interest rates is observed quite often. Moreover, the 10-day interest rate seems to be more volatile than the 180-day interest rate.

Table 2 shows the results of parameters estimated in the one-factor jump-diffusion model. For the whole sample period, all parameters in the model, except for measurement error terms, are statistically significant. The parameter α is estimated as 1.6358 which is strongly significant,[3] implying that a strong mean reversion is found in the interest rate dynamics. The estimate of 1.6358 for α implies a mean half life of 0.4237 years[4] in the interest rate process. β is estimated as 9.69%, which means that the long-term mean of instantaneous short-term interest rate (the state variable) is about 9.69%. σ is the volatility of the diffusion part in the interest rate process, which is estimated as 5.36%. Thus the instantaneous short-term interest rate is at higher level and is much more volatile than the observed interest rates. The jump intensity parameter λ is estimated as 2.0977, which implies that on average, jumps happen twice a week or so. The average jump magnitude θ is estimated as 0.0138, and its standard deviation δ is 0.0097. The market price of interest rate risk ξ is estimated as a negative value, which is expected. $\sigma(\varepsilon_1), \sigma(\varepsilon_2)$, and $\sigma(\varepsilon_3)$ are the standard

Table 1. Summary Statistics of Sampled Interest Rates for Parameter Estimation.

Maturity	10-day	30-day	90-day	180-day
All: 7/23/1983–9/23/1997; Sample No. 741				
Average	0.0646	0.0690	0.0702	0.0712
Standard Deviation	0.0222	0.0218	0.0220	0.0226
Maximum	0.1670	0.1740	0.1611	0.1550
Minimum	0.0220	0.0238	0.0235	0.0224
Skewness	0.8447	1.0133	0.7728	0.4937
Excess Kurtosis	2.0320	2.6575	1.9606	1.1247
Sub-period 1: 7/23/1983–12/29/1990; Sample No. 389				
Average	0.0614	0.0678	0.0678	0.0680
Standard Deviation	0.0277	0.0275	0.0282	0.0293
Maximum	0.1670	0.1740	0.1611	0.1550
Minimum	0.0220	0.0238	0.0235	0.0224
Skewness	1.0380	1.0234	0.9208	0.7285
Excess Kurtosis	1.1344	1.3054	0.7490	0.0139
Sub-period 2: 1/5/1991–9/27/1997; Sample No. 352				
Average	0.0682	0.0703	0.0729	0.0748
Standard Deviation	0.0127	0.0127	0.0114	0.0103
Maximum	0.1250	0.1135	0.1075	0.1070
Minimum	0.0450	0.0470	0.0510	0.0545
Skewness	0.7385	0.4517	0.0382	–0.0553
Excess Kurtosis	0.8128	–0.1463	–0.8412	–0.6912

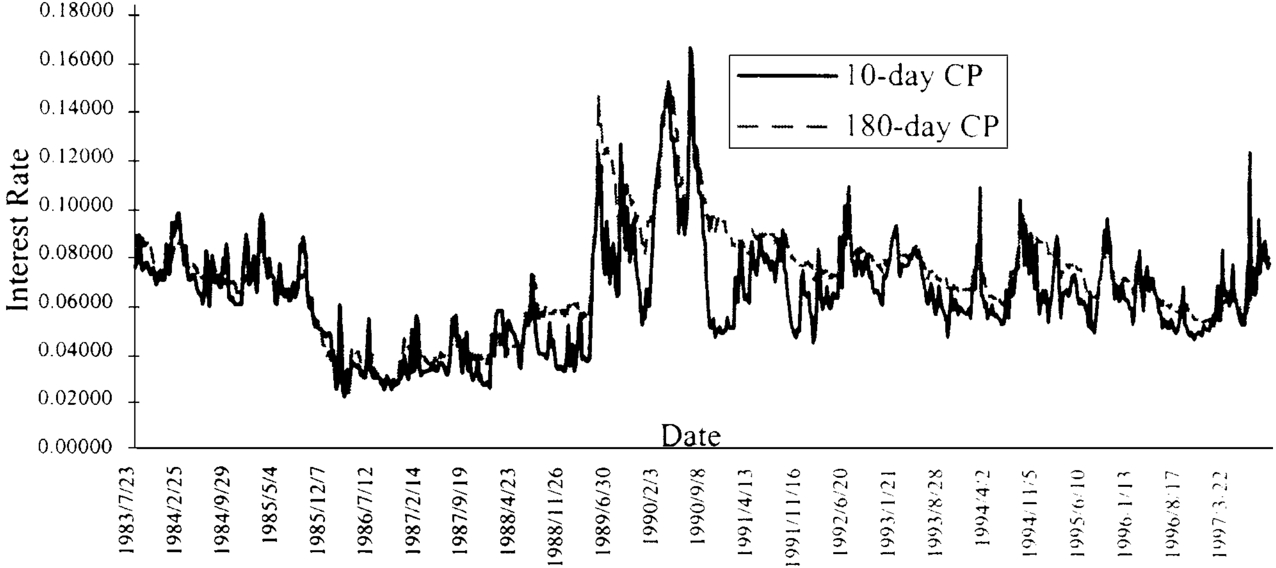


Fig. 1. 10-Day CP Rate Versus 180-Day CP Rate.

Table 2. Parameters Estimated for One-Factor Jump-Diffusion Model.

	All 7/23/1983–9/23/1997			Sub-period 1 7/23/1983–12/29/1990			Sub-period 2 1/5/1991–9/27/1997		
	Coefficient Estimated	Standard Error	T-test	Coefficient Estimated	Standard Error	T-test	Coefficient Estimated	Standard Error	T-test
α	1.6358	0.1746	9.36*	2.1142	0.1010	20.91*	1.7388	0.1853	9.38*
β	0.0969	0.0224	4.32*	0.0786	0.0132	5.93*	0.1110	0.0213	5.21*
σ	0.0536	0.0012	45.01*	0.0342	0.0012	26.72*	0.0475	0.0015	31.08*
λ	2.0977	0.3180	6.59*	1.8695	0.4503	4.15*	2.2215	0.4070	5.45*
θ	0.0138	0.0011	13.03*	0.0197	0.0018	10.78*	0.0145	0.0017	8.22*
δ	0.0097	0.0003	27.13*	0.0119	0.0001	18.29*	0.0129	0.0005	24.99*
ξ	–1.3286	0.5433	–2.44*	–2.2136	0.5272	–4.19*	–1.2584	0.7370	–1.70
$\sigma(\varepsilon_1)$	0.0000	0.0000	1.13	0.0000	0.0000	0.15	0.0000	0.0000	0.30
$\sigma(\varepsilon_2)$	0.0000	0.0000	0.97	0.0000	0.0000	1.14	0.0000	0.0000	0.32
$\sigma(\varepsilon_3)$	0.0000	0.0000	1.18	0.0000	0.0000	0.38	0.0000	0.0000	0.45
ρ_1	–0.0633	0.6273	–0.10	0.4894	0.7204	0.67	–1.2716	1.3387	–0.94
ρ_2	1.0519	0.1345	7.82*	0.9006	0.0454	19.81*	0.9594	0.2870	3.34*
ρ_3	1.6512	0.3771	4.38*	1.1250	0.2786	4.03*	1.6651	0.3828	4.34*

* significant at 1%.

deviation of the innovations in the measurement errors for logarithms of discount bond prices for the 30-day, 90-day, and the 180-day Commercial paper respectively. The standard errors of the model in terms of discount bond price and annualized yield on 30-day, 90-day, and 180-day Commercial papers were also calculated and are shown in Table 4. In terms of price, the standard errors are 0.0611, 0.2398, and 0.6266 respectively, which are correspondingly equal to 6, 24, and 63 basis points. In terms of annualized yield, the standard errors are 0.0073, 0.0096, and 0.0125 respectively, which are correspondingly equal to 73, 96, and 125 basis points. These measurement errors are economically significant implying that the one-factor model does not fit the term structure very well. This result is similar to that of Chen and Scott (1993).

Table 2 also shows the results for the two sub-periods. Although interest rates are more volatile in the first sub-period, the estimated standard deviation of the diffusion component σ is lower. Meanwhile, the jump magnitude parameter θ is estimated higher, although its standard deviation parameter δ and the jump intensity parameter λ are estimated slightly lower compared to those for the second sub-period. This implies that the high volatility of interest rates in the first sub-period is in fact caused by jump risks rather than by diffusion dynamics. Moreover the long-term mean parameter β for the first sub-period is estimated lower compare to that for the second sub-period. This simply reflects the fact that interest rates in the second sub-period are in general higher than those in the first sub-period. In conclusion, estimated parameters for the two sub-periods are comparable other than those reflect the market observations. This confirms the stability and reliability of the model estimation.

Table 3 shows the results of parameters estimated in the two-factor jump-diffusion model. For the whole sample period, all parameters in the model, except for a measurement error term, are statistically significant. For the first factor, the result is quite similar to that in the one-factor model, except that the long-term mean of the factor is 6.56%, which is lower than that in the one-factor model. Moreover the average jump magnitude θ_1 is negative. For the second factor, the parameter α_2 is estimated as 0.1582 implies a mean half life of 4.38 years in the state variable process. The long-term mean β_2 is estimated as 0.62%, σ_2 is 0.57%. Thus the second factor is at low level and much less volatile with weaker mean-reversion characteristics than the first factor. Moreover, the jump intensity parameter for the second factor λ_2 is estimated as 0.1808, which implies that on average, a jump happens every 5.5 weeks or so. The average jump magnitude θ_2 is estimated as 0.0014, and its standard deviation δ_2 is 0.0001, which are also smaller than that of the first factor. Thus

Table 3. Parameters Estimated for Two-Factor Jump-Diffusion Model.

	All 7/23/1983–9/23/1997			Sub-period 1 7/23/1983–12/29/1990			Sub-period 2 1/5/1991–9/27/1997		
	Coefficient Estimated	Standard Error	T-test	Coefficient Estimated	Standard Error	T-test	Coefficient Estimated	Standard Error	T-test
α_1	1.6433	0.0003	5243.02*	1.4990	0.0165	90.72*	1.4886	0.0041	361.14*
β_1	0.0656	0.0011	57.95*	0.0638	0.0202	3.15*	0.0981	0.0044	22.00*
σ_1	0.0532	0.0002	266.48*	0.0357	0.0014	25.69*	0.0896	0.0003	236.33*
λ_1	1.8283	0.0259	70.64*	2.0844	0.2535	8.22*	2.1370	0.1093	19.55*
θ_1	− 0.0085	0.0008	− 10.64*	0.0122	0.0036	3.39*	0.0063	0.0009	6.92*
δ_1	0.0109	0.0028	3.89*	0.0097	0.0077	1.26	0.0116	0.0001	137.54*
ξ_1	− 0.4036	0.0335	− 12.06*	− 1.9457	0.1576	− 12.34*	− 5.0180	0.0737	− 68.01*
α_2	0.1582	0.0002	754.80*	0.1070	0.0003	291.70*	0.1047	0.0000	1340.45*
β_2	0.0062	0.0001	61.07*	0.0167	0.0013	12.23*	0.0130	0.0003	41.23*
σ_2	0.0057	0.0000	3686.44*	0.0154	0.0002	59.01*	0.0141	0.0001	78.80*
λ_2	0.1808	0.0021	84.29*	0.1630	0.0319	5.10*	0.1588	0.0055	28.77*
θ_2	0.0014	0.0000	1607.50*	0.0160	0.0001	135.17*	0.0161	0.0000	595.12*
δ_2	0.0001	0.0000	405.60*	0.0531	0.0001	452.47*	0.0531	0.0000	1908.59*
ξ_2	− 0.0517	0.0030	− 17.12*	− 0.4230	0.0095	− 44.40*	− 0.4081	0.0023	− 174.54*
$\sigma(\varepsilon_1)$	0.0000	0.0000	0.85	0.0062	0.0006	9.84*	0.0021	0.0013	1.60
$\sigma(\varepsilon_2)$	0.0008	0.0000	149.57*	0.0005	0.0000	9.65*	0.0000	0.0001	0.10
ρ_1	0.9997	0.0151	66.04*	0.9541	152.3674	0.00	0.6484	0.3047	2.12*
ρ_2	0.6567	0.0016	406.44*	0.6720	0.9095	0.73	0.8889	0.0974	9.12*

* significant at 1%.

the second factor is less volatile, has weaker mean-reversion, and jumps less frequently than the first factor. $\sigma(\varepsilon_1)$ and $\sigma(\varepsilon_2)$ are the standard deviation of the innovations in the measurement errors for logarithms of discount bond prices for the 30-day and the 180-day Commercial paper respectively. The standard errors of the model in terms of discount bond price and annualized yield on 30-day, 90-day, and 180-day Commercial papers were also calculated and are shown in Table 4. In terms of price, the standard errors are 0.0566, 0.1061, and 0.1455 respectively, which are correspondingly equal to 5, 10, and 14 basis points. In terms of annualized yield, the standard errors are 0.0068, 0.0042, and 0.0029 respectively, which are correspondingly equal to 68, 42, and 29 basis points. These measurement errors are much less economically significant than that of the one-factor model, implying that the two-factor model is superior to the one-factor model in fitting the term structure of interest rates. This result is also consistent with that of Chen and Scott (1993).

As to the estimations of the two-factor model for the two sub-periods, Table 3 shows the results are quite similar to those of the one-factor model. In particular, for the first factor, the estimated standard deviation of the diffusion component σ_1 for the first sub-period is lower, while the jump magnitude parameter θ_1 is estimated higher, with its standard deviation parameter δ_1 and the jump intensity parameter λ_1 estimated slightly lower compared to those for the second sub-period. This means, as in the case of the one-factor model, the high volatility of interest rates in the first sub-period is mainly caused by jumps rather than by diffusion evolutions. Similarly the long-term mean parameter β_1 for the first sub-period is estimated lower compare to that for the second sub-period, which also reflects the market observations. On the other hand, as in the case of the whole sample period, the second factor estimated for the two sub-period, is less volatile, has weaker mean-reversion, and jumps less frequently than the first factor. As a result, estimated parameters for the two sub-periods are comparable other than those reflect the market observations. This further confirms the stability and reliability of the model estimation.

No matter how many factors should be incorporated in modeling the term structure, we need to understand the hedging implications of jumps component. Since the results show that the jump-diffusion model is significant, either for the one-factor model or the two-factor model, with the two-factor model fitting better. We should use a two-factor jump-diffusion model to compute any hedge ratios if hedges need to be implemented. Under such setting, however, we can not obtain faster and easier results when interest rate hedge ratios are computed, because we can not get any option formula with closed-form solution. In other words, the hedge ratio calculations will become more difficult

Table 4. Standard Errors of Term Structure Estimation.

	10-day		30-day		90-day		180-day	
	In price	In yield	In price	In yield	In price	In yield	In price	In yield
				1-Factor Model				
Average Error	0.0000	0.0000	− 0.0257	0.0031	− 0.0248	0.0010	0.0842	− 0.0017
Standard Error	0.0000	0.0000	0.0611	0.0073	0.2398	0.0096	0.6266	0.0125
				2-Factor Model				
Average Error	0.0000	0.0000	− 0.0264	0.0032	− 0.0318	0.0013	0.0568	− 0.0011
Standard Error	0.0000	0.0000	0.0566	0.0068	0.1061	0.0042	0.1455	0.0029

and complex than original diffusion models with closed-form option solution when jump risks are taken into account.

By using the Fourier inversion formula for distributions, Scott (1997) shows that if jumps occurs independently of the interest rate and the stock price, the European stock options with Poisson jump process can be valued numerically. Maybe we can follow such methodology to derive bond price solution analytically if jump components are significant. As a consequence, the relevant hedge ratios can still be computed numerically.

For the explanation of the factors, Fig. 2 plots the 10-day interest rate versus the sum of the two factors (which is associated with the instantaneous short-term interest rate). From Fig. 2, we can find that the two series are highly correlated. Figure 3 plots the 180-day interest versus the second factor. Also the two series are highly correlated. Moreover, we also run regression analysis to investigate the relationship between the observed interest rates and the estimated factors. The result is in Table 5. From Table 5, all coefficients estimated are statistically significant at the level of 1%, implying the two factors are correlated with the observed interest rates. Comparing the coefficients of the two factors, the coefficients of the second factor are equally high for models using the four interest rates as dependent variables. While the coefficients of the first factor are lower when using longer-term interest rates as dependent variables. This implies that the first factor is more associated with the shorter-term interest rates, while the second factor is associated with the general term structure of interest rate. Incidentally, we investigate the relationships between the stock return and the two factors. The results show that the coefficient either for the first factor and the second factor is negative, which is as expected.

5. SUMMARY

Financial variables such as stock prices, foreign exchange rates, and interest rates are always assumed to follow a diffusion process with continuous time paths when pricing financial assets. Despite their attractive statistical properties and computation convenience, more and more empirical evidence shows that pure diffusion models are not appropriate for these financial variables. For interest rates, jump-diffusion processes are particularly meaningful since the interest rate is an important economic variable, which is, to some extent, controlled by the government as an instrument for its financial policy.

Although theoretical derivations for the jump-diffusion term structure models have been developed, the associated empirical work has not been

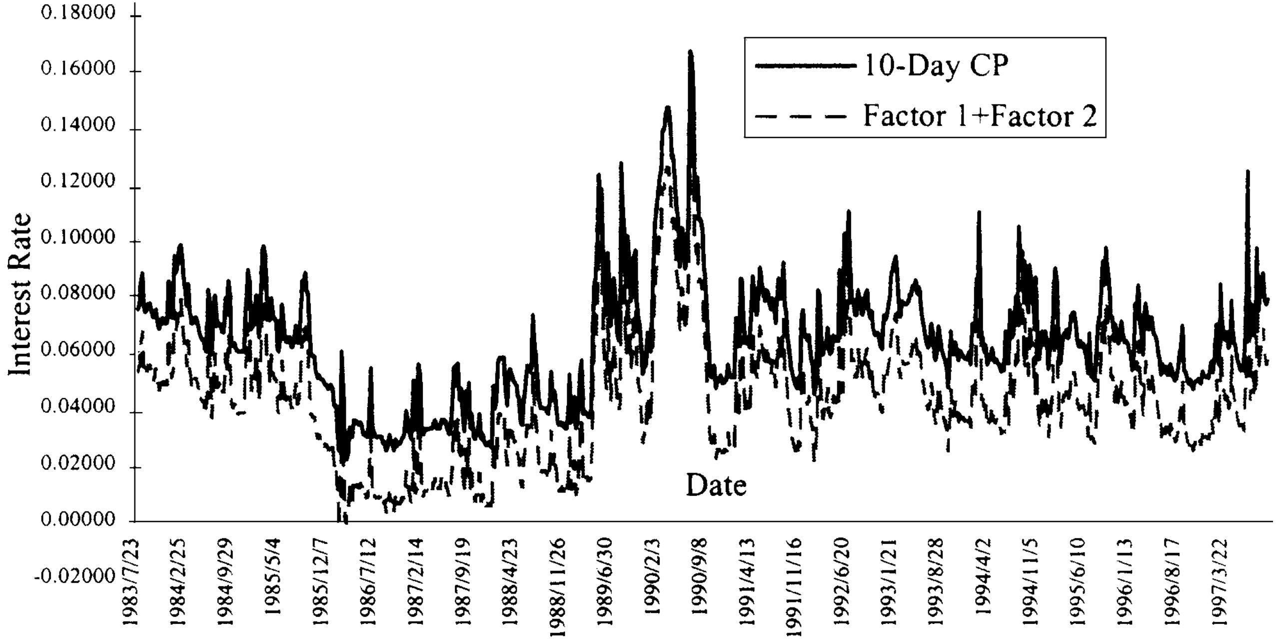


Fig. 2. 10-Day CP Rate Versus Factor 1 + Factor 2.

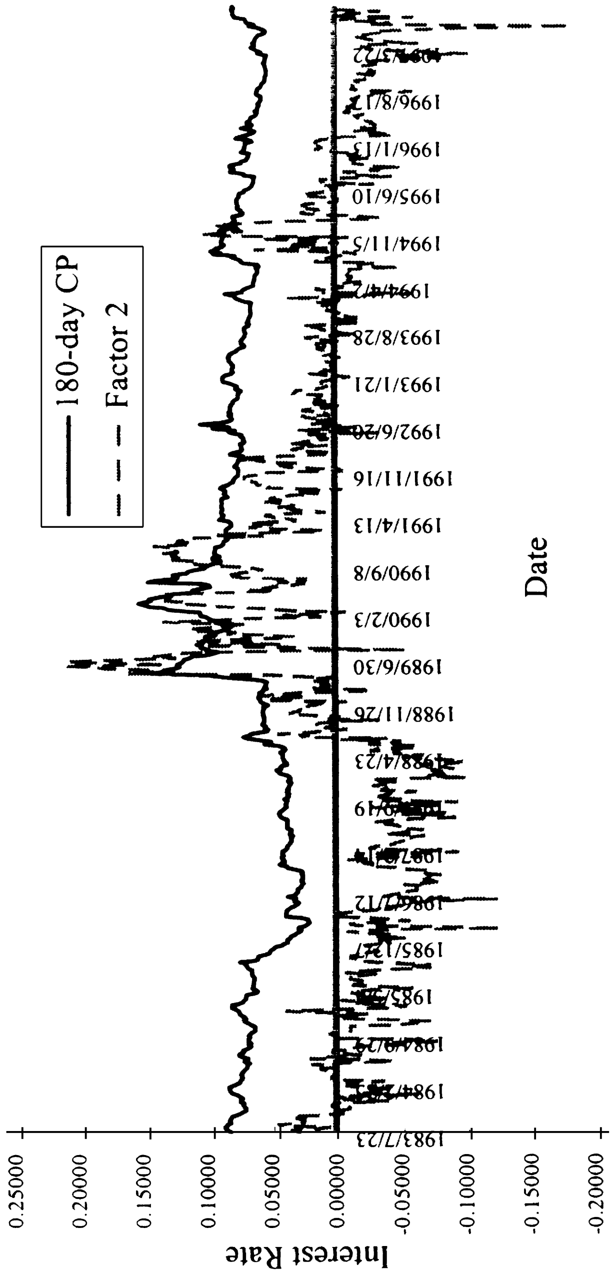


Fig. 3. 180-Day CP Rate Versus Factor 2.

Table 5. Interest Rate Factor Analysis.

Dependent Variable	Regression Coefficient					
	Intercept		Factor 1		Factor 2	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
10-day CP Rate	0.022597	1.01E-06	0.975503	2.12E-05	0.995758	2.07E-05
30-day CP Rate	0.031387	0.000469	0.865993	0.009836	0.951256	0.009611
90-day CP Rate	0.037098	0.000287	0.749967	0.006022	0.944200	0.005884
180-day CP Rate	0.040423	0.000199	0.691713	0.004171	0.937985	0.004075
Stock Return	0.014328	0.003919	− 0.22349	0.082176	− 0.26715	0.080295

found. In this study, we investigated a jump-diffusion process, which is a mixture of an O-U process with mean-reverting characteristics used by Vasicek (1977) and a compound Poisson jump process, for interest rates. Closed-form approximate solutions for discount bond prices were derived by Baz and Das (1996). Essentially the approximate model is a one-factor term structure model. It has the disadvantage that all bond returns are perfectly correlated, and it may not be adequate to characterize the term structure of interest rates and its changing shape over time. However, the model can at least incorporate jump risks into the term structure model, making the model more complete relative to the pure diffusion Vasicek model. In addition, just as the simple diffusion Vasicek model, the short-term interest rate can move to negative values under the extended jump-diffusion Vasicek model.

To overcome the drawback of one-factor model, we extended the Baz and Das (1996) one-factor jump-diffusion model to a multi-factor jump-diffusion model, and developed a methodology for estimating the extended Vasicek jump-diffusion term structure of interest rates model and completed an empirical study for Taiwan money market interest rates. The state variables (such as the instantaneous short-term interest rate, and other factors) that drives the term structure of interest rates dynamics was not observable, and the observed bond prices are functions of the state variable. Thus we needed to use the change of variable technique to obtain the likelihood function in terms of the observed bond prices, in order to conduct a maximum likelihood estimate. The estimation procedure of this study is similar to Chen and Scott (1993).

We use weekly interest rates on the 10-day, 30-day, 90-day, and the 180-day commercial papers to estimate parameters in the one-factor and the two-factor jump-diffusion models. The sample period is from July 23, 1983 to September 27, 1997. The results show that the jump-diffusion model is significant, either for the one-factor model or the two-factor model, with the two-factor model fitting better. This is as expected, since one-factor models do not fit the versatile term structure of interest rates very well. For the two-factor model, compared to the second factor, the first factor exhibits characteristics of stronger mean-reversion, higher volatility, and more frequent jumps in the process. In explanations, the first factor is more associated with shorter-term interest rates, and the second factor is associated with both short-term and longer-term interest rates.

Since the assumption of an appropriate stochastic process for the interest rate and the estimation of its associated parameters are of critical importance when pricing and hedging with term structure of interest rates and interest rate derivatives, the results and the methodology for estimating parameters in the jump-diffusion process have important implications for the area of financial engineering.

NOTES

1. In the case of systematic jump risk, the bond price is given by Eq. (10), with λ replaced by λ^*, the risk-neutral jump intensity rate, and $\lambda^* = \lambda\left(1 - \dfrac{\phi_J \delta^2}{(\theta^2 + \delta^2)B^2/2 - \theta B}\right)$, where ϕ_J is the market price of interest rate jump risk, which is assumed to be proportional to the variance of the jump increment. In empirical estimation, one cannot identify whether λ or λ^* is estimated. Fortunately this is not relevant in pricing term structure derivatives.

2. To obtain the approximate conditional density function for the short-term interest rate r_t, consider the last term of M in Eq. (7), $\sum_{j=1}^{n} e^{\alpha T_j}$, where T_j denotes the time when the j-th jump happens, which is an independent uniform distribution conditional on the number of jumps n in the interval $(0, t)$ is known. Assuming jumps in interest rate spread equally over the time interval $(0, t)$, then the value of $\sum_{j=1}^{n} e^{\alpha T_j}$ is given by

$$E\left[\sum_{j=1}^{n} e^{\alpha T_j}\right] = \sum_{j=1}^{n} E(e^{\alpha T_j}) = n \cdot E(e^{\alpha\tau}) = n \cdot \int_0^t e^{\alpha T} \frac{1}{t}\, dT = \frac{n}{\alpha t}(e^{\alpha t} - 1).$$

Similarly, for the last term of S in Eq. (7)

$$E\left[\sum_{j=1}^{n} e^{2\alpha T_j}\right] = \frac{\mathrm{n}}{2\alpha t}(e^{2\alpha t} - 1).$$

Substituting these into Eq. (7), results in Eq. (18), the approximate density function.

3. A formal statistical test for the hypothesis $\alpha = 0$ is not simple because the parameter value lies on the boundary of the parameter space. The problem of testing the hypothesis $\alpha = 0$ with a *t*-statistic or a likelihood ratio test is similar to the problem of testing for unit roots in time series, and the test statistics do not have the familiar large-sample asymptotic distributions.

4. The mean half life is the expected time for the process to return halfway to its long-term mean β. The mean half life is calculated as $-\ln(0.5)/\alpha$.

REFERENCES

Ahn, C., & Thompson, H. (1988). Jump-Diffusion Processes and the Term Structure of Interest Rates. *Journal of Finance*, *43*, 155–174.

Akgiray, V., & Booth, G. (1986). Stock Price Processes with Discontinuous Time Paths: An Empirical Examination. *Financial Review*, *21*, 163–184.

Akgiray, V., & Booth, G. (1988). Mixed Diffusion-Jump Process Modeling of Exchange Rate Movements. *Review of Economics and Statistics*, *70*, 631–637.

Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. A Wiley-Interscience Publication, John Wiley & Sons.

Ball, C. A., & Torous, W. N. (1983). A Simplified Jump Process for Common Stock Returns. *Journal of Financial and Quantitative Analysis*, March, 53–63.

Ball, C. A., & Torous, W. N. (1985a). On Jumps in Stock Returns. *Journal of Financial and Quantitative Analysis*, *10*, 337–351.

Ball, C. A., & Torous, W. N. (1985b). On Jump in Common Stock Prices and Their Impact on Call Pricing. *Journal of Finance*, *40*, 155–173.

Baz, J., & Das, S. R. (1996). Analytical Approximations of the Term Structure for Jump-Diffusion Processes: A Numerical Analysis. *Journal of Fixed Income*, June, 78–86.

Chen, R. R., & Scott, L. (1993). Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *Journal of Fixed Income*, December, 14–31.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, *51*, 385–408.

Das, S. R. (1994). Poisson-Gaussian Processes and the Bond Markets. Working paper No. 95–034, Harvard Business School.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. New York, John Wiley & Sons.

Hamilton, J. (1988). Rational Expectations Econometrics Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates. *Journal of Economic Dynamics and Control*, *12*, 385–423.

Jarrow, R. A., & Rosenfeld, E. R. (1984). Jump Risks and the Intertemporal Capital Asset Pricing Model. *Journal of Business*, July, 337–351.

Jorion, P. (1988). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies*, *1*(4), 427–445.

Lin, B. H., & Yeh, S. K. (1997). An Empirical Study on Jump-Diffusion Process with Heteroskedasticity in Taiwan Stock Market. Conference paper, APFA Annual Meeting, Tokyo, 1998.

Merton, R. C. (1976). Option Pricing when Underlying Returns are Discontinuous. *Journal of Financial Economics*, *3*, 125–144.
Park, A., Ann, C. M., & Fujihara, R. (1993). Optimal Hedged Portfolios: the Case of Jump-Diffusion Risks. *Journal of International Money and Finance*, *12*, 493–510.
Ross, S. M. (1993). *Introduction to Probability Models*, Fifth Edition. Academic Press, Inc.
Scott, L. (1997). Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods. *Mathematical Finance*, forthcoming.
Tucker, A. L., & Pond, I. (1988). The Probability Distribution of Foreign Exchange Price Changes: Tests of Candidate Processes. *Review of Economics and Statistics*, *70*, 638–647.
Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure of Interest Rates. *Journal of Financial Economics*, *5*, 177–188.

APPENDIX A

Following Arnold (1974), we first assume

$$r_t = \phi(t)\left(r_0 + \int_0^t \phi(u)^{-1}\alpha\beta du + \int_0^t \phi(u)^{-1}\sigma\, dW + \sum_{j=1}^{N(t)} \phi(T_j)^{-1} J_j\right) \quad \text{(A1)}$$

where $\phi(t) = \exp(-\int_0^t \alpha du) = e^{-\alpha t}$. Next we let

$$R_t = r_0 + \int_0^t \phi(u)^{-1}\alpha\beta du + \int_0^t \phi(u)^{-1}\sigma\, dW + \sum_{j=1}^{N(t)} \phi(T_j)^{-1} J_j \quad \text{(A2)}$$

The differential form for Eq. (A2) is

$$dR_t = \phi(u)^{-1}(\alpha\beta dt + \sigma\, dW + JdN)$$

According to Eq. (A1) $r_t = \phi(t)R_t$, thus

$$\begin{aligned} dr_t &= \phi'(t)R_t dt + \phi(t)dr_t \\ &= -\alpha\phi(t)R_t dt + \alpha\beta dt + \sigma\, dW = JdN \\ &= -ar_t dt + abdt + \sigma\, dW + JdN \\ &= \alpha(\beta - r_t)dt + \sigma\, dW + JdN \end{aligned} \quad \text{(A3)}$$

APPENDIX B

Let $X(t)$ be $e^{-\alpha t}(r_0+\int_0^t e^{\alpha u}\alpha\beta du+\int_0^t e^{\alpha u}\,\sigma\, dW(u))$, and $Y(t)$ be $e^{-\alpha t}\sum_{j=1}^{N(t)} e^{\alpha T_j} J_j$, then $r_t=X(t)+Y(t)$. $X(t)$ is independent with $Y(t)$, the distribution of $X(t)$ is Normal distribution with a mean $E[X(t)]=e^{-\alpha t}r_0+\beta(1-e^{-\alpha t})$ and a variance $Var[X(t)]=\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t})$. The probability density function for r_t can be derived as

$$\Pr(r_t \le r)=\Pr(X(t)+Y(t) \le r)$$

$$=\sum_{n=0}^{\infty} \Pr[X(t)+Y(t) \le r \mid N(t)=n]\cdot \Pr[N(t)=n]$$

$$=\sum_{n=0}^{\infty}\left[\frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!}\right]\cdot \Pr(X(t)+Y(t) \le r \mid N(t)=n] \tag{B1}$$

Moreover

$$\Pr[X(t)+Y(t) \le r \mid N(t)=n]$$

$$=\Pr[X(t)+e^{-\alpha t}\sum_{j=1}^{n} e^{\alpha T_j}J_j \le r]$$

$$=\int_0^t\int_0^t\cdots\int_0^t \Pr[X(t)+e^{-\alpha t}\sum_{j=1}^{n} e^{\alpha T_j}J_j \le r \mid T_1=\tau_2, T_2=\tau_2,\cdots,T_n=\tau_n]$$

$$\cdot f_{T_1}(\tau_1)\cdot f_{T_2}(\tau_2)\cdot\cdots\cdot f_{T_n}(\tau_n)d\tau_1 d\tau_2\cdots d\tau_n$$

$$=\int_0^t\int_0^t\cdots\int_0^t \Pr[X(t)+e^{-\alpha t}\sum_{j=1}^{n} e^{\alpha \tau_\xi}J_j \le r]\cdot\frac{1}{t^n}d\tau_1 d\tau_2\cdots d\tau_n \tag{B2}$$

J_j is assumed to be Normally distributed with mean θ and standard deviation δ. It is also assumed to be independent with $X(t)$. Moreover, as specified above, $X(t)$ is also Normally distributed. Thus under the condition that τ_j is known, the variable $X(t)+e^{-\alpha t}\sum_{j=1}^{n} e^{\alpha\tau_j} J_j$ is also Normally distributed. Thus Eq. (B2) can be written as

$$\Pr[X(t)+Y(t) \leq r \mid N(t)=n]$$

$$=\int_0^t \int_0^t \cdots \int_0^t \Omega(M, S)\cdot\frac{1}{t^n}d\tau_1 d\tau_2 \cdots d\tau_n \tag{B3}$$

where $\Omega(M,S)$ denotes the cumulative density function for a Normal distribution with mean M and variance S and

$$M=e^{-\alpha t}r_0+\beta(1-e^{-\alpha t})+\theta e^{-\alpha t}\sum_{j=1}^{n} e^{\alpha\tau_j} \tag{B4}$$

$$S=\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t})+\delta^2 e^{-2\alpha t}\sum_{j=1}^{n} e^{2\alpha\tau_j} \tag{B5}$$

It follows that Eq. (B1) can be expressed as

$$\Pr[r_t \leq r]=\sum_{n=0}^{\infty}\left[\frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!}\right]\cdot\int_0^{\mathrm{t}}\int_0^{\mathrm{t}}\cdots\int_0^{\mathrm{t}}\Omega(r_t; M, S)\cdot\frac{1}{t^n}d\tau_1 d\tau_2\cdots d\tau_n \tag{B6}$$

The probability density function for r_t is

$$f(t_t)=\sum_{n=0}^{\infty}\left[\frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!}\right]\cdot\int_0^{\mathrm{t}}\int_0^{\mathrm{t}}\cdots\int_0^{\mathrm{t}}\bar{\omega}(r_t; M, S)\cdot\frac{1}{t^n}d\tau_1 d\tau_2\cdots d\tau_n \tag{B7}$$

where $\omega(r_t;M, S)$ represents the probability density function for a Normal distribution.

APPENDIX C

According to Ross (1993), the conditional mean and variance for r_t can be derived by

$$E(r_r) = E\{E[r_t \mid N(t) = n]\} = E\{E(E[r_t \mid N(t) = n \mid T_j = \tau_j]) \tag{C1}$$

$$\begin{aligned} Var(r_t) &= E\{Var[r_t \mid N(t) = n]\} + Var\{E[r_t \mid N(t) = n]\} \\ &= E\{E(Var[r_t \mid N(t) = n \mid T_j = \tau_j])\} \\ &+ E\{Var(E[r_t \mid N(t) = n \mid T_j = \tau_j])\} \\ &+ Var\{E(E[r_t \mid N(t) = n \mid T_j = \tau_j])\} \end{aligned} \tag{C2}$$

Based on Eqs (C1) and (C2) and the density function for r_t in Eq. (B7), the result is

$$E(r_t) = e^{-\alpha t} r_0 + (\beta + \frac{\lambda\theta}{\alpha})(1 - e^{-\alpha t}) \tag{C3}$$

$$Var(r_t) = \frac{\sigma^2 + \lambda(\theta^2 + \delta^2)}{2\alpha}(1 - e^{-2\alpha t}) \tag{C4}$$

CROSS HEDGING AND VALUE AT RISK: WHOLESALE ELECTRICITY FORWARD CONTRACTS

Chi-Keung Woo, Ira Horowitz and Khoa Hoang

ABSTRACT

We consider the problem of an electric-power marketer offering a fixed-price forward contract to provide electricity purchased from a fledgling spot electricity market that is unpredictable and potentially volatile. Using a spot-price relationship between two wholesale electricity markets, we show how the marketer may hedge against the spot-price volatility, determine a forward price, assess the probability of making an ex post *profit, compute the contract's expected profit, and calculate the contract's value at risk. Such information is useful to the marketer's decision making. The empirical evidence from highly volatile spot-price data supports our contention that the spot-price relationship is not spurious and can be used for the purpose of risk hedging, pricing, and risk assessment.*

1. INTRODUCTION

Recent regulatory reforms at the federal level have led to a profound restructuring of the U.S. electric power industry (Joskow, 1997; Woo et al., 1997). The passage of the 1978 Public Utilities Regulatory Policy Act sparked the early development of a competitive power-generation industry. The 1992 Energy Policy Act gives the Federal Energy Regulatory Commission (FERC)

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 283–301.

ISBN: 0-7623-0798-6

broad power to mandate open and comparable access to transmission owned by electric utilities. In 1996, the FERC issued Orders 888 and 889 that specify the terms and conditions and the *pro forma* tariff for transmission open access. These reforms foster the development of competitive wholesale electricity markets at major delivery points of the high-voltage transmission grids in the U.S.

Wholesale electricity trading has grown exponentially, dominated by power marketers not affiliated with electric utilities (Seiple, 1996). In addition to spot-market trading, a power marketer may offer fixed-price forward contracts to wholesale buyers such as electric utilities that resell the electricity to their retail customers. Because a forward contract obligates the marketer to deliver electricity at a specific location, it exposes the marketer to volatile spot electricity prices at the market in that location. Such exposure is self-evident when the marketer does not own generation and must make spot-market purchases to meet the delivery obligation. Even when the marketer owns generation, however, the electricity generated has an opportunity cost equal to the spot price. Thus, spot-price volatility, which implies volatility in the marketer's costs, will in either event directly affect a forward contract's profit variance.

As seen from Fig. 1, spot-price volatility may be substantial, with daily spot prices that can, within a few days, spike several hundred folds, from under $50 per megawatt-hour (MWH) to $2,600 per MWH. Even though such price spikes may reflect the operation of a rational market (Michaels & Ellig, 1999), they nonetheless subject the marketer to significant potential loss that can easily result in bankruptcy. For example, a 100 MW forward contract with 16-hour delivery at a fixed price of $50/MWH yields daily revenue of $80,000 (= 100 MW $\times$ 16 hours $\times$ $50/MWH). When the spot price spikes to $2,600/MWH, as it did at the ComEd market in late June 1998, the daily cost rockets to $4,160,000 (= 100 MW $\times$ 16 hours $\times$ $2,600/MWH). The total loss on that fateful day can be a staggering $4.08 million.

This paper considers the problem of determining the financial risks of a forward contract under which a power marketer agrees to provide electricity at a fixed price per MWH for a given delivery period, say one year, that begins sometime after contract signing. The marketer's problem is that although there is a pre-established delivery price, the price or opportunity cost that it will have to pay for that electricity on any given day is subject to the vagaries of highly volatile and unpredictable spot markets. Therefore the marketer's profit on the forward contract is inherently uncertain, and its problem is to quantify the financial risks associated with a given forward-contract price in this uncertain

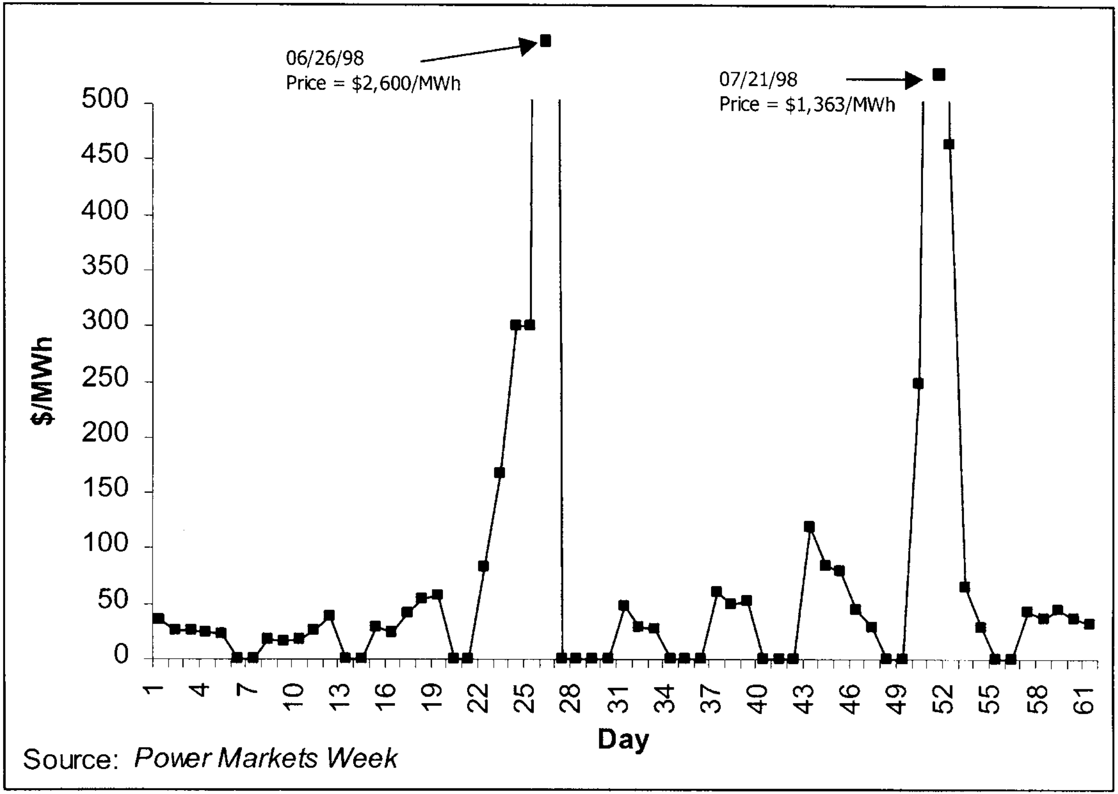


Fig. 1. Daily on-peak spot prices at the ComEd market during 06/01/98–07/31/98.

environment. Knowing such risks will help the marketer to make an informed decision on whether to offer the contract at that forward price.

The solution to the problem posed herein should be of great interest to both practitioners and academics. From the perspective of a practitioner (e.g. a dogmatic power marketer), if a simple cross-hedging strategy is effective and the related financial risk computations are straightforward, they deserve the practitioner's serious consideration for adoption.

From the perspective of an academic, our empirical evidence tests the hypothesis that two geographically close and inter-connected spot markets are integrated and form a single aggregate market. Market integration rationalizes the use of the futures contract for the foreign market to hedge against the large spot-price volatility in the local market. More importantly, almost perfect correlation between the two spot prices implies that a simple hedging strategy is almost 100% effective. Evidence of almost perfect price correlation questions the need for a more complicated and probably more costly strategy involving multiple instruments (e.g. electricity futures with delivery points outside the two spot markets included in the analysis, natural-gas futures, and possibly over-the-counter call options in the local market). To be sure, should the data reject the hypothesis of market integration, this would cast doubt on the common belief of increasing competition in wholesale electricity markets and lead us to conclude that cross hedging is useless in the presence of such market fragmentation.

2. CROSS HEDGING AND RISK ASSESSMENT

Consider a power marketer who wishes to sell a forward contract at a fixed price of $\$G$/MWH in local market "1". There is no forward trading in that local market, although active forward trading takes place in a foreign market "2". The forward price at market "2" is $\$F$/MWH.

To hedge against the local market's spot price volatility, the marketer implements a simple cross-hedging strategy (see, for example, Anderson & Danthine, 1981, for a theoretical exposition) by conducting the following transactions:

- Buy spot electricity on day t at $\$P_{1t}$/MWH in the local market to serve the local forward contract. This results in a day-t per MWH profit of $\pi_{1t} = G - P_{1t}$.
- For each MWH sold under the local market's forward contract, buy β MWH at the fixed price $\$F$/MWH in the foreign market and resell the contract's electricity at the foreign market's day-t spot price of $\$P_{2t}$/MWH. This results in a day-t per MWH profit of $\pi_{2t} = \beta\ (P_{2t} - F)$.

These two transactions lead to a combined per MWH profit on day t equal to

$$\pi_t = \pi_{1t} + \pi_{2t} = (G - P_{1t}) + \beta\,(P_{2t} - F). \tag{1}$$

The marketer's two-pronged problem is to determine both G and β, given the inherent volatility in the two spot prices. What is apparent from Eq. (1), however, is that once β has been determined, the variance in profit, σ_π^2, will be a function of both the price variances in the respective markets, σ_1^2 and σ_2^2, and the covariance between the two spot-market prices, σ_{12}. Specifically:

$$\sigma_\pi^2 = \sigma_1^2 + \beta^2\sigma_2^2 - 2\beta\sigma_{12}. \tag{2}$$

Hence, the profit variance will be lower than otherwise, *ceteris paribus*, when the spot prices are positively correlated. As reported for selective markets in Woo et al. (1997), this is indeed often the case for adjacent spot electricity markets. Suppose, then, that an increase in P_{2t} is on average accompanied by a rise in P_{1t}. The incremental loss in π_{1t} is ΔP_{1t} and the incremental gain in π_{2t} is $\beta\Delta P_{2t}$. If $\Delta P_{1t} = \beta\Delta P_{2t}$, the combined profit π_t would on average remain unchanged.

The success of the marketer's cross-hedging strategy assumes a spot-price relationship between the local and foreign markets. Let α denote a parameter, and let ε_t denote a day-t random-error term with the usual IID independence and normality properties. In particular, with E the expectations operator, $\mathrm{E}[\varepsilon_t] = 0$ and $\mathrm{E}[\varepsilon_t^2] = \sigma_\varepsilon^2 \geq 0$. Suppose the spot-price relationship remains unchanged over the local forward contract's delivery period and that it is described by the linear regression equation

$$P_{1t} = \alpha + \beta P_{2t} + \varepsilon_t. \tag{3}$$

Here, P_{1t} is the local market's average of hourly spot prices for the M delivery hours on day t, and P_{2t} is the foreign market's average of hourly spot price for the same M delivery hours on that day. It is immediately verified that $\Delta P_{1t} = \beta\Delta P_{2t}$. Hence, on average the spot-price fluctuations in the two markets will cancel each other out insofar as their impact on the marketer's total profits is concerned, provided that its purchases in the foreign market are set equal to β, the regression's slope parameter. Then, substituting Eq. (3) into Eq. (1) yields:

$$\pi_t = G - \alpha - \beta F - \varepsilon_t \tag{4}$$

As reflected in Eq. (4), because of the linear relationship between the daily spot prices in the two markets, over the local forward contract's delivery period, daily profit does not depend on those spot prices. Moreover, as the first three terms on the right-hand side of Eq. (4) are fixed, π_t varies only with the locational basis risk ε_t. Arising as it does from random error, this basis risk cannot be removed by cross hedging. Thus β is the optimal hedge ratio that minimizes the variance of π_t (Siegel & Siegel, 1990, Chapter 3). With the parameter vector $\boldsymbol{V} = (\alpha, \beta)$ known, the variance of π_t is now given by $\sigma_\pi^2 = \sigma_\varepsilon^2$, as opposed to the formulation of Eq. (2).

In practice, however, the vector $\boldsymbol{V}$ is not known. Nonetheless, the marketer does have two clear alternatives for dealing with this aspect of the problem. In the first of these, the marketer applies ordinary least squares (OLS) to estimate Eq. (3). With e_t denoting the day-t residual from a sample of n historical observations, and $\boldsymbol{V}_e = (a, b)$ the OLS estimate of $\boldsymbol{V}$, the estimated Eq. is:

$$P_{1t} = a + bP_{2t} + e_t \qquad ')$$

Then, the marketer behaves *as if* Eq. (3′) is the true regression equation. The variance of the residuals, $s_e^2 = \sum e_t^2/(n-1)$ with n being the size of the sample used to estimate equation (3′), is used as the unbiased estimator of σ_ε^2 in the above machinations and through the subsequent analysis (Woo et al., 2001).

Alternatively, in Bayesian fashion the uncertain $\boldsymbol{V}$ is considered to be a random vector. The marketer is then asked to assess a multi-normal density over $\boldsymbol{V}$. As Berger (1985, p. 168) remarks: "The most natural focus is to use the data to estimate the prior distribution or the posterior distribution". That distribution can then be revised in typical Bayesian fashion as additional sample information is received in the form of further observations on the prices in the two markets (Raiffa & Schlaifer, 1961, Chapter 12). Under this second, empirical Bayes, approach the marketer's assessed prior distribution over $\boldsymbol{V}$ has a mean of $\boldsymbol{V}_e = E[\boldsymbol{V}]$, and the prior variances of the two parameters, σ_α^2 and σ_β^2, are assigned by using the respective standard errors from the estimated regression, s_a^2 and s_b^2. The OLS covariance, s_{ab}, is used as the prior covariance, $\sigma_{\alpha\beta}$. In the most readily applied version of the subsequent procedure, the variance of the residuals, s_e^2, is taken to be the known population variance of the random-error term (Raiffa & Schlaifer, 1961, pp. 310–312).

From Eq. (4), the expected daily per MWH profit is computed to be:

$$\mathrm{E}[\pi_t] = G - \mathrm{E}[\alpha] - \mathrm{E}[\beta]F - \mathrm{E}[\varepsilon_t] = G - a - bF. \tag{5a}$$

The daily per MWH profit variance is now computed as:

$$s_\pi^2 = s_a^2 + F^2 s_b{}^2 + 2Fs_{ab} + s_e^2. \tag{5b}$$

There is a higher profit variance in this mode of analysis than the $s_\pi^2 = s_e^2$ that would be used in the first mode. The increased variance stems from the fact that in the first mode the marketer has, in effect, assigned a dogmatic prior over $\boldsymbol{V}$. The marketer's dogmatism is reflected in its behaving as if the estimated regression coefficients are indeed the true regression coefficients. That behavior results in the marketer's setting the standard errors of those coefficients, and the covariance between them, equal to zero.

As the linear combination of normal densities, the daily per MWH profit is also normally distributed:

$$\pi_t \sim \mathrm{N}(G - a - bF,\ s_a^2 + F^2 s_b^2 + 2F s_{ab} + s_e^2). \tag{5c}$$

We may therefore directly compute the probability of making a per MWH daily profit above a predetermined threshold π_T (say, zero):

$$\mathrm{prob}(\pi_t > \pi_T) = \mathrm{prob}\{(\pi_t - \mathrm{E}[\pi_t])/s_\pi > (\pi_T - \mathrm{E}[\pi_t])/s_\pi\}. \tag{6}$$

We can use Eq. (6) in combination with Eq. (5a) to solve for a particular value for G that would yield a target probability of positive profit. Suppose G is initially set at $(a + bF)$. The latter results in an expected profit of zero, with a variance of s_π^2 computed from Eq. (5b). At $G = a + bF$, $\mathrm{prob}(\pi_t > 0) = 0.5$. To improve $\mathrm{prob}(\pi_t > 0)$ to a target of $p_T > 0.5$, we increase G by zs_π, where z is the standard normal variate corresponding to p_T. In particular, $z = 1.65$ if $p_T = 0.95$. Thus the pricing rule for positive profit with this degree of certainty is:

$$G_{0.95} = a + bF + 1.65 s_\pi. \tag{7}$$

Thus far we have been dealing with the expected value and volatility of the daily per MWH profit in connection to the two forward prices (G, F) and the estimated inter-market spot-price relationship. But the local forward contract may have more than one delivery day. Suppose the contract has L delivery days. The forward contract is profitable if $\mu = \sum \pi_t / L$, the average of the daily per MWH profits over the L delivery days, is positive. As π_t is normally distributed, so is μ, with $\mathrm{E}[\mu] = G - a - bF$ and $s_\mu^2 = s_\pi^2 / L$; or,

$$\mu \sim \mathrm{N}(G - a - bF,\ (s_a^2 + F^2 s_b^2 + 2F s_{ab} + s_e^2)/L).$$

Proceeding in a manner similar to the case of the daily per MWH profit, the marketer can derive the probability that the forward contract with L delivery days will be *ex post* profitable. Letting μ_T denote the threshold for the average of daily per MWH profits:

$$\mathrm{prob}(\mu > \mu_T) = \mathrm{prob}\{(\mu - \mathrm{E}[\mu])/s_\mu > (\mu_T - \mathrm{E}[\mu])/s_\mu\}.$$

We recognize that besides the number of delivery days, the marketer's total risk exposure critically depends on the megawatt (MW) size of the local forward contract. Suppose the contract's total electricity delivery at the 100% rate is $Q = KLM$, where K = MW size, L = number of delivery days in the contract period, and M = number of delivery hours per delivery day. The contract's expected total profit is:

$$\Pi = Q\mathrm{E}[\mu] = Q(G - a - bF).$$

The total profit variance is:

$$s_{\Pi}^2 = Q^2 s_{\mu}^2.$$

Following Jorion (1997), we use prob($\Pi <$ VAR) = 0.05 to derive the contract's value at risk:

$$\mathrm{VAR} = Q(G - a - bF - 1.65 s_{\mu}) \tag{8}$$

Equation (8) states the marketer's maximum loss under normal circumstances. In other words, there is less than a 5% chance that the marketer's *ex post* loss after the realization of market spot prices will exceed the VAR in Eq. (8).

The profitable price $G_{0.95}$ in Eq. (7) and the VAR in Eq. (8) are closely related. The relationship is that if $G = G_{0.95}$, then VAR = 0. Should G be above (below) $G_{0.95}$, VAR is positive (negative). This shows how the marketer can systematically revise G to improve the probability of profit and to reduce the VAR.[1]

3. AN ILLUSTRATION

We use two different pairs of geographically separated real-world markets to illustrate our approach. The first pair, in the Mid-West, comprises Commonwealth Edison (ComEd) in the role of market "1", the local market with spot trading only, and Cinergy in the role of market "2", the foreign market with both spot and futures trading. The second pair, in the Pacific Northwest, comprises Mid-Columbia (Mid-C) in the role of the local market and the California-Oregon border (COB) in the role of the foreign market. For each pair we retain the assumption that the power marketer would sell the forward contact in the local market and cross hedge via the purchase of a strip contract (a series of 12 monthly forward contracts) with delivery in the foreign market. The local forward contract is for 16-hour on-peak (06:00–22:00, Monday to Friday, except public holidays) delivery throughout a local contract period that is assumed to be one year. Computations for a local contract with a shorter time period are entirely analogous.

The first pair, ComEd/Cinergy, represents emerging markets that one would expect to be associated with larger basis risk than would be the case with the more mature Mid-C/COB tandem. The former pair has tended to have the more volatile and higher spot prices than the latter pair. In particular, the ComEd/

Cinergy pair has proved to be very vulnerable to such unanticipated events as the hot weather and plant outages in June and July of 1998. We would therefore expect cross hedging to be less effective for that pair than for Mid-C/COB, and consequently we would expect more risks for the power marketer in ComEd versus its Mid-C counterpart.

3.1. Estimation

The presumptive first stage in our procedure is to obtain the estimated spot-price regression, Eq. (3′) and the associated standard errors, using the daily spot-price data for each market pair. The resulting estimates, however, can be misleading and subject to spurious interpretation if the price series (P_{1t}, P_{2t}) are random walks (e.g. $P_{it} = P_{i(t-1)} + \text{error}$, $i = 1, 2$) that may drift apart over time. Such drifting can cause the basis risk as reflected in the residuals to grow over time, rendering the cross hedge useless. To guard against this possibility, prior to relying on the estimated parameters from equation (3′) we first test the null hypotheses that (P_{1t}, P_{2t}) are random walks and that (P_{1t}, P_{2t}) drift apart over time.

The test statistic for the random-walk hypothesis is the Augmented Dickey-Fuller (ADF) statistic whose critical value for testing at the 1% level is -3.4355. The ADF statistic is the *t*-statistic for the OLS estimate of δ in the regression $\Delta v_t = \delta v_{t-1} + \phi \Delta v_{t-1} + \text{white noise}$. The variable v_t is the residual in the estimated regression $P_{it} = \gamma_0 + \gamma_1 P_{i(t-1)} + \eta_t$, where η_t is a random-error term with the usual normality properties. This test, then, is a unit-root test.

Table 1 reports some summary statistics of daily on-peak spot prices, as well as the ADF statistics, for all four markets for our sample period of June 1, 1998 to May 31, 1999. The sample period is so chosen as to highlight how cross hedging may be used in a highly uncertain market-price environment.

The summary statistics indicate that the spot-price distributions are skewed to the right, with the median price below the mean price. The non-normal price distribution, however, should not prevent our empirical implementation. First, least-squares regression yields unbiased coefficient estimates. Second, such measurements of risk as profit volatility, probability of positive profit and value at risk are based on the multi-normal prior density. And, finally, the average of the daily per MWH profits is normally distributed under the Central Limit Theorem.

The standard deviations in Table 1 reflect the fact that Mid-C and COB spot prices are generally lower and less volatile than are the ComEd and Cinergy spot prices, which fuels our expectation of lower $G_{0.95}$ and VAR for the Mid-C market versus the ComEd market. The ADF statistics indicate that none of the

Table 1. Summary Statistics and ADF Statistics of Daily On-peak Spot Prices (U.S.$/MWH) by Market for the Sample Period of 06/01/98–05/31/99.

Statistics	COB	Mid-C	Cinergy	ComEd
Sample size	256	256	256	250
Mean	29.31	27.61	55.74	47.94
Minimum	8.15	7.42	14.15	15.50
First quartile	21.51	18.75	18.51	19.50
Median	26.89	25.24	23.31	23.51
Third quartile	31.56	30.73	29.25	28.54
Maximum	85.91	87.65	2040.48	2600.00
Standard deviation	12.87	14.19	201.79	187.45
ADF statistic for testing H_0: The price series is a random walk	− 4.26*	− 9.25*	− 4.58*	− 9.35*

Notes:
(a) The ComEd market only has 250 observations because of missing price data.
(b) The ADF statistic's critical value is − 3.4355 at the 1% level.
(c) "*" = "Significant at the 1% level and the null hypothesis is rejected."
(d) Source: Power Markets Week.

four price series is a random walk, implying that we need not discredit any estimated spot-price relationship on random-walk grounds.

To understand the behavior of the so-called regional basis (i.e. the difference between two locational spot prices), we plot the daily spot prices for each market pair in Figs. 2.A and 2.B. These figures indicate that the regional basis widens during high- price days. Our review of the events during those days suggests that the size of the regional basis increases with the extent of transmission congestion that prevents the unfettered flow of power between the two inter-connected markets.

We apply the cointegration test to test the hypothesis that the (P_{1t}, P_{2t}) pair drift apart over time. The test statistic is an ADF statistic, with − 3.9001 being the critical test value at the 1% level.[2] The estimation results for both market pairs, as well as the ADF statistics for the two regressions, are given in Table 2. The ADF statistics indicate that we can safely reject the hypothesis that the two spot-price series in either market pair drift apart. The statistics lead us to infer that both market pairs comprise substantially, if imperfectly, related markets.

Moreover, the markets within each pair would appear to be integrated such that they are slightly differentiated markets within a larger overall market.

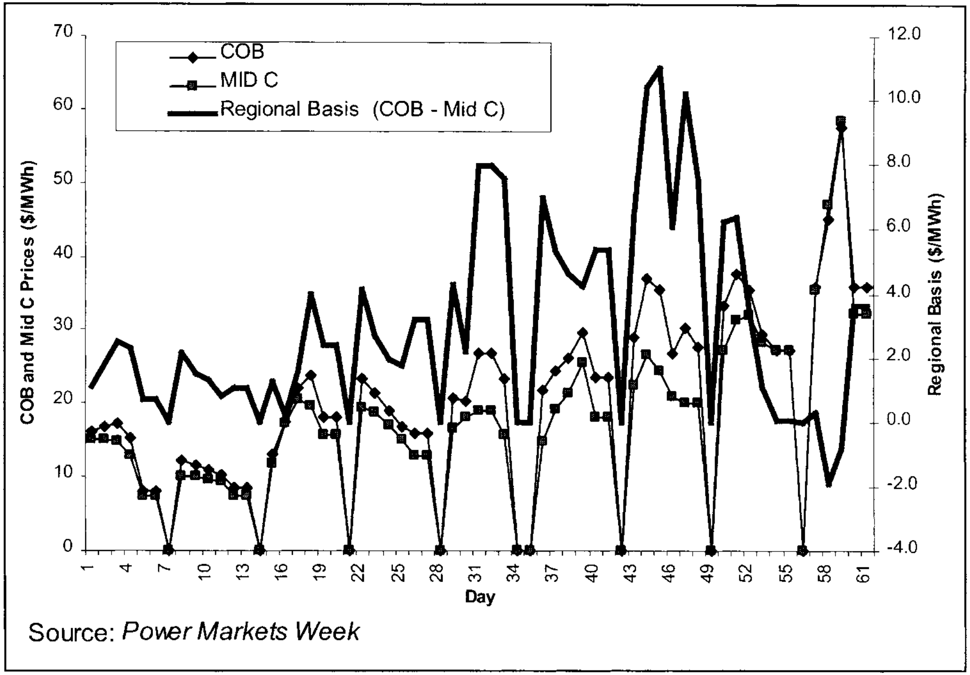
Table 2. On-peak Spot-Price Regressions by Market Pair.

Independent variables	Dependent variable Mid-C price	Dependent variable ComEd price
Intercept	− 4.097 (− 9.64)*	− 5.01 (− 1.94)
COB price	1.082 (81.43)*	
Cinergy price		1.093 (73.87)*
Adjusted R^2	0.9630	0.9563
Mean squared error	7.46	1533.8
ADF statistic for testing H_0: The two price series drift apart without limit	− 6.92*	− 11.53*

Notes:
(a) t-statistics in ( ) and "*" = "Significant at 1% and the null hypothesis that the coefficient equal to zero is rejected".
(b) The ADF statistic's critical value is − 3.9001 at the 1% level.

Evidence of market integration would support the use of the futures contract for the foreign market to hedge against the price volatility in the local market. The first evidence of market integration is that the value for b is 1.082 for the Mid-C/COB estimated regression and is 1.093 for the ComEd/Cinergy regression, implying that a $1 price change in either foreign market is accompanied by an approximately equal change in its companion local market. Consistent with the law of one price, absent persistent congestion that hinders inter-market trading, the price in one market should, on average, equal the price in an adjacent market plus the cost of transmission. The second piece of evidence is that the values for a of − 4.097 for the Mid-C/COB estimated regression and of − 5.010 for the ComEd/Cinergy regression, do in fact approximate the average costs of transmission between those market pairs.

Despite the evidence of market integration, the simple cross hedging described in Section 2 might be ineffective if the ties between the two markets are relatively loose. In the present instance, however, the adjusted-R^2 values for both regressions exceed 0.95. As discussed and described in detail in Woo et al. (1997), the latter confirm the tight ties that exist between these market pairs. The marketer can therefore be confident of nearly 100% hedge effectiveness. Thus these two market pairs do not call for a more complicated and probably more costly strategy involving multiple instruments (e.g. electricity futures with delivery points outside the two spot markets included in the analysis,

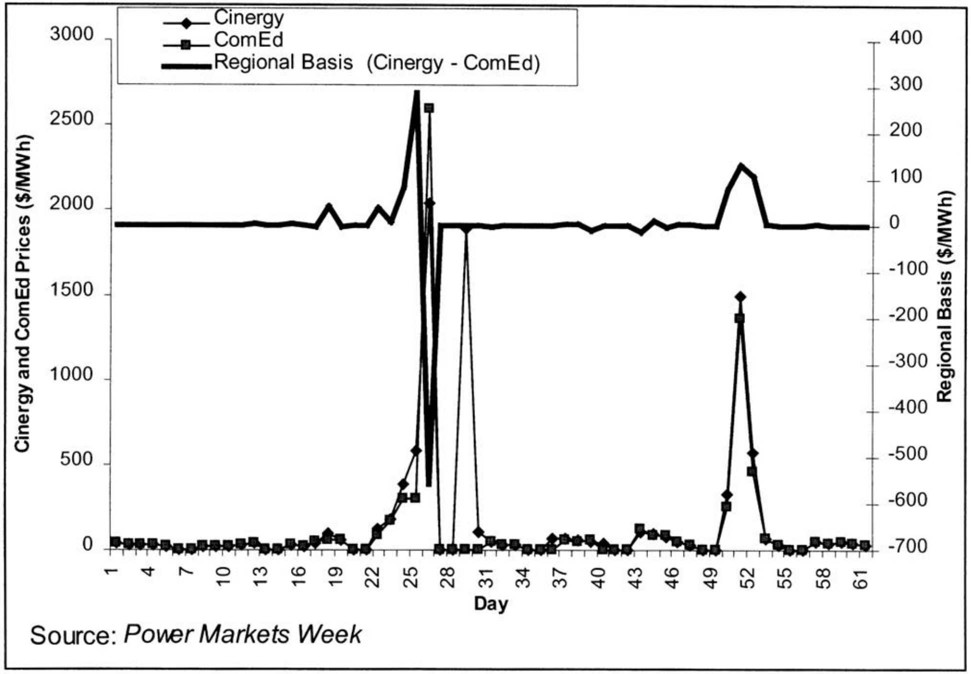


Note: Spot prices and regional basis for Mid-C/COB market pair during the 30 Days (06/01/98 – 07/31/98) with the highest regional basis, sorted in ascending order by the size of regional basis.

Fig. 2A. Regional Basis and Spot Prices for the Mid-C/COB Market Pair During 06/01/98–07/31/98.

Fig. 2*A*. Continued.

Day [A]	Date [B]	COB [C]	Mid-Columbia [D]	Regional Basis [E] = Abs([C] − [D])	Day [A]	Date [B]	COB [C]	Mid-Columbia [D]	Regional Basis Abs ([C] − [D])
1	07/15/98	36	25	11	16	07/08/98	26	21	5
2	07/14/98	37	27	10	17	06/29/98	21	17	4
3	07/17/98	30	20	10	18	07/09/98	30	25	4
4	07/01/98	27	19	8	19	06/22/98	23	19	4
5	07/02/98	27	19	8	20	06/18/98	24	20	4
6	07/03/98	23	16	8	21	07/30/98	36	32	4
7	07/18/98	28	20	8	22	07/31/98	36	32	4
8	07/06/98	22	15	7	23	07/22/98	36	32	3
9	07/13/98	29	22	6	24	06/26/98	16	13	3
10	07/21/98	38	31	6	25	06/27/98	16	13	3
11	07/20/98	33	27	6	26	06/23/98	21	19	3
12	07/16/98	27	21	6	27	06/03/98	17	15	2
13	07/10/98	24	18	5	28	06/19/98	18	16	2
14	07/11/98	24	18	5	29	06/20/98	18	16	2
15	07/07/98	25	19	5	30	06/04/98	15	13	2



Note: Spot prices and regional Basis for ComEd/Cinergy market pair during the 30 Days (06/01/98 – 07/31/98) with the highest regional basis, sorted in ascending order by the size of regional Basis.

Fig. 2B. Regional Basis and Spot Prices for the ComEd/Cinergy Market Pair During 06/01/98–07/31/98.

Fig. 2*B*. Continued.

Day [A]	Date [B]	Cinergy [C]	ComEd [D]	Regional Basis Abs ([C] − [D])	Day [A]	Date [B]	Cinergy [C]	ComEd [D]	Regional Basis Abs ([C] − [D])
1	06/26/98	2040	2600	560	16	07/16/98	49	45	4
2	06/25/98	588	300	288	17	06/19/98	54	58	3
3	07/21/98	1493	1363	130	18	07/07/98	64	61	3
4	07/22/98	572	465	107	19	07/01/98	45	48	3
5	06/24/98	383	300	83	20	06/17/98	39	42	3
6	07/20/98	330	250	80	21	07/27/98	46	44	3
7	06/18/98	96	55	41	22	07/17/98	32	29	2
8	06/22/98	121	83	38	23	06/12/98	40	38	2
9	07/13/98	106	120	14	24	06/15/98	31	29	2
10	07/14/98	97	85	12	25	07/29/98	47	45	2
11	07/09/98	44	52	9	26	07/30/98	38	36	2
12	06/23/98	175	168	7	27	07/28/98	39	37	1
13	07/08/98	55	50	5	28	06/02/98	27	26	1
14	07/15/98	76	80	4	29	06/05/98	21	22	1
15	07/23/98	69	65	4	30	06/11/98	26	25	1

natural-gas futures, and possibly over-the-counter call options in the local market). While this finding is specific to the Mid-C/COB and ComEd/Cinergy market pairs, it does indicate the usefulness of the simple strategy.

In light of the evidence of market integration and the goodness of fit of both estimated regressions, a power marketer may now feel comfortable in using the regression results to assign its non-dogmatic prior densities and to thence implement cross hedging and assess the risks inherent in the local contract.

3.2. Empirical Results

By altering the price of the local contract, we develop alternative estimates for the contract's expected total profit, the probability of a positive profit, and VAR. For the sake of illustration, we make the following assumptions:

- The local forward contract's size is 100 MW, enough to serve approximately 30,000 residential homes. The one-year contract has 256 delivery days and 16 delivery hours per day. Thus the contract's total electricity delivery is 409,600 MWH.
- The 12-month forward price with delivery beginning 08/99 is \$35/MWH at the COB market and \$39/MWH at the Cinergy market (Source: Settlement prices on 07/06/99 based on the information on www.powermarketers.com).

Table 3.A shows that when the Mid-C contract is priced at or below \$33/MWH, the contract is not likely (less than a 1% chance) to be profitable. At \$30/MWH, the expected total loss can be as large as \$1.5 million and the VAR is \$1.6 million. At \$33.75/MWH, the Mid-C contract's expected profit is zero and the VAR is $-$\$0.115 million. But if the Mid-C contract price can be raised to \$34/MWH, the probability of positive profit rises to $p^+ > 0.90$ and the VAR dwindles to \$0.016 million. At $G_{0.95} = \$34.04$, the probability of positive profit is $p^+ = 0.95$ and the VAR becomes zero. Adding \$1/MWH to $G_{0.95}$ ensures the contract's profitability with almost certainty ($p^+ > 0.99$).

Table 3.B shows that when the ComEd contract is priced below \$37/MWH, the contract's probability of positive profit is $p^+ < 0.40$. At \$36/MWH, the expected total loss is \$0.674 million but the VAR is large at \$2.33 million. At \$37.64/MWH, the ComEd contract is expected to break even but it still has a VAR equal to \$1.65 million, over 13 times the $-$\$0.115 million VAR for the Mid-C contract at the break-even price of \$34.76/MWH. If the ComEd contract price can be raised to \$41.69/MWH, the probability of positive profit rises to over 0.95 and the VAR diminishes to \$0.00. To make the ComEd contract profitable with almost certainty ($p^+ > 0.99$), the power marketer would need to raise the price by another \$2.4/MWH to \$44/MWH. Thus the ComEd contract

Table 3A. Forward Price, Probability of Positive Profit, Expected Total Profit, and Value at Risk of a 1-year, 100-MW Forward Contract with 16-hour On-peak Delivery at Mid-columbia.

Mid-C forward price ($/MWH)	Probability of positive profit (%)	Expected total profit ($000)	Value at risk with 5% chance ($000)
30	0.0	− 1,539	− 1,655
31	0.0	− 1,130	− 1,246
32	0.0	− 720	− 836
33	0.0	− 310	− 426
33.76	50	0	− 115
34	92	98.8	− 16
34.05	**95**	**121**	**0**
35	Over 99	508	393
36	Over 99	918	802

Note: The price for a 12-month forward contract with COB delivery is assumed to be $35/MWH. At this COB price, the variance of the average of daily per MWH profit is $s_\mu^2 = \$0.032$/MWH. The Mid-C price in **bold** is $G_{0.95}$ that achieves a 95% probability of positive profit and a zero VAR.

is more risky than the Mid-C contract because the ComEd/Cinergy market pair has more volatile and higher spot prices than does the Mid-C/COB pair.

Having determined the alternative forward price levels at the Mid-C and ComEd markets, it might have been interesting to compare them with actual forward prices observed in these markets. Such a comparison would shed light on the price-making behavior and risk aversion of power marketers. For instance, if the actual forward prices tend to be at the low end of those in Tables 3A and 3B, we can infer that power marketers price aggressively and are not risk averse. Unfortunately, forward prices for Mid-C and ComEd are unavailable; otherwise, cross hedging would have become unnecessary.

4. CONCLUSION

Volatile spot prices and incomplete forward markets motivate our consideration of an electric-power marketer who sells a fixed-price contract to a wholesale buyer. The marketer implements cross hedging to reduce the risk of the fixed-price contract with delivery to a local market that does not have forward trading. Using a spot-price relationship between two wholesale electricity markets, we show how the marketer may hedge against the spot-price volatility, determine a forward price, assess the probability of making an *ex post* profit,

Table 3B. Forward Price, Probability of Positive Profit, Expected Total Profit, and Value at Risk of a 1-year, 100-MW Forward Contract with 16-hour On-peak Delivery at ComEd.

ComEd forward price ($/MWH)	Probability of positive profit (%)	Expected total profit ($000)	Value at risk with 5% chance ($000)
36	25	− 674	− 2,332
37	40	− 264	− 1,923
37.64	50	0	− 1,658
38	56	145	− 1,513
39	71	554	− 1,103
40	83	964	− 694
41	91	1,374	− 284
41.69	**95**	**1,658**	**0**
42	96	1,783	125
43	99	2,193	534
44	Over 99	2,603	945

Note: The price for a 12-month forward contract with Cinergy delivery is assumed to be $39/MWH. At this Cinergy price, the variance of the average of daily per MWH profit is $s_{\mu}^2 = \$6.02$/MWH. The ComEd price in **bold** is $G_{0.95}$ that achieves a 95% probability of positive profit and a zero VAR.

compute the contract's expected profit, and calculate the contract's value at risk. Such information is useful to the marketer's decision making.

We demonstrate our approach using spot-price data for two pairs of wholesale markets: Mid-C/COB and ComEd/Cinergy. The regression results for both market pairs support our contention that the spot-price relationship is not spurious and can be used for the purpose of cross hedging, pricing, and risk assessment. The empirical results confirm our expectation that a forward contract with ComEd delivery is more risky than one with Mid-C delivery because the ComEd/Cinergy market pair's spot-price volatility is substantially larger than that of the Mid-C/COB market pair. Finally, the simple cross-hedging strategy is effective for these market pairs and calls into question the need for a more complicated and probably more costly strategy.

NOTES

1. Equation (7) shows that any G is an increasing function of s_{π}. Hence, the non-dogmatic marketer will, *ceteris paribus*, choose a higher G than will that marketer's

dogmatic counterpart. The higher G is reflective of the non-dogmatic marketer's uncertainty as to the true values of the regression parameters α and β. Indeed, our approach implicitly assumes that the marketer will set the price *and* the amount to purchase in the second market (that is, β) simultaneously, with the profit in the first market *always* being uncertain because one never knows the true α, and the profit earned in the second market uncertain *a priori* because one never knows the true β.

An alternative approach to the problem would assert that once the regression has been estimated the marketer knows with certainty that the profit in the second market will be $b(P_{2t} - F)$. In this case, the total profit will be given by: $G - \alpha + (b - \beta)P_{2t} - bF - \varepsilon_t$. Repeating the earlier machinations, it is immediately determined that the expected profit is again computed as in Eq. (5a). It can also be verified that the variance differs from that of Eq. (5b), only insofar as the term $F^2 s_b^2$ is replaced by $P_{2e}^2\, s_b^2$, where $P_{2\mathrm{e}}$ is the expected price in the second market. Most assuredly, that expected price will be less than F. Hence, this approach will yield a lower variance than will the approach that we adopt. Thus, our approach is the most inherently conservative approach to the problem, because it is that which gives the highest variance.

2. See Davidson and MacKinnon (1993) for a description of the unit root and cointegration tests, and Woo et al. (1997) for their applications in an analysis of electricity spot prices.

REFERENCES

Anderson, R. W., & Danthine, J.-P. (1981). Cross Hedging. *Journal of Political Economy*, *89*, 1182–1196.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Jorion, P. (1997). *Value at Risk*. Chicago: Irwin.

Joskow, P. L. (1997). Restructuring, Competition, and Regulatory Reform in the U.S. Electricity Sector. *Journal of Economic Perspectives*, *11*, 119–138.

Michaels, R. J., & Ellig, J. (1999). Price Spike Redux: A Market Emerged, Remarkably Rational. *Public Utilities Fortnightly*, *137*, 40–47.

Raiffa, H., & Schlaifer, R. (1961). *Applied Statistical Decision Theory*. Boston: Harvard University.

Seigel, D. R., & Seigel, D. F. (1990). *The Futures Markets*. Chicago: Probus.

Seiple, C. (1996). Liquid Power Markets: A Rising Tide. *Public Utilities Fortnightly*, *134*, 12–13.

Woo, C. K., Lloyd-Zannetti, D., & Horowitz, I. (1997). Electricity Market Integration in the Pacific Northwest. *Energy Journal*, *18*, 75–101.

Woo, C. K., Horowitz, I., & Hoang, K. (2001). Cross Hedging and Forward Contract Pricing of Electricity. *Energy Economics*, *23*, 1–15.

This Page Intentionally Left Blank

USING MICROSOFT EXCEL AND DECISION TREES TO DEMONSTRATE THE BINOMIAL OPTION PRICING MODEL

John C. Lee

ABSTRACT

In the early 1990s Lotus 123 was "the" software application that every businessperson had to have. By the mid 1990s Microsoft Excel took over the honor with a vengeance. Now in the early 2000s, Lotus 123 is basically nowhere to be found. This paper will demonstrate why Microsoft Excel became so popular by creating the four Decision Trees for the Binomial Option Pricing Model. At the same time, this paper will discuss the Binomial Option Pricing Model in a less mathematical fashion. All the mathematical calculations will be taken care by the Microsoft Excel program that is presented in this paper. This relieves the reader from the computational burden of the Binomial Option Pricing Model and allows the reader to concentrate on the concepts of the Binomial Option Pricing Model.

1. INTRODUCTION

Microsoft Excel is one of the most powerful and valuable tools available to the business users. The financial industry in New York City has recognized this value. We can see this by going to one of the many jobs sites on the Internet.

Advances in Investment Analysis and Portfolio Management, Volume 8, pages 303–329.

ISBN: 0-7623-0798-6

Two Internet sites that demonstrate the value of some one who knows Microsoft Excel very well are www.dice.com and www.net-temps.com. For both of these Internet sites, search by New York City and VBA, which is Microsoft Excel's programming language, and you will see many jobs posting requiring VBA. In these VBA job posting you will see many jobs that only require Microsoft Excel VBA and most of these jobs pay over $70,000, around twice the national average.

The academic world has begun to realize the value of Microsoft Excel. There are now many books that use Microsoft Excel to do statistical analysis and financial modeling. This can be shown by going to the Internet site www.amazon.com and search for books by "Data Analysis Microsoft Excel" and search for books by "Financial Modeling Microsoft Excel".

The Binomial Option pricing model is one the most famous models used to price options. Only the Black-Scholes model is more famous. One problem with learning the Binomial Option pricing model is that it is computationally intensive. This results in a very complicated formula to price an option.

The complexity of the Binomial Option pricing model makes it a challenge to learn the model. Most books teach the Binomial Option model by describing the formula. This is a not very effective because it usually requires the learner to mentally keep track of many details, many times to the point of information overload. There is a well-known principle in psychology that the average number of things that a person can remember at one time is seven.

As a teaching aid, many books include Decision Trees. Because of the computational intensity of the model, most books do not present Decision Trees with more than 3 periods. One problem with this is that the Binomial Option model is best when the periods are large.

This paper will do two things. It will first demonstrate the power of Microsoft Excel. It will do this by demonstrating that it is possible to create large Decision Trees for the Binomial Pricing Model using Microsoft Excel. A 10 period Decision Tree would require 2047 call calculations and 2047 put calculations. This paper will also show the Decision Tree for the price of a stock and the price of a bond, each requiring 2047 calculation. Therefore there would be $2{,}047 * 4 = 8{,}188$ calculations for a complete set of 10 period Decision Trees.

The second thing that this paper will do is present the Binomial Option model in a less mathematical matter. It will try to make it so that the reader will not have to keep track of many things at one time. It will do this by using Decision Trees to price call and put options.

Included with this paper is a Microsoft Excel workbook called *binomia-loptionpricingmodel.xls* that contains the VBA code to create the Decision

Trees for the Binomial Option Pricing Model. The password for the workbook is *bigsky* for those who want to study the code. This workbook requires Microsoft Excel 97 or higher to run.

Section 2 discusses the basic concepts of call and put options. Sections 3 and 4 demonstrate the one period call and put option-pricing models. Section 5 demonstrates the two-period option-pricing model. Section 6 demonstrates a Microsoft Excel program that creates the Decision Trees for the Binomial Option Pricing model for *n* periods. Section 7 presents the Excel VBA code that created the Decision Trees. Finally, section 8 summarizes the paper

2. CALL AND PUT OPTIONS

A *call option* gives the owner the right but not the obligation to buy the underlying security at a specified price. The price in which the owner can buy the underlying price is called the *exercise price*. A call option becomes valuable when the exercise price is less than the current price of the underlying stock price.

For example, a call option on an IBM stock with an exercise price of $100 when the stock price of an IBM stock is $110 is worth $10. The reason it is worth $10 is because a holder of the call option can buy the IBM stock at $100 and then sell the IBM stock at the prevailing price of $110 for a profit of $10. Also, a call option on an IBM stock with an exercise price of $100 when the stock price of an IBM stock is $90 is worth $0.

A *put option* gives the owner the right but not the obligation to sell the underlying security at a specified price. A put option becomes valuable when the exercise price is more than the current price of the underlying stock price.

For example, a put option on an IBM stock with an exercise price of $100 when the stock price of an IBM stock is $90 is worth $10. The reason it is worth $10 is because a holder of the put option can buy the IBM stock at the prevailing price of $90 and then sell the IBM stock at the put price of $100 for a profit of $10. Also, a put option on an IBM stock with an exercise price of $100 when the stock price of the IBM stock is $110 is worth $0.

Figures 1 and 2 are charts showing the value of call and put options of the above IBM stock at varying prices.

3. CALL OPTION PRICING – ONE PERIOD

What should be the value of these options? Lets look at a case where we are only concern with the value of options for one period. In the next period a stock price can either go up or go down. Lets look at a case where we know for

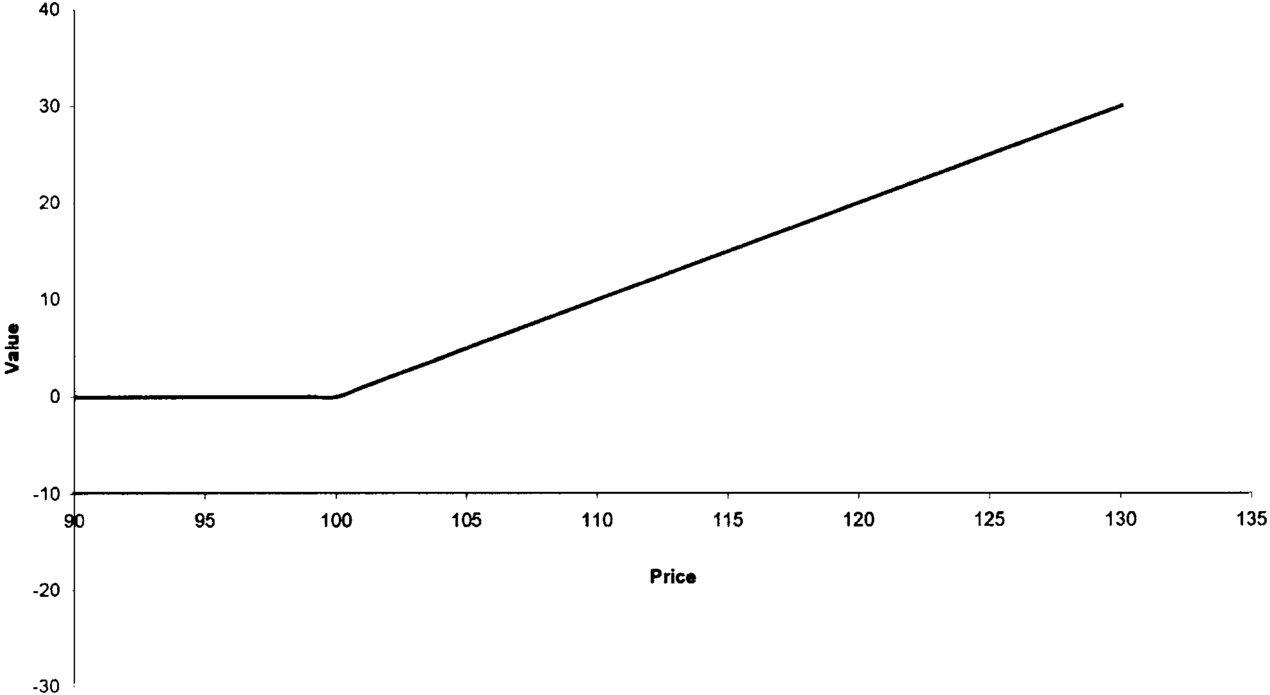


Fig. 1. Value of Call Option.

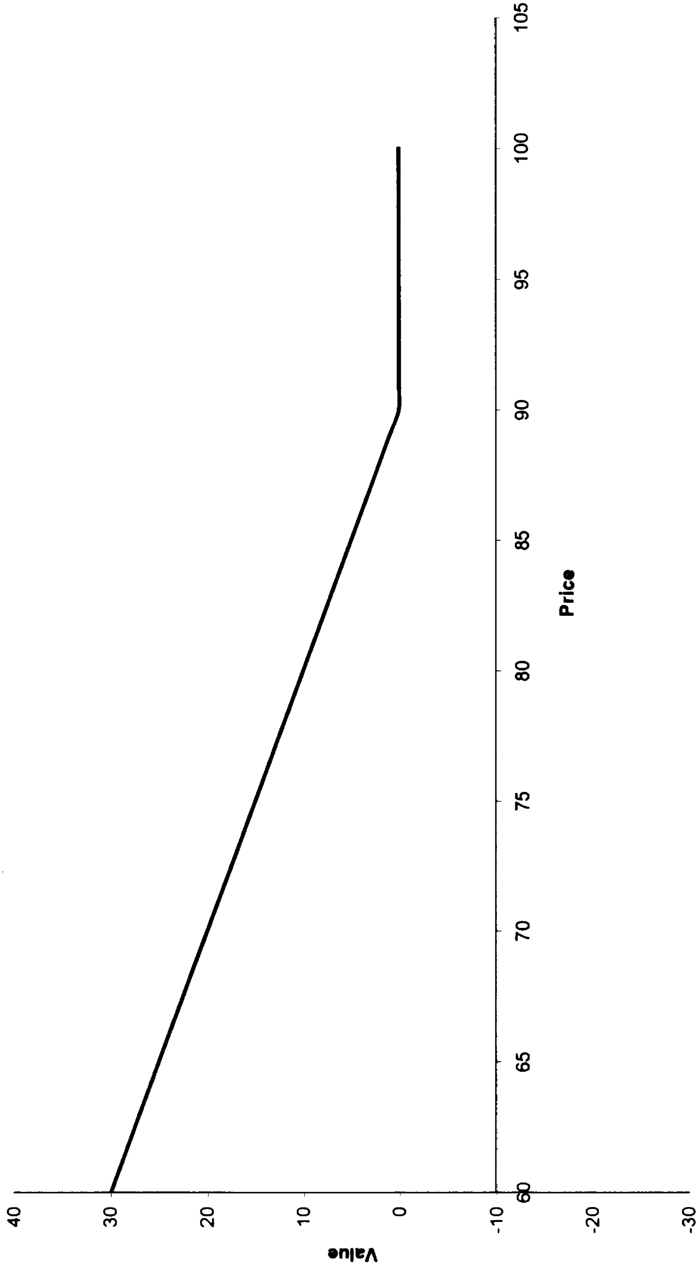


Fig. 2. Put Option Value.

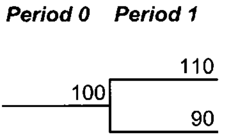


Fig. 3. Stock Price.

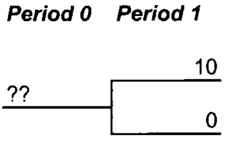


Fig. 4. Call Option Price.

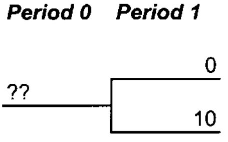


Fig. 5. Put Option Price.

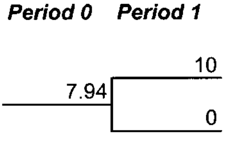


Fig. 6. Call Option Price.

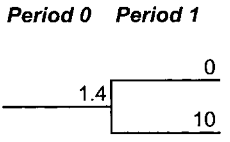


Fig. 7. Put Option Price.

certain that a stock with a price of \$100 will either go up 10% or go down 10% in the next period and the exercise after one period is \$100. Figures 3, 4 and 5 shows the Decision Tree for the stock price, the call option price and the put option price respectively.

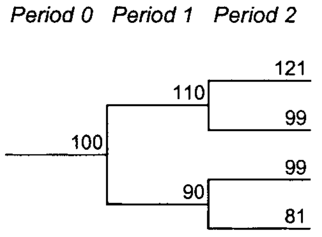


Fig. 8. Stock Price.

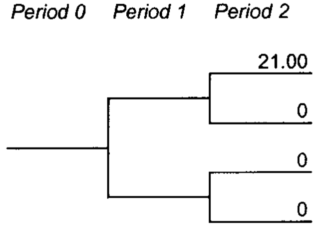


Fig. 9. Call Option.

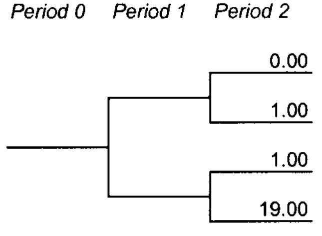


Fig. 10. Put Option.

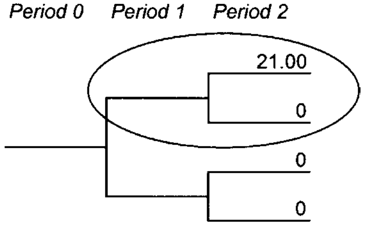


Fig. 11. Call Option.

Lets first consider the issue of pricing a call option. Using a one period Decision Tree we can illustrate the price of a stock if it goes up and the price of a stock if it goes down. Since we know the possible endings values of a stock, we can derive the possible ending values of a call option. If the stock price increases to \$110, the price of the call option will then be \$10 (\$110–\$100). If the stock price decreases to \$90, the value of the call option will worth \$0 because it would be below the exercise price of \$100. We have just discussed the possible ending value of a call option in period 1. But, what we are really interested is what is the value now of the call option knowing the two resulting value of a call option.

To help determine the value of a one period call option, its useful to know that it is possible to replicate the resulting two state of the value of the call option by buying a combination of stocks and bonds. Below is the formula to replicate the situation where the price increases to \$110. We will assume that the interest rate for the bond is 7%.

$$110S + 1.07B = 10$$

$$90S + 1.07B = 0$$

We can use simple algebra to solve for both S and B. The first thing that we need to do is to rearrange the second equation as follows,

$$1.07B = -90S$$

With the above equation, we can rewrite the first equation as

$$\begin{aligned} 110S + (-90S) &= 10 \\ 20S &= 10 \\ S &= 0.5 \end{aligned}$$

We can solve for B by substituting the value 0.5 for S in the first equation.

$$\begin{aligned} 110(0.5) + 1.07B &= 10 \\ 55 + 1.07B &= 10 \\ 1.07B &= -45 \\ B &= -42.05607 \end{aligned}$$

Therefore, from the above simple algebraic exercise, we should at period 0 buy 0.5 shares of IBM stock and borrow 42.05607 at 7% to replicate the payoff of the call option. This means the value of a call option should be $0.5*100 - 42.05607 = 7.94393$.

If this were not the case, there would then be arbitrage profits. For example, if the call option were sold for \$8 there would be a profit of 0.056607. This would result in the increase in the selling of the call option. The increase in the supply of call options would push the price down for the call options. If the call option were sold for \$7, there would be a saving of 0.94393. This saving would result in the increase demand for the call option. This increase demand would result in the price of the call option to increase. The equilibrium point would be 7.94393.

Using the above mentioned concept and procedure, Benninga (2000) has derived a one-period call option model as

$$C = q_u \,\mathrm{Max}[\mathrm{S}(1+\mathrm{u})\mathrm{X}, 0] + q_d \,\mathrm{Max}[\mathrm{S}(1+\mathrm{d}) - \mathrm{X}, 0] \tag{1}$$

where,

$$q_u = \frac{i-d}{(1+i)(u-d)}$$

$$q_d = \frac{u-i}{(1+i)(u-d)}$$

$u =$ increase factor

$d =$ down factor

$i =$ interest rate

If we let $i = r$, $p = (r-d)/(u-d)$, $1-p = (u-r)/(u-d)$, $\mathrm{R} = 1/(1+r)$, $C_u = \mathrm{Max}[\mathrm{S}(1+\mathrm{u}) - \mathrm{X}, 0]$ and $C_d = \mathrm{Max}[\mathrm{S}(1+\mathrm{d}) - \mathrm{X}, 0]$ then we have

$$C = [pC_u + (1-p)C_d]/\mathrm{R}, \tag{2}$$

where, $C_u =$ call option price after increase

$C_d =$ call option price after decrease

Equation (2) is identical to Eq. (6B.6) in Lee et al. (2000, p. 234).[1]

Below calculates the value of the above one period call option where the strike price, X, is \$100 and the risk free interest rate is 7%. We will assume that that the price of a stock for any given period will either increase or decrease by 10%.

$X = \$100$
$\mathrm{S} = \$100$
$u = 1.10$
$\mathrm{d} = 0.9$
$\mathrm{R} = 1 + \mathrm{r} = 1 + 0.07$

$$p = (1.07 - 0.90)/(1.10 - 0.90)$$
$$C = [0.85(10) + 0.15(0)]/1.07 = \$7.94$$

Therefore from the above calculations, the value of the call option is \$7.94. Figure 6 shows the resulting Decision Tree for the above call option.

4. PUT OPTION PRICING – ONE PERIOD

Like the call option, it is possible to replicate the resulting two state of the value of the put option by buying a combination of stocks and bonds. Below is the formula to replicate the situation where the price decreases to \$90.

$$110S + 1.07B = 0$$

$$90S + 1.07B = 10$$

We will use simple algebra to solve for both S and B. The first thing we will do is to rewrite the second equation as follows,

$$1.07B = 10 - 90S$$

The next thing to do is to substitute the above equation to the first put option equation. Doing this would result in the following,

$$110S + 10 - 90S = 0$$

The following solves for S,

$$20S = -10$$

$$S = -0.5$$

Now lets solve for B by putting the value of S into the first equation. This is shown below.

$$110(-0.5) + 1.07B = 0$$

$$1.07B = 55$$

$$B = 51.04$$

From the above simple algebra exercise we have $S = -0.5$ and $B = 51.04$. This tells us that we should in period 0 lend \$51.04 at 7% and sell 0.5 shares of stock to replicate the put option payoff for period 1. And, the value of the put option should be $100*(20.5) + 51.40 = -50 + 51.40 = 1.40$.

Using the same arbitrage argument that we used in the discussing of the call option, 1.40 has to be the equilibrium price of the put option.

As with the call option, Benninga (2000) has derived a one-period put option model as

$$P = q_u \, \mathrm{Max}[\mathrm{X} - \mathrm{S}(1 + \mathrm{u}), 0] + q_d \, \mathrm{Max}[\mathrm{X} - \mathrm{S}(1 + \mathrm{d}), 0] \tag{3}$$

where,

$$q_u = \frac{i - d}{(1 + i)(u - d)}$$

$$q_d = \frac{u - i}{(1 + i)(u - d)}$$

$$u = \text{ increase factor}$$

$$d = \text{down factor}$$

$$i = \text{ interest rate}$$

If we let $i = r$, $p = (r - \mathrm{d})/(u - \mathrm{d})$, $1 - p = (u - \mathrm{r})/(u - \mathrm{d})$, $\mathrm{R} = 1/(1 + r)$, $P_u = \mathrm{Max}[\mathrm{X} - \mathrm{S}(1 + \mathrm{u}), 0]$ and $P_d = \mathrm{Max}[\mathrm{X} - \mathrm{S}(1 + \mathrm{d}), 0]$ then we have

$$P = [pP_u + (1 - p)P_d]/\mathrm{R}, \tag{4}$$

where, P_u = put option price after increase

P_d = put option price after decrease

Below calculates the value of the above one period call option where the strike price, X, is \$100 and the risk-free interest rate is 7%.

$$P = [0.85(0) + 0.15(10)]/1.07 = \$1.40$$

From the above calculation the put option pricing Decision Tree would look like the following.

Figure 7 shows the resulting Decision Tree for the above put option.

5. OPTION PRICING – TWO PERIOD

We now will look at pricing options for two periods. Figure 8 shows the stock price Decision tree based on the parameters indicated in the last section. This Decision Tree was created based on the assumption that a stock price will either increase by 10% or decrease by 10%.

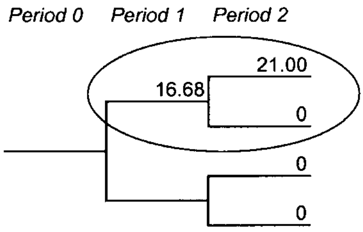


Fig. 12. Call Option.

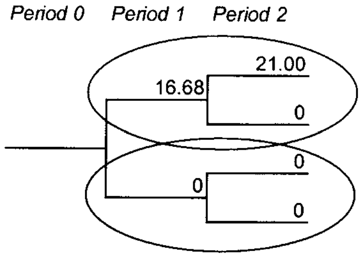


Fig. 13. Call Option.

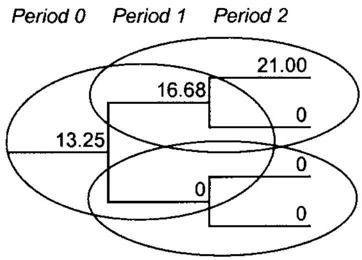


Fig. 14. Call Option.

How do we price the value of a call and put option for two periods?

The highest possible value for our stock based on our assumption is \$121. We get this value first by multiplying the stock price at period 0 by 110% to get the resulting value of \$110 of period 1. We then again multiply the stock price in period 1 by 110% to get the resulting value of \$121. In period two, the value of a call option when a stock price is \$121 is the stock price minus the exercise price, \$121 − 100, or \$21 dollars. In period two, the value of a put option when a stock price \$121 is the exercise price minus the stock price, \$100 − \$121, or − \$21. A negative value has no value to an investor so the value of the put option would be \$0.

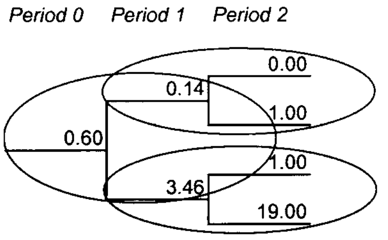


Fig. 15. Put Option.

Table 1.

Periods	Calculations
6	127
7	255
8	511
9	1023
10	2047
11	4065
12	8191

The lowest possible value for our stock based on our assumptions is \$81. We get this value first by multiplying the stock price at period 0 by 90% (decreasing the value of the stock by 10%) to get the resulting value of \$90 of period 1. When then again multiply the stock price in period 1 by 90% to get the resulting value of \$81. In period two, the value of a call option when a stock price is \$81, is the stock price minus the exercise price, \$81 − \$100, or − \$19. A negative value has no value to an investor so the value of a call option would be \$0. In period two, the value of a put option when a stock price is \$81 is the exercise price minus the stock price, \$100 − \$ 81, or \$19. We can derive the call and put option value for the other possible value of the stock in period 2 in the same fashion.

Figures 9 and 10 show the possible call and put option values for period 2.

We cannot calculate the value of the call and put option in period 1 the same way we did in period 2 because it's not the ending value of the stock. In period 1 there are two possible call values. One value is when the stock price increased and one value is when the stock price decreased. The call option Decision tree shown above shows two possible values for a call option in period 1. If we just focus on the value of a call option when the stock price increase from period one, we will notice that it is like the Decision Tree for a call option for one period. This is shown in Fig. 11.

Using the same method for pricing a call option for one period, the price of a call option when stock price increase from period 0 will be $16.68. The resulting Decision Tree is shown in Fig. 12.

In the same fashion we can price the value of a call option when a stock price decrease. The price of a call option when a stock price decreases from period 0 is $0. The resulting Decision Tree is shown in Fig. 13.

In the same fashion we can price the value of a call option in period 0. The resulting Decision Tree is shown in Fig. 14.

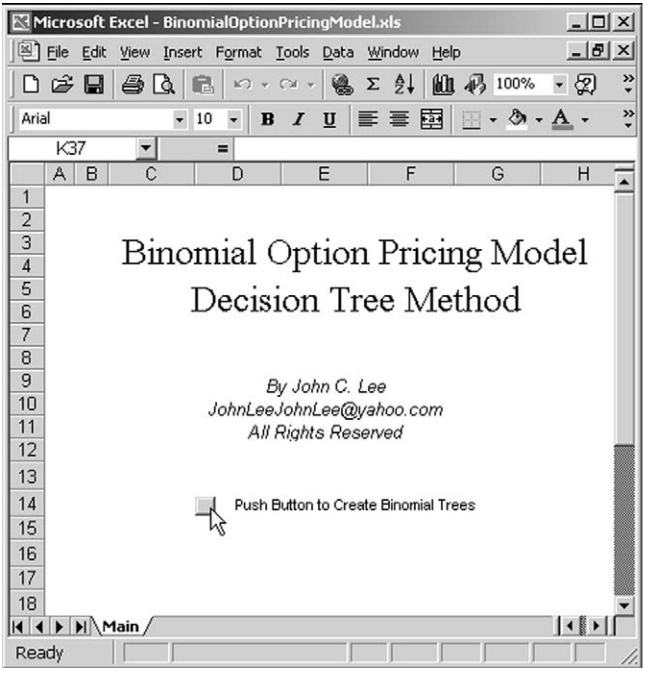


Fig. 16. Excel File *binomialoptionpricingmodel.xls*.

We can calculate the value of a put option in the same manor as we did in calculating the value of a call option. The Decision Tree for a put option is shown in Fig. 15.

6. USING MICROSOFT EXCEL TO CREATE THE BINOMIAL OPTION CALL TREES

In the previous section we priced the value of a call and put option by pricing backwards, from the last period to the first period. This method of pricing call and put options will work for any *n* period. To price the value of a call options for 2 periods required 7 sets of calculations. The number of calculations increases dramatically as *n* increases. Table 1 lists the number of calculations for specific number of periods.

After two periods it becomes very cumbersome to calculate and create the Decision Trees for a call and put option. In the previous section we saw that calculations where very repetitive and mechanical. To solve this problem, this

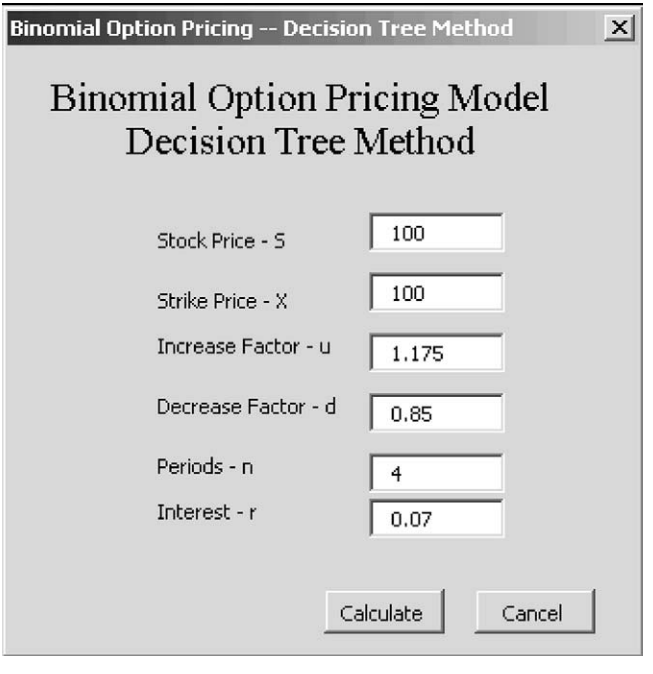


Fig. 17. Dialog Box Showing Parameters for the *Binomial Option Pricing Model*.

paper will use Microsoft Excel to do the calculations and create the Decision Trees for the call and put options. We will also use Microsoft Excel to calculate and draw the related Decision Trees for the underlying stock and bond.

To solve this repetitive and mechanical calculation of the Binomial Option Pricing Model, we will look at a Microsoft Excel file called *binomialoptionpricingmodel.xls* that is included with this paper.

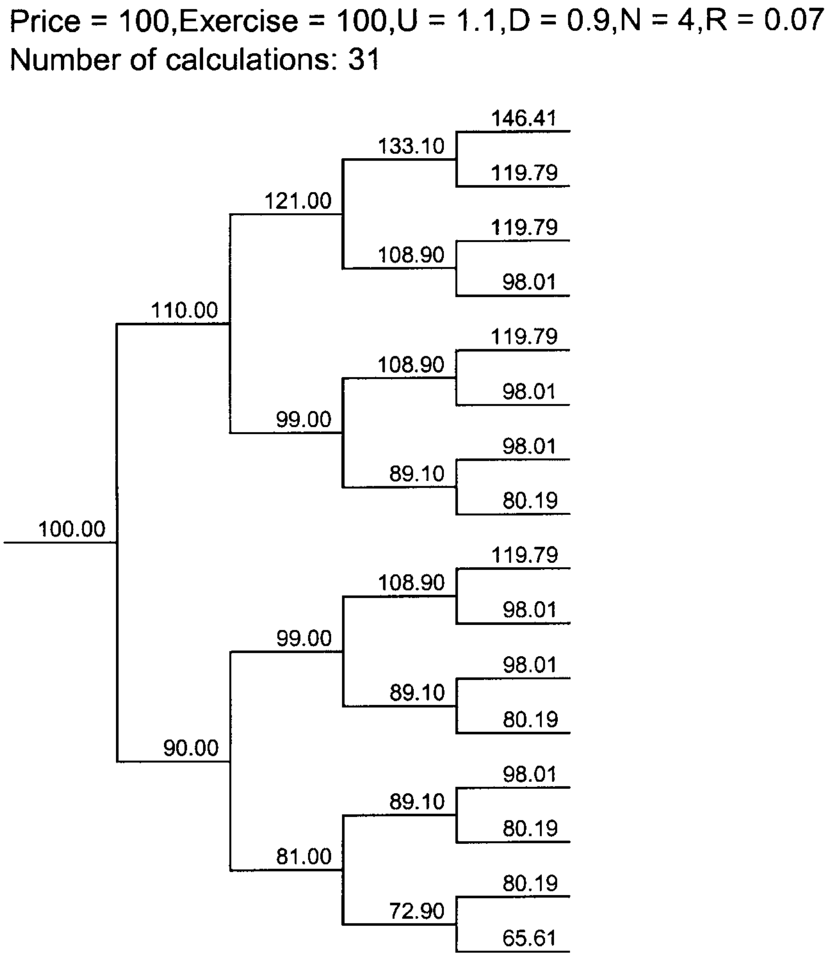


Fig. 18. Stock Price Decision Tree.

We will use this Excel file to produce four Decision Trees for the IBM stock that was discussed in the previous sections. The four Decision Trees are,

(1) Stock Price
(2) Call Option Price
(3) Put Option Price
(4) Bond Price

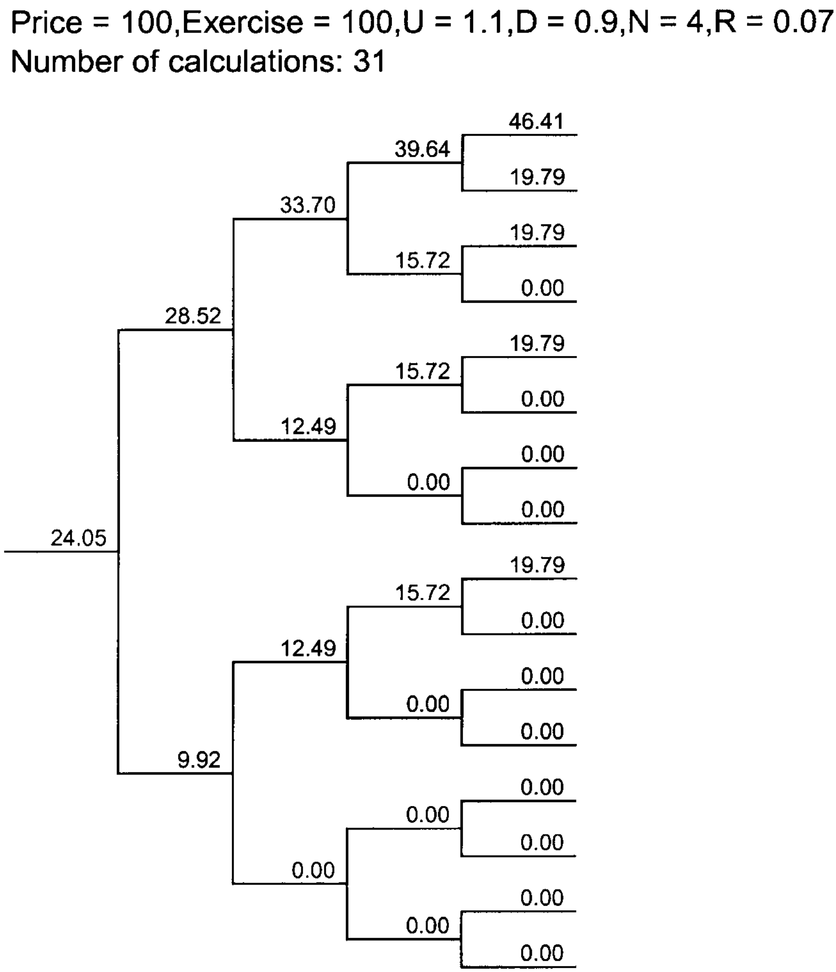


Fig. 19. Call Opton Pricing Decision Tree.

This section will demonstrate how to use the *binomialoptionpricingmodel.xls* Excel file to create the four Decision Trees.

Figure 16 shows the Excel file *binomialoptionpricingmodel.xls* after the file is opened.

Pushing the button shown in Fig. 16 will get the dialog box shown in Fig. 17.

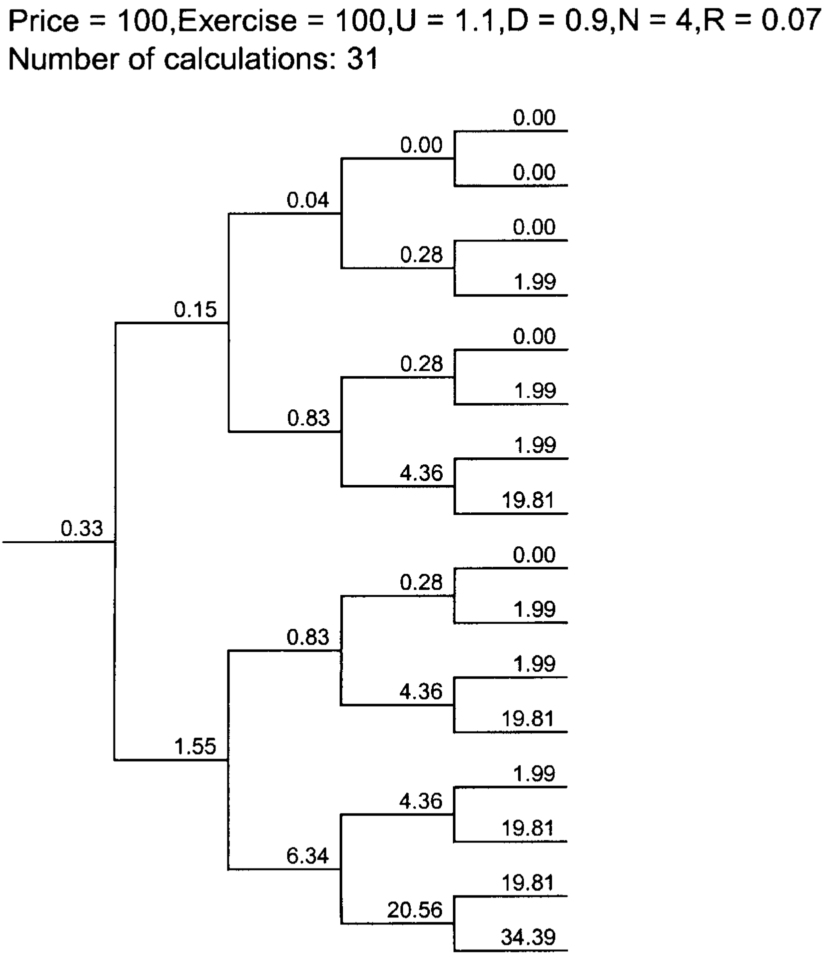


Fig. 20. Put Option Pricing Decision Tree.

The dialog box shown in Fig. 17 shows the parameters for the *Binomial Option Pricing Model*. These parameters are changeable. The dialog box in Fig. 17 shows the default values.

Pushing the *calculate* button shown in Fig. 17 will produce the four Decision Trees shown in Figs 18, 19,20 and 21.

The Table at the beginning of this section indicated 31 calculations were required to create a Decision Tree that has four periods. This section showed

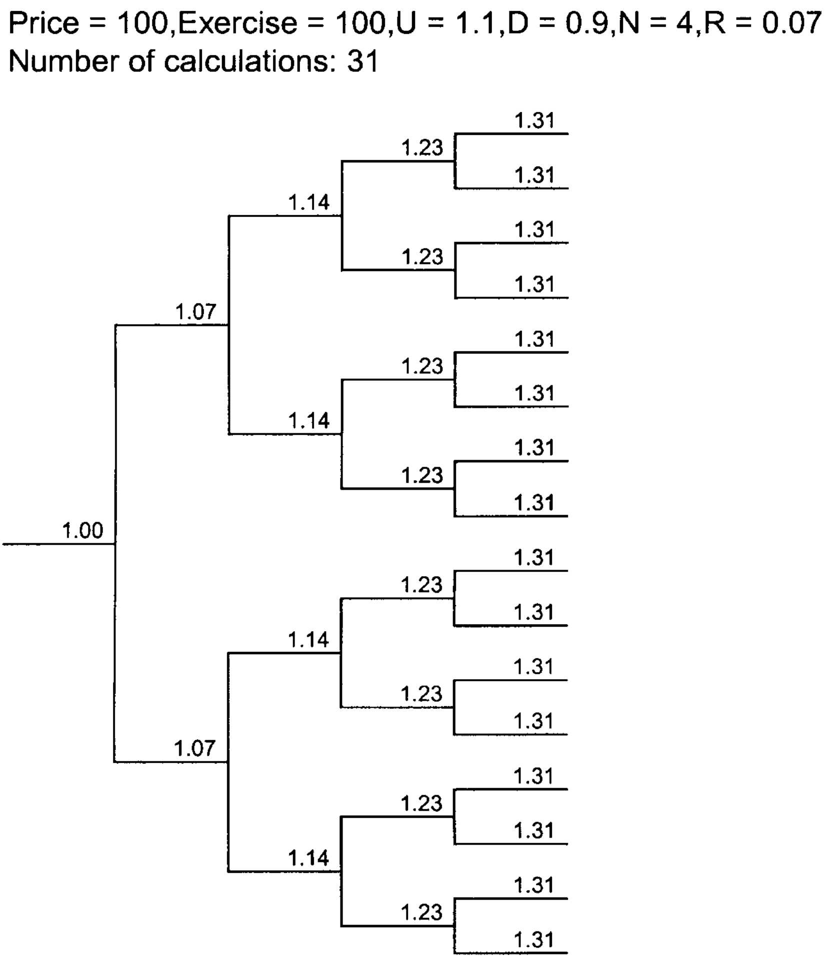


Fig. 21. Bond Pricing Decision Tree.

four Decision Trees. Therefore, the Excel file did 31 * 4 = 121 calculations to create the four Decision Trees.

Benninga (2000, p. 260) has defined the price of a call option in a Binomial Option Pricing model with n periods as,

$$C = \sum_{i=0}^{n} \binom{n}{i} q_u^i q_d^{n-i} \max[S(1+u)^i(1+d)^{n-i} - X, 0] \tag{5}$$

and the price of a put option in a Binomial Option Pricing model with n periods as,

$$P = \sum_{i=0}^{n} \binom{n}{i} q_u^i q_d^{n-i} \max[X - S(1+u)^i(1+d)^{n-i}, 0] \tag{6}$$

Lee et al. (2000, p. 237) has defined the pricing of a call option in a Binomial Option Pricing model with n period as,

$$C = \frac{1}{R^n} \sum_{k=0}^{n} \frac{n!}{k!(n-k!)} p^k (1-p)^{n-k} \max[0, (1+u)^k(1+d)^{n-k}, S - X] \tag{7}$$

The definition of the pricing of a put option in a Binomial Option Pricing model with n period would then be defined as,

$$P = \frac{1}{R^n} \sum_{k=0}^{n} \frac{n!}{k!(n-k!)} p^k (1-p)^{n-k} \max[0, X - (1+u)^k(1+d)^{n-k}, S] \tag{8}$$

7. EXCEL VBA CODE – BINOMIAL OPTION PRICING MODEL

It is important to note that the thing that makes Microsoft Excel powerful is that it offers a powerful professional programming language called Visual Basic for Applications. This section shows the VBA code that generated the Decision Trees for the Binomial Option pricing model. This code is in the form *frmBinomiaOption*. The procedure *cmdCalculate_Click* is the first procedure to run.

```
'/***************************************************************************
'/              Binomial Option Pricing Model
'/                   Decision Tree Method
'/
'/                    by John C. Lee
'/                JohnLeeJohnLee@yahoo.com
'/                   All Rights Reserved
'/***************************************************************************
Option Explicit
Dim mwbTreeWorkbook As Workbook
Dim mwsTreeWorksheet As Worksheet
Dim mwsCallTree As Worksheet
Dim mwsPutTree As Worksheet
Dim mwsBondTree As Worksheet
Dim mdblPFactor As Double
Dim mBinomialCalc As Long
'/**********************************************
'/Purpose:  Keep track the numbers of binomial calc
'/**********************************************
Property Let BinomialCalc(l As Long)
    mBinomialCalc = l
End Property
Property Get BinomialCalc() As Long
    BinomialCalc = mBinomialCalc
End Property
Property Set TreeWorkbook(wb As Workbook)
    Set mwbTreeWorkbook = wb
End Property
Property Get TreeWorkbook() As Workbook
    Set TreeWorkbook = mwbTreeWorkbook
End Property
Property Set TreeWorksheet(ws As Worksheet)
    Set mwsTreeWorksheet = ws
End Property
Property Get TreeWorksheet() As Worksheet
    Set TreeWorksheet = mwsTreeWorksheet
End Property
Property Set CallTree(ws As Worksheet)
    Set mwsCallTree = ws
End Property
Property Get CallTree() As Worksheet
    Set CallTree = mwsCallTree
End Property
Property Set PutTree(ws As Worksheet)
    Set mwsPutTree = ws
End Property
Property Get PutTree() As Worksheet
    Set PutTree = mwsPutTree
End Property
Property Set BondTree(ws As Worksheet)
    Set mwsBondTree = ws
End Property
Property Get BondTree() As Worksheet
    Set BondTree = mwsBondTree
End Property

Property Let PFactor(r As Double)
    Dim dRate As Double

    dRate = ((1 + r) - Me.txtBinomialD) / (Me.txtBinomialU - Me.txtBinomialD)
    Let mdblPFactor = dRate
End Property
Property Get PFactor() As Double
    Let PFactor = mdblPFactor
```

```
End Property
Private Sub cmdCalculate_Click()
    Me.Hide
    BinomialOption
    Unload Me
End Sub

Private Sub cmdCancel_Click()
    Unload Me
End Sub


Private Sub UserForm_Initialize()
   With Me
        .txtBinomialS = 100
        .txtBinomialX = 100
        .txtBinomialD = 0.9
        .txtBinomialU = 1.1
        .txtBinomialN = 4
        .txtBinomialr = 0.07
    End With
    Me.Hide
End Sub

Sub BinomialOption()
    Dim wbTree As Workbook
    Dim wsTree As Worksheet
    Dim rColumn As Range
    Dim ws As Worksheet

    Set Me.TreeWorkbook = Workbooks.Add
    Set Me.BondTree = Me.TreeWorkbook.Worksheets.Add
    Set Me.PutTree = Me.TreeWorkbook.Worksheets.Add
    Set Me.CallTree = Me.TreeWorkbook.Worksheets.Add
    Set Me.TreeWorksheet = Me.TreeWorkbook.Worksheets.Add

    Set rColumn = Me.TreeWorksheet.Range("a1")

    With Me
        .BinomialCalc = 0
        .PFactor = Me.txtBinomialr
        .CallTree.Name = "Call Option Price"
        .PutTree.Name = "Put Option Price"
        .TreeWorksheet.Name = "Stock Price"
        .BondTree.Name = "Bond"
    End With

    DecitionTree rCell:=rColumn, nPeriod:=Me.txtBinomialN + 1, _
                dblPrice:=Me.txtBinomialS, sngU:=Me.txtBinomialU, _
                sngD:=Me.txtBinomialD

    DecitionTreeFormat
    TreeTitle wsTree:=Me.TreeWorksheet, sTitle:="Stock Price "
    TreeTitle wsTree:=Me.CallTree, sTitle:="Call Option Pricing"
    TreeTitle wsTree:=Me.PutTree, sTitle:="Put Option Pricing"
    TreeTitle wsTree:=Me.BondTree, sTitle:="Bond Pricing"

    Application.DisplayAlerts = False
    For Each ws In Me.TreeWorkbook.Worksheets
        If Left(ws.Name, 5) = "Sheet" Then
            ws.Delete
        Else
            ws.Activate
            ActiveWindow.DisplayGridlines = False
            ws.UsedRange.NumberFormat = "#,##0.00_);(#,##0.00)"
        End If
    Next
    Application.DisplayAlerts = True
    Me.TreeWorksheet.Activate
End Sub
Sub TreeTitle(wsTree As Worksheet, sTitle As String)
```

```
    wsTree.Range("A1:a5").EntireRow.Insert (xlShiftDown)
    With wsTree
        With .Cells(1)
               .Value = sTitle
               .Font.Size = 20
               .Font.Italic = True
        End With
        With .Cells(2, 1)
             .Value = "Decision Tree"
             .Font.Size = 16
             .Font.Italic = True
        End With
        With .Cells(3, 1)
             .Value = "Price = " & Me.txtBinomialS & _
                    ",Exercise = " & Me.txtBinomialX & _
                    ",U = " & Me.txtBinomialU & _
                    ",D = " & Me.txtBinomialD & _
                    ",N = " & Me.txtBinomialN & _
                    ",R = " & Me.txtBinomialr
             .Font.Size = 14
        End With
        With .Cells(4, 1)
            .Value = "Number of calculations: " & Me.BinomialCalc
            .Font.Size = 14
        End With
    End With
End Sub
Sub BondDecisionTree(rPrice As Range, arCell As Variant, iCount As Long)
    Dim rBond As Range
    Dim rPup As Range
    Dim rPDown As Range

    Set rBond = Me.BondTree.Cells(rPrice.Row, rPrice.Column)
    Set rPup = Me.BondTree.Cells(arCell(iCount - 1).Row, arCell(iCount - 1).Column)
    Set rPDown = Me.BondTree.Cells(arCell(iCount).Row, arCell(iCount).Column)

    If rPup.Column = Me.TreeWorksheet.UsedRange.Columns.Count Then
        rPup.Value = (1 + Me.txtBinomialr) ^ (rPup.Column - 1)
        rPDown.Value = rPup.Value
    End If

    With rBond
        .Value = (1 + Me.txtBinomialr) ^ (rBond.Column - 1)
        .Borders(xlBottom).LineStyle = xlContinuous
    End With
            rPDown.Borders(xlBottom).LineStyle = xlContinuous
            With rPup
                .Borders(xlBottom).LineStyle = xlContinuous
                .Offset(1, 0).Resize((rPDown.Row - rPup.Row), 1). _
                    Borders(xlEdgeLeft).LineStyle = xlContinuous
            End With

End Sub


Sub PutDecisionTree(rPrice As Range, arCell As Variant, iCount As Long)
    Dim rCall As Range
    Dim rPup As Range
    Dim rPDown As Range

    Set rCall = Me.PutTree.Cells(rPrice.Row, rPrice.Column)
    Set rPup = Me.PutTree.Cells(arCell(iCount - 1).Row, arCell(iCount - 1).Column)
    Set rPDown = Me.PutTree.Cells(arCell(iCount).Row, arCell(iCount).Column)

    If rPup.Column = Me.TreeWorksheet.UsedRange.Columns.Count Then
        rPup.Value = WorksheetFunction.Max(Me.txtBinomialX - arCell(iCount - 1), 0)
        rPDown.Value = WorksheetFunction.Max(Me.txtBinomialX - arCell(iCount), 0)
    End If

    With rCall
```

```
            .Value = (Me.PFactor * rPup + (1 - Me.PFactor) * rPDown) / (1 +
Me.txtBinomialr)
            .Borders(xlBottom).LineStyle = xlContinuous
    End With
            rPDown.Borders(xlBottom).LineStyle = xlContinuous
            With rPup
                .Borders(xlBottom).LineStyle = xlContinuous
                .Offset(1, 0).Resize((rPDown.Row - rPup.Row), 1). _
                    Borders(xlEdgeLeft).LineStyle = xlContinuous
            End With

End Sub
Sub CallDecisionTree(rPrice As Range, arCell As Variant, iCount As Long)
    Dim rCall As Range
    Dim rCup As Range
    Dim rCDown As Range

    Set rCall = Me.CallTree.Cells(rPrice.Row, rPrice.Column)
    Set rCup = Me.CallTree.Cells(arCell(iCount - 1).Row, arCell(iCount - 1).Column)
    Set rCDown = Me.CallTree.Cells(arCell(iCount).Row, arCell(iCount).Column)

    If rCup.Column = Me.TreeWorksheet.UsedRange.Columns.Count Then
        With rCup
            .Value = WorksheetFunction.Max(arCell(iCount - 1) - Me.txtBinomialX, 0)
            .Borders(xlBottom).LineStyle = xlContinuous
        End With
        With rCDown
                .Value = WorksheetFunction.Max(arCell(iCount) - Me.txtBinomialX, 0)
                .Borders(xlBottom).LineStyle = xlContinuous
        End With
    End If

    With rCall
        .Value = (Me.PFactor * rCup + (1 - Me.PFactor) * rCDown) / (1 + Me.txtBinomialr)
        .Borders(xlBottom).LineStyle = xlContinuous
    End With

    rCup.Offset(1, 0).Resize((rCDown.Row - rCup.Row), 1). _
                    Borders(xlEdgeLeft).LineStyle = xlContinuous
End Sub

Sub DecitionTreeFormat()
    Dim rTree As Range
    Dim nColumns As Integer
    Dim rLast As Range
    Dim rCell As Range
    Dim lCount As Long
    Dim lCellSize As Long
    Dim vntColumn As Variant
    Dim iCount As Long
    Dim lTimes As Long
    Dim arCell() As Range
    Dim sFormatColumn As String
    Dim rPrice As Range


    Application.StatusBar = "Formatting Tree.. "
    Set rTree = Me.TreeWorksheet.UsedRange
    nColumns = rTree.Columns.Count


    Set rLast = rTree.Columns(nColumns).EntireColumn.SpecialCells(xlCellTypeConstants, 23)
    lCellSize = rLast.Cells.Count
    For lCount = nColumns To 2 Step -1
        sFormatColumn = rLast.Parent.Columns(lCount).EntireColumn.Address
        Application.StatusBar = "Formatting column " & sFormatColumn
        ReDim vntColumn(1 To (rLast.Cells.Count / 2), 1)

        Application.StatusBar = "Assigning values to array for column " & _
            rLast.Parent.Columns(lCount).EntireColumn.Address
        vntColumn = rLast.Offset(0, -1).EntireColumn.Cells(1). _
```

```
        Resize(rLast.Cells.Count / 2, 1)
        rLast.Offset(0, -1).EntireColumn.ClearContents

        ReDim arCell(1 To rLast.Cells.Count)
        lTimes = 1
        Application.StatusBar = "Assigning cells to arrays. Total number of cells: " _
         & lCellSize
        For Each rCell In rLast.Cells
            Application.StatusBar = "Array to column " & sFormatColumn & " Cells " _
            & rCell.Row
            Set arCell(lTimes) = rCell
            lTimes = lTimes + 1
        Next
        lTimes = 1
        Application.StatusBar = "Formatting leaves for column " & sFormatColumn
        For iCount = 2 To lCellSize Step 2

            Application.StatusBar = "Formatting leaves for cell " & arCell(iCount).Address
            If rLast.Cells.Count <> 2 Then
                Set rPrice = arCell(iCount).Offset(-1 * ((arCell(iCount).Row - _
                arCell(iCount - 1).Row) / 2), -1)
                rPrice.Value = vntColumn(lTimes, 1)
            Else
                Set rPrice = arCell(iCount).Offset(-1 * ((arCell(iCount).Row - _
                arCell(iCount - 1).Row) / 2), -1)
                rPrice.Value = vntColumn
            End If


            arCell(iCount).Borders(xlBottom).LineStyle = xlContinuous
            With arCell(iCount - 1)
                .Borders(xlBottom).LineStyle = xlContinuous
                .Offset(1, 0).Resize((arCell(iCount).Row - arCell(iCount - 1).Row), 1). _
                    Borders(xlEdgeLeft).LineStyle = xlContinuous
            End With
            lTimes = 1 + lTimes

            CallDecisionTree rPrice:=rPrice, arCell:=arCell, iCount:=iCount
            PutDecisionTree rPrice:=rPrice, arCell:=arCell, iCount:=iCount
            BondDecisionTree rPrice:=rPrice, arCell:=arCell, iCount:=iCount
        Next

        Set rLast = rTree.Columns(lCount - 1).EntireColumn. _
               SpecialCells(xlCellTypeConstants, 23)
               lCellSize = rLast.Cells.Count
    Next ' / outer next

    rLast.Borders(xlBottom).LineStyle = xlContinuous
    Application.StatusBar = False
End Sub

'/*****************************************************************
'/Purpse: To calculate the price value of every state of the binomial
'/        decision tree
'/*****************************************************************
Sub DecitionTree(rCell As Range, nPeriod As Integer, _
            dblPrice As Double, sngU As Single, sngD As Single)
    Dim lIteminColumn As Long


    If Not nPeriod = 1 Then
        'Do Up
        DecitionTree rCell:=rCell.Offset(0, 1), nPeriod:=nPeriod - 1, _
                dblPrice:=dblPrice * sngU, sngU:=sngU, _
                sngD:=sngD
        'Do Down
        DecitionTree rCell:=rCell.Offset(0, 1), nPeriod:=nPeriod - 1, _
                dblPrice:=dblPrice * sngD, sngU:=sngU, _
                sngD:=sngD
    End If
```

```
    lIteminColumn = WorksheetFunction.CountA(rCell.EntireColumn)


    If lIteminColumn = 0 Then
        rCell = dblPrice

    Else
        If nPeriod <> 1 Then
            rCell.EntireColumn.Cells(lIteminColumn + 1) = dblPrice
        Else
            rCell.EntireColumn.Cells(((lIteminColumn + 1) * 2) - 1) = dblPrice
            Application.StatusBar = "The number of binomial calcs are : " & _
           Me.BinomialCalc & " at cell " & _
           rCell.EntireColumn.Cells(((lIteminColumn + 1) * 2) - 1).Address
        End If

    End If
    Me.BinomialCalc = Me.BinomialCalc + 1
End Sub
```

8. SUMMARY

In this paper we demonstrated why Microsoft Excel is a very powerful application and why the Financial Industry in New York City value people that know Microsoft Excel very well. Microsoft Excel gives the business user the ability to create powerful application quickly with out relying on the Information Technology (IT) department. Prior to Microsoft Excel, business user would have to rely heavily on the Information Technology department. There are two problems with relying on the IT department. The first problem is that the tools that the IT department was using resulted in longer development time. The second problem was that the IT department was not as familiar with the business processes as the business users.

Simultaneously this paper demonstrated, with the aid of Microsoft Excel and Decision Trees, the Binomial Option model in a less mathematically fashion. This paper allowed the reader to focus more on the concepts by studying the associated Decision Trees, which were created by Microsoft Excel. This paper also demonstrate that using Microsoft Excel release the reader from the computation burden of the Binomial Option Model.

This paper also published the Microsoft Excel VBA code that created the Binomial Option Decision Trees. This allows for those who are interested to study the many advance Microsoft Excel VBA programming concepts that were used to create the Decision trees. One major computer science programming concept used by Microsoft Excel VBA is recursive programming. Recursive programming is the ideal of a procedure calling its self many times. Inside the procedure there are statements to decide when not to call it self.

NOTE

1. Please note that in Lee et al. (2000, p. 234) $u = 1 +$ percentage of price increase, $d = 1 -$ percentage of price increase.

REFERENCES

Benninga, S. (2000). *Financial Modeling*. Cambridge: MIT Press.

Daigler, R. T. (1994). *Financial Futures and Options Markets Concepts and Strategies*. New York: Harper Collins.

Jarrow, R., TurnBull, S. (1996). *Derivative Securities*. Cincinnati: South-Western College Publishing.

Lee, C.-F., Lee, J. C., Lee, A. C. (2000). *Statistics for Business and Financial Economics*. New Jersey: World Scientific.

Wells, E., Harshbarger, S. (1997). *Microsoft Excel 97 Developer's Handbook*. Redmond: Microsoft Press.

Walkenbach, J. (1999). *Microsoft Excel 2000 Power Programming with VBA*. New York: IDG Books WorldWide.

This Page Intentionally Left Blank