ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

Volume 9

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 9

ADVANCES IN INVESTMENT ANALYSIS AND PORTFOLIO MANAGEMENT

EDITED BY

CHENG-FEW LEE
Department of Finance, Rutgers University, USA

2002

JAI
An Imprint of Elsevier Science

Amsterdam – Boston – London – New York – Oxford – Paris
San Diego – San Francisco – Singapore – Sydney – Tokyo

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

First edition 2002

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: 0-7623-0887-7

∞ 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

Anlin Chen	Department of Business Management, National Sun Yat-Sen University, Taiwan
James F. Cotter	Wayne Calloway School of Business and Accountancy, Wake Forest University, USA
John Elder	Department of Finance, College of Business Administration, Dakota State University, USA
Christophe Faugère	School of Business, University of Albany, USA
Jonathan Fletcher	Department of Accounting and Finance, University of Sthrathclyde, UK
Jean L. Heck	Department of Finance, College of Commerce and Finance, Villanova University, USA
Michael M. Holland	Department of Finance, College of Commerce and Finance, Villanova University, USA
Ken Hung	Department of Business, Management National Dong Hwa University, Taiwan
Christine X. Jiang	Area of Finance, The Fogelman College of Business and Economics, The Univeristy of Memphis, USA
Jang-Chul Kim	Fogelman College of Business and Economics, University of Memphis, USA
Clarence C. Y. Kwan	Michael G. DeGroote School of Business, McMaster University, Canada

Cheng-Few Lee	Department of Finance and Economics, Graduate School of Management, Rutgers University, USA
Yi-Tsung Lee	Department of Accounting, National Chengchi University, Taiwan
Karyl Leggio	University of Missouri, USA
Li Li	Department of Finance and Economics, Graduate School of Management, Rutgers University, USA
Gwohorng Liaw	Department of Economics, Tunghai University, Taiwan
Donald Lien	Department of Economics, School of Business, University of Kansas, USA
Y. Angela Liu	Department of Business Administration, National Chung Chen University, Taiwan
Ming-Shiun Pan	Department of Finance, Decision Sciences, and Information Systems, USA
Mahmut Parlar	Michael G. DeGroote School of Business, McMaster University, Canada
David R. Shaffer	Department of Finance, College of Commerce and Finance, Villanova University, USA
Hany Shawky	School of Business, University at Albany, USA
Chin W. Yang	Department of Economics, Clarion University of Pennsylvania, USA
Felicia A. Yang	Department of Economics, University of Pennsylvania, 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 twelve papers. The abstract of each chapter is as follows:

Chapter 1. Christophe Faugère and Hany Shawky develop an endogenous growth model that incorporates random technological shocks to the economy. These random technological shocks affect both production and the depreciation of capital. We show the existence of a long-run steady-state growth path, and characterize it. An optimal growth rate for the economy and the long-run expected stock return are both derived. We then turn to study the volatility of expected stock returns around this steady state. Once tested, the model shows that deviations of de-trended capital stock, deviations of shocks from expected values and deviations of labor force growth from steady state together explain about 20% of the deviations of stock returns from long-term expected values. Our estimates also implies that investors have levels of risk aversion consistent with the literature, and that labor growth fluctuations are not significant, due to crowding out effects.

Chapter 2. Chin W. Yang, Ken Hung and Felicia A. Yang examine the equivalence property between the angle-maximization portfolio technique and the Markowitz risk minimization model is proved. Via reciprocal and monotonic transformation, they can be made equivalent with or without different types of short sale. Since the Markowitz portfolio model is formulated in the standard convex quadratic programming, the equivalence property would enable us to apply the same well-known mathematic properties to the angle maximization model and enjoy the same convenient computational advantage of the quadratic program (e.g., Markowitz's critical line algorithm).

Chapter 3. Donald Lien and Karyl Leggio consider optimal ratios for different lengths of hedging horizon when the highest frequency data is generated by a cointegrated system. It is found that, after reparameterization, a temporal aggregation of cointegrated systems remains a cointegrated system. This result provides a convenient method to estimate n-day hedge ratio for any integer n. The only remaining issue concerns the possible incorrect lag selections. Empirical results from ten futures contracts however indicate lag selections have no effect on the estimated hedge ratios.

Chapter 4. Cheng-Few Lee and Li Li test various CAPM-based market-timing and selectivity models, we find that about 12% of the funds have a statistically significant Alpha with about 4% of the funds having a significantly positive Alpha, and 8% of the funds having a significantly negative Alpha. About 15% of funds show significant timing ability with about 9% funds having a significantly positive timing coefficient and 6% of the funds having a significantly negative timing coefficient. The Asset Allocation funds demonstrate the most timing ability and the Aggressive Growth funds demonstrate the least timing ability.

Chapter 5. John Elder investigates the extent to which three observable macroeconomic factors can explain the time-varying risk premia in the short-end of the term structure. We employ an empirical model that is motivated by a dynamic asset pricing model with time-varying risk premia and time-invariant reward-to volatility measures. We find that, in our model, two factors explain up to

65% of the temporal variation in Treasury bill returns, with the short-end of the term structure responding significantly to contemporaneous innovations the funds rate and shifts (or twists) in the yield curve. Our primary new findings are that a factor based on shifts in the yield curve may explain the time variation in risk premia at the very short end of the term structure, and that a factor based on innovations in the federal funds rate may be weakly linked to the time-varying risk premia over the post-1966 sample, when the federal funds market first began to function as a major source of bank liquidity. This latter result is somewhat sensitive to the sample period.

Chapter 6. Christine X. Jiang and Jang-Chul Kim use a sample of stock splits on NYSE listed ADRs between 1994 and 1999, we study the change in liquidity following stock splits. Our findings suggest that cost to liquidity demanders measured by percentage quoted and effective bid-ask spreads, split-factor adjusted quoted depth and trading volume increases for split-up securities. However, we observe that raw trading volume and depth both go up after splits, suggesting that liquidity may increase because market makers/brokers' higher incentives in promoting the shares for larger payments on order flows. In addition, number of small trades and number of shareholders go up 28% and 21%, respectively while institutional holdings pre- and post-splits are not significantly different, also consistent with the notion that splits provide an incentive for brokers to promote the stocks, and their efforts seem to target small investors.

Chapter 7. Clarence C. Y. Kwan and Mahmut Parlar consider portfolio selection with round-lot requirements in analytical settings where short sales are disallowed and allowed. In either case, by exploiting some analytical properties of the objective function in portfolio optimization, we are able to approximate the round-lot solution without the encumbrance of any algorithmic complexities that are often associated with integer programming. The efficient heuristic we use to solve the resulting nonlinear integer programming problem examines only the corner points of a 'hypercube' surrounding the optimal fractional solution found without the round-lot requirements. Then, by characterizing the covariance structure of security returns with the single index model, we establish the correspondence between the round-lot solution and the solution without round-lot requirements for which security selection criteria in terms of risk-return trade-off are available. This correspondence, in turn, provides useful information regarding the sensitivity of the round-lot solution in response to changes in return expectations. Given these nice features, the analysis should enhance the practical relevance of portfolio modeling for assisting investment decisions.

Chapter 8. Jean L. Heck, Michael M. Holland, and David R. Shaffer examine that while a major consequence of the use of debt by a business is generally assumed to be a change to the risk of default, theoretical work relating this risk to the lender's required rate of return is notably sparse. This paper defines an equilibrium model to value debt given a non-zero probability of default by extending previous research and then formulates the corresponding appropriate security market line. Also, a model to value debt is synthesized that compensates a lender for both capital market risk and default risk.

Chapter 9. Yi-Tsung Lee and Gwohorng Liaw look at how some studies, such as Bagwell (1992) and Bernardo and Cornell (1997), provided evidences that the shareholders' valuations differ dramatically. They argued that the valuations differ substantially, implying a significantly small supply or demand elasticity. However, Kandel et al. (1999) indicated quite an elastic demand for stocks of Israeli IPOs that were conducted as non-discriminatory auctions. To resolve these controversial findings, this paper discusses the procedure of measuring price elasticity and provides some measures of elasticity. In addition to indicating that Bagwell's measure tends to underestimate the actual elasticity, this study supplements previous work by testing under another auction mechanism, discriminatory pricing rule, and our results are consistent with Kandel et al.'s findings.

Chapter 10. Anlin Chen and James F. Cotter show that private information as well as public information is important in revising the terms of the offer during the pre-selling period (or the waiting period) and that when the revealed private information is positive, the underwriter compensates the investors for this information by underpricing the issue more than when the information is negative. Even though the cost of compensating positive information is quite high, the issuer still benefits from the positive infornation in that the wealth transferred to the investors is smaller under underwriter's information acquisition activities. Furthermore, IPO long-run performance is negatively related to the positive information revealed during the waiting period and the underwriter prestige. Finally, IPO firms without receiving significant information during the waiting period survive longer after issuance.

Chapter 11. Ming-Shiun Pan and Y. Angela Liu examines the term structure of correlations of weekly returns for six national stock markets namely, Australia, Hong Kong, Japan, Malaysia, Singapore, and the U.S. We decompose stock indexes into permanent and temporary components using a canonical correlation analysis and then calculate short- and long-horizon return correlations from these two price components. The empirical results for the sample period of January 1988 to December 1994 reveal that the relationships of return correlations among these stock markets are not stable across return horizons. While correlations, in general, tend to increase with return horizons, there are several cases showing that correlations decline when investment horizons increase.

Chapter 12. Jonathan Fletcher examines the out of sample performance of monthly asset allocation strategies within UK industry portfolios using linear asset pricing models and a characteristic-based model of stock returns to forecast expected returns. We find that strategies that use conditional versions of the asset pricing models outperforms the strategy that uses the characteristics-based model in terms of higher Sharpe performance and more positive abnormal returns. In addition, these strategies provide significant positive Jensen (1968) and Ferson and Schadt (1996) performance measures even with binding investment constraints. Our results support the usefulness of conditional asset pricing models in mean-variance analysis.

This Page Intentionally Left Blank

ENDOGENOUS GROWTH AND STOCK RETURNS VOLATILITY IN THE LONG RUN

Christophe Faugère and Hany Shawky

ABSTRACT

We develop an endogenous growth model that incorporates random technological shocks to the economy. These random technological shocks affect both production and the depreciation of capital. We show the existence of a long-run steady-state growth path, and characterize it. An optimal growth rate for the economy and the long-run expected stock return are both derived. We then turn to study the volatility of expected stock returns around this steady state. Once tested, the model shows that deviations of de-trended capital stock, deviations of shocks from expected values and deviations of labor force growth from steady state together explain about 20% of the deviations of stock returns from long-term expected values. Our estimates also implies that investors have levels of risk aversion consistent with the literature, and that labor growth fluctuations are not significant, due to crowding out effects.

1. INTRODUCTION

The efficient markets hypothesis implies that stock market prices should follow a random walk and thus, stock returns should be unpredictable. However, many recent studies such as Fama and French (1988a, b), Keim and Stambaugh

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 1–20.

ISBN: 0-7623-0887-7

(1986), French, Schwert and Stambaugh (1987), Campbell and Shiller (1988), Chen, Roll and Ross (1986), Lo and MacKinlay (1988) and Fama (1990) document returns predictability. These studies have shown that state variables such as aggregate production growth and yield spreads are empirically useful in predicting stock and bond returns.

In an attempt to explain the time-varying behavior of stock returns, two broad categories of asset pricing models emerged. Consumption-based asset pricing models such as in Merton (1973), Lucas (1978), Breeden (1979) and Ceccetti, Lam and Mark (1990) relate the returns on financial assets to the intertemporal marginal rate of substitution of consumers using a consumption growth function. Production-based models on the other hand, relate the marginal rate of transformation to asset returns using production functions as in Cochrane (1991), Balvers, Cosimano and McDonald (1990) and Restoy and Rockinger (1994).

Cochrane (1991) finds that historical stock returns are related to economic variables such as growth rates of GNP and the investment to capital ratios. Unfortunately however, he does not provide a formal structure for explaining the real source and nature of economic fluctuations that might impact expected asset returns. Shawky and Peng (1995) use a real business cycle model with exogenous technical progress and show that technological shocks are critical factors in explaining asset returns.

We develop an endogenous growth model with technological shocks that affect both the final output and the depreciation rate of capital. We characterize the properties of a particular steady state growth path, where expected growth is constant in the long run. We then examine the volatility of stock returns around the steady state as a function of deviations of the de-trended capital stock, deviations from expected values of labor growth, and deviations of shocks from their long-term mean. This leads to an empirically testable hypothesis.

Our empirical results indicate that deviations of capital stock and technological shocks from their long run mean are significant variables, but that the deviations in labor growth are not, due perhaps to crowding out effects. The theoretical model predicts both the optimal growth rate of the economy as well as the long-term average stock market return. These are empirically testable implications. In fact, our results imply that investors exhibit a fair degree of consumption smoothing behavior. We also find that to be in accordance with the sample's expected stock return, our technology must exhibit some degree of increasing returns to capital.

This paper is organized in five sections. In Section 2 we set up the model and derive optimality conditions that are consistent with endogenous economic growth. We develop the concept of steady-growth in Section 3 and proceed to derive its equilibrium conditions. The time-series data, the methodology and the empirical results are presented and analyzed in Section 4. The final section provides a summary and some concluding remarks.

2. A MODEL OF ENDOGENOUS GROWTH

Consider a stochastic growth model similar to that in Brock-Mirman (1972), where technology is affected by a random shock every period. We also assume that growth is self-sustaining in a sense defined later on. Our goal is to investigate how stock returns are affected by long run technological trends.

Formally our goal is to search for the optimal policy that solves

$$J(Y_0) = \max E_0\left[\sum_{t=0}^{\infty} \rho^t U(C_t)\right]$$
$$C_t = (1 - i_t) \times Y_t \tag{1}$$
$$Y_t = \theta_t f(v_t, K_t)$$

Where C_t, K_t and Y_t are *per-unit-of-labor* consumption, capital stock and output, and v_t is the rate of capacity utilization. Labor L_t is assumed to evolve exogenously over time. The control variable is the investment rate i_t. The variable θ_t is a multiplicative random shock that is i.i.d. and is defined over a compact range $[\underline{\theta}, \overline{\theta}]$. The value of θ_t is realized at the beginning of period t. We assume that consumers' preferences are represented by $U(C_t) = C_t^{1-\gamma}/(1-\gamma)$, with $\gamma > 0$ representing the coefficient of relative risk aversion (CRRA).

We assume that the production function has constant returns to scale in capital and labor. In particular we use the following *per-capita* formulation:

$$f(v_t, K_t) = (A(v_t K_t)^{\alpha} + B)^{1/\alpha} \tag{2}$$

with $0 < \alpha < 1$ for now.[1] A crucial advantage of this formulation is that it allows the economy to grow at an endogenously sustained rate. In the traditional economic growth literature, an economy can only sustain growth by resorting to exogenous technical progress. In a sense, the fundamental source of economic growth is determined outside the model. The endogenous growth literature however, has sought to incorporate the sources of economic growth by featuring externalities (public spending, learning by doing) or certain factors of production that can be accumulated forever (human capital).

A critical feature for achieving endogenous growth is that these externalities counteract the natural tendency for decreasing returns to capital. In our model, given a stream of technological shocks, the economy will generate endogenous growth due to sufficiently high marginal returns to capital in the long run.[2] Capital utilization rates are included in the model because it is a way to measure the actual flow of services provided by the capital stock in place. These rates are exogenously determined and incorporating them in the model leads to a better estimate of the production function.[3]

We assume that capital depreciates at a stochastic rate δ_t, *and evolves according to*:[4]

$$K_{t+1} = [i_t Y_t + (1 - \delta_t) K_t] \frac{L_t}{L_{t+1}} \tag{3}$$

It is further assumed that the depreciation rate δ_t, is perfectly negatively correlated with the shocks θ_t, and hence we write $\delta_t = 1 - \mu\theta_t$, so that the above relationship becomes:

$$K_{t+1} = [i_t Y_t + \mu\theta_t K_t] \frac{L_t}{L_{t+1}} \tag{4}$$

The parameter μ must be such that $\mu\theta_t < 1$. The intuition for having a stochastic depreciation rate is that the outstanding stock of capital is generally subjected to the same type of transitory technological shocks as output.[5] For example, the productivity of labor measured in output/hour might be temporarily raised as a result of corporate downsizing. The productivity of capital might also be temporarily raised as a result of a credit crunch. A rise in productivity might induce some firms to slow down the rate of depreciation of certain capital goods.[6]

Next, we are solving for the social planner's optimum, as a way to characterize the optimal paths of consumption and investment in this economy.

A. Optimality Conditions

The standard first order condition is:[7]

$$\rho E_t \left\{ \frac{U'(C_{t+1})}{U'(C_t)} \theta_{t+1} f_2(v_{t+1}, K_{t+1}) \right\} \frac{L_t}{L_{t+1}} = 1 \tag{5}$$

Letting $\theta_{t+1} f_2(v_{t+1}, K_{t+1}) = (1 + R_{t+1})$ measure one plus the stock market return, we get:

$$\rho E_t \left\{ \frac{U'(C_{t+1})}{U'(C_t)} (1 + R_{t+1}) \right\} \frac{L_t}{L_{t+1}} = 1 \tag{6}$$

If we define $X_{t+1} = \dfrac{C_{t+1}}{C_t}$ and let $Z_{t+1} = X_{t+1}^{-\gamma}(1 + R_{t+1})$, then the first order condition becomes:

$$\rho E_t\{Z_{t+1}\} \frac{L_t}{L_{t+1}} = 1 \tag{7}$$

We will assume as in Hansen-Singleton (1983) that Z_{t+1} is log normally distributed $\ln(Z_{t+1}) \sim N(\mu_t, \sigma^2)$ conditional on the information available at t. Following their approach we can deduce a new first order condition as:

$$E_t\{R_{t+1}\} = \gamma E_t\left\{\ln\left(\frac{C_{t+1}}{C_t}\right)\right\} + \ln\left(\frac{L_{t+1}}{L_t}\right) - \ln(\rho) - \sigma^2/2 \tag{8}$$

Therefore:

$$R_{t+1} = \gamma E_t\left\{\ln\left(\frac{C_{t+1}}{C_t}\right)\right\} + \ln\left(\frac{L_{t+1}}{L_t}\right) - \ln(\rho) - \sigma^2/2 + \varepsilon_{t+1} \tag{9}$$

Where $\varepsilon_{t+1} = R_{t+1} - E\{R_{t+1}\}$. Equation (9) is identical to Hansen-Singleton (1983), except for the term involving labor growth. Hansen and Singleton state that it was not their goal to solve for an explicit representation of equilibrium prices in terms of the underlying shocks to technology. We take their model a step further by looking at the determinants of consumption growth in terms of technological progress and shocks.

3. STEADY STATE GROWTH

A characteristic of most industrialized economies is that per-capita real variables exhibit sustained growth over long periods of time. We will use the concept of steady state growth to describe a situation in which all state variables grow at the same constant *expected* rate. This is a novel approach in a growth model with stochastic shocks. Traditionally, the long-term stability of the economy refers to the convergence of cumulative distributions of shocks to a stationary distribution, as in Brock and Mirman (1972).

In order to characterize steady state growth, we need to transform the economy by detrending real variables.[8] Let g denote a particular growth rate and define new normalized variables as:

$$y_t = Y_t/(1+g)^t \quad c_t = C_t/(1+g)^t \quad k_t = K_t/(1+g)^t$$

We define a *Fulfilled Expectations Steady state* (FESS) as a vector $(\theta, g, n, \bar{i}, \bar{k}, \bar{y})$, where θ is the expected value of the random shock, g is the long-run expected growth rate of consumption, n is the long run growth rate of the labor force, and the vector $(\bar{i}, \bar{k}, \bar{y})$ is defined as follows:

$$\lim_{t\to\infty} \ln(i_t) = \ln(\bar{i}); \lim_{t\to\infty} \ln(k_t) = \ln(\bar{k})$$

$$\lim_{t\to\infty} \ln(y_t) = \lim_{t\to\infty} E_t\{\ln(y_{t+1})\} = \ln(\bar{y})$$

$$\text{with} \lim_{t\to\infty} E_t\{\ln(C_{t+1}/C_t)\} = g \text{ and } \lim_{t\to\infty} E_t\{\ln(L_{t+1}/L_t)\} = n$$

$$\text{and} \lim_{t\to\infty} \ln(\theta_t) = E_t\{\ln(\theta_{t+1})\} = \ln(\theta)$$

A Fulfilled Expectations Steady state is an equilibrium where the sequence of shock realizations converge to the expected value of the shock, and the long run expected growth rate is actually realized.[9] Our next proposition proves the existence of such a steady state.

Proposition 1: Assume there is a sequence of ex-post shocks which converges to θ, such that capacity utilization rates converge to a constant and the stock of capital grows at a constant rate in the long run, then a FESS $(\theta, g, n, \bar{i}, \bar{k}, \bar{y})$ exists and:

$$g = \lim_{t\to\infty} E_t\{\ln(C_{t+1}/C_t)\} = (1/\gamma)\left[\ln\left(\frac{\rho A^{1/\alpha}\bar{v}\theta}{1+n}\right) + \sigma^2/2\right]$$

$$\bar{R} = \lim_{t\to\infty} E_t\{R_{t+1}\} = \ln(A^{1/\alpha}\bar{v}\theta) = \gamma \ln(1+g) + \ln(1+n) - \ln(\rho) - \sigma^2/2$$

Proof: Assume that there exists a growth rate $\kappa > 0$ such that:

$$\lim_{t\to\infty} \ln(\theta_t) = \ln(\theta) \text{ and } \lim_{t\to\infty} \ln(\tilde{k}_t) = \ln(\tilde{k})$$

With $\tilde{k}_t = K_t/(1+\kappa)^t$. Therefore actual sequences of capital stocks grow to infinity. As $\lim_{t\to\infty} \ln(v_t) = \ln(\bar{v})$ this implies that:

$$\lim_{t\to\infty} E_t\{R_{t+1}\} = \lim_{t\to\infty} E_t\{\ln(\theta_{t+1} f_2(v_{t+1}, K_{t+1}))\} = \ln(A^{1/\alpha}\bar{v}\theta) \tag{10}$$

Where $\ln(\theta) = E_t\{\ln(\theta_{t+1})\}$. We conclude from the Euler equation that:

$$\lim_{t\to\infty} E_t\{\ln(C_{t+1}/C_t)\} = (1/\gamma)\left[\ln\left(\frac{\rho A^{1/\alpha}\bar{v}\theta}{1+n}\right) + \sigma^2/2\right] = g \tag{11}$$

So that consumption grows at a constant expected growth rate.[10] From Eqs (10) and (11) we can easily derive the second equality expressing the long-term expected rate $\bar{R}$ as a function of the growth rate and other parameters. From the capital accumulation equation we know that:

$$\lim_{t\to\infty} \ln(\tilde{k}_{t+1}) = \lim_{t\to\infty} \ln(i_t(Av_t^{\alpha} + BK_t^{-\alpha})^{1/\alpha} + \mu) + \lim_{t\to\infty} \ln(\tilde{k}_t)$$
$$+ \lim_{t\to\infty} \ln(\theta_t) - \kappa - \lim_{t\to\infty} \ln(L_{t+1}/L_t) \tag{12}$$

As $\lim_{t\to\infty} \ln(\theta_t) = \ln(\theta)$ this implies

$$\lim_{t\to\infty} \ln(A^{1/\alpha} i_t v_t + \mu) = n + \kappa - \ln(\theta) \tag{13}$$

so that $\lim_{t\to\infty} \ln(i_t v_t) = \text{constant}$. Since $\ln(v_t)$ converges to a constant $\ln(\bar{v})$ thus the log of the rate of investment $\ln(i_t)$ converges to a constant $\ln(\bar{i})$. If output is to grow at the same rate as consumption we have to set $\kappa = g$. In order to fulfill the last condition of existence of an FESS:

$$\lim_{t\to\infty} \ln(y_t) = \lim_{t\to\infty} E_t\{\ln(y_{t+1})\} = \lim_{t\to\infty} \ln(\bar{y}) \tag{14}$$

We need the following to hold true:

$$\lim_{t\to\infty} \ln(\theta_t) + \alpha^{-1} \lim_{t\to\infty} \ln[A + B(v_t K_t)^{-\alpha}] + \lim_{t\to\infty} \ln(v_t) + \lim_{t\to\infty} \ln(k_t)$$
$$= \lim_{t\to\infty} E_t \ln(\theta_{t+1}) + \alpha^{-1} \lim_{t\to\infty} \ln[A + B(v_{t+1} K_{t+1})^{-\alpha}] + \lim_{t\to\infty} \ln(v_{t+1})$$
$$+ \lim_{t\to\infty} \ln(k_{t+1}) = \ln(\bar{y}) \tag{15}$$

For some $\bar{y}$. Again, this is true when $\lim_{t\to\infty} \ln(\theta_t) = \ln(\theta)$. *Q.E.D.*

For the FESS to exist we impose a transversality condition that $\lim_{t\to\infty} (\rho(1+g)^{\gamma})^t = 0$. In other words, we need $\rho(1+g)^{\gamma} < 1$.[11] Proposition 1 gives exact closed form solutions for the optimal expected growth rate and the long run expected stock return. The long run expected stock return equals the long run expected productivity of capital. From a comparative statics perspective, we see that the long run rate of growth would rise with larger expected productivity, discount factor, and variance σ^2. The expected growth rate would drop with faster population growth, and a larger degree of risk aversion.[12]

A. *Deviations from the Steady State*

We follow the Real Business Cycle literature (Kydland-Prescott, 1982), and linearize the economy around the steady state (FESS). Even though an economy subjected to arbitrary shocks does not necessarily converge to the FESS, this steady state offers an interesting benchmark to look at macroeconomic fluctuations. It reproduces the stylized facts of actual economies, while still accounting for the random nature of shocks. One additional advantage is that by linearizing, we can construct a simple testable hypothesis about these fluctuations, without having to know the actual shape of optimal solutions.

The first step is to rewrite the first order conditions using normalized variables. Let us recall that $c_t = C_t/(1+g)^t$, and then we have:

$$R_{t+1} = \gamma E_t\left\{\ln\left(\frac{c_{t+1}}{c_t}\right)\right\} + \ln\left(\frac{L_{t+1}}{L_t}\right) - \ln\left(\frac{(1+g)^\gamma}{\rho}\right) - \sigma^2/2 + \varepsilon_{t+1} \quad (16)$$

Let $e_t = C_t/Y_t$ be the consumption rate, then we have:

$$R_{t+1} = \gamma E_t\left\{\ln\left(\frac{e_{t+1}}{e_t}\right)\right\} + \gamma E_t\left\{\ln\left(\frac{y_{t+1}}{y_t}\right)\right\} + \ln\left(\frac{L_{t+1}}{L_t}\right)$$

$$- \ln\left(\frac{(1+g)^\gamma}{\rho}\right) - \sigma^2/2 + \varepsilon_{t+1} \quad (17)$$

We denote with a 'hat' variables that represent deviations from the FESS. Thus $\hat{R}_{t+1} = (R_{t+1} - \bar{R})$ is the deviation of the stock market return, from its long run trend $\bar{R} = \ln(A^{1/\alpha}\bar{v}\theta)$. The variable $\hat{l}_{t+1} = (\ln(L_{t+1}/L_t) - n)$ is the deviation of labor force growth rate from its long-term value. We also define $\hat{y}_t = \ln(y_t/\bar{y})$, $\hat{k}_t = \ln(k_t/\bar{k})$, $\hat{v}_t = \ln(v_t/\bar{v})$, and $\hat{\theta}_t = \ln(\theta_t/\bar{\theta})$.

Because the representative agent's problem can be solved *after* we normalize the variables, analogous first order conditions imply that the optimal consumption rate decision can be rewritten as $e_t = 1 - i_t(y_t) = e(y_t)$.[13] If we define a monotonic transformation $\ln(e_t) = \mathrm{Q}(\ln(y_t))$, then the function $\ln(e_t)$ can be linearized around the FESS so that in effect we have:

$$\hat{e}_t = \ln(e_t/\bar{e}) \approx a \times \ln(y_t/\bar{y}) = a \times \hat{y}_t \text{ and } E_t \ln(e_{t+1}/\bar{e}) \approx a \times E_t \ln(y_{t+1}/\bar{y}) \quad (18)$$

Where $a = \mathrm{Q}'(\ln(\bar{y}))$. The variable a represents the elasticity of the *rate* of consumption with respect to income, along the FESS.[14] In the long run, we obtain the following expression for the return on the market:

$$\hat{R}_{t+1} = \gamma(1+a)E_t\left\{\ln\left(\frac{y_{t+1}}{y_t}\right)\right\} + \hat{l}_{t+1} + \varepsilon_{t+1} \quad (19)$$

This equation is similar to the Balvers et al. (1990), which suggests that the rate of return is mainly conditioned by output growth. The main difference is that we have labor growth as another explanatory variable, and we explicitly model the technological sector. Next, we expand the first term on the right hand side of (19) substituting in the specific production function:

$$E_t\left\{\ln\left(\frac{y_{t+1}}{y_t}\right)\right\} = E_t\{\ln(\theta_{t+1})\} + \alpha^{-1}\ln(A + B(v_{t+1}K_{t+1})^{-\alpha}) + \ln(v_{t+1}) + \ln\left(\frac{k_{t+1}}{y_t}\right) \tag{20}$$

The last term in (20) becomes:

$$\ln\left(\frac{k_{t+1}}{y_t}\right) = \ln(i_t + (\mu\theta_t k_t/y_t)) - \ln\left((1+g)\frac{L_{t+1}}{L_t}\right) \tag{21}$$

Thus, as the capital stock K_t grows without bounds, the previous expression (21) becomes:

$$E_t\left\{\ln\left(\frac{y_{t+1}}{y_t}\right)\right\} = \ln(A^{1/\alpha}\theta) + \ln(i_t v_t + \mu/A^{1/\alpha}) + \hat{v}_{t+1} - \hat{v}_t - \hat{l}_{t+1} - \ln((1+g)(1+n)) \tag{22}$$

Similar to the argument made previously about the consumption function, we can deduce that the optimal investment rate policy $i_t = \mathrm{I}(y_t)$ is a function of the normalized variable y_t, and $\mathrm{I}(y_t)$ is continuous. We also know that $y_t = \theta_t(A(v_t k_t)^\alpha + B(1+g)^{-\alpha t})^{1/\alpha}$. We can linearize this last function around the FESS, and express the second term on the RHS of (22) as:

$$\ln(i_t v_t + \mu/A^{1/\alpha}) = \ln(\overline{iv} + \mu/A^{1/\alpha}) + b_1\hat{\theta}_t + b_2\hat{k}_t + b_3\hat{v}_t \tag{23}$$

Where the coefficients b_is represent the elasticities of the *effective rate* of investment *iv*, with respect to θ, *k* and *v*, along the FESS. Finally, inserting this back into the linearized Euler condition (19) leads to:

$$\hat{R}_{t+1} = \gamma(1+a)(b_1\hat{\theta}_t + b_2\hat{k}_t + \hat{v}_{t+1} + (b_3 - 1)\hat{v}_t) + (1 - \gamma(1+a))\hat{l}_{t+1} + \gamma(1+a)\ln\left(\frac{\theta(A^{1/\alpha}\overline{iv} + \mu)}{(1+g)(1+n)}\right) + \varepsilon_{t+1} \tag{24}$$

Once we substitute the value for the growth rate g into this equation we obtain:

$$\hat{R}_{t+1} = \gamma(1+a)(b_1\hat{\theta}_t + b_2\hat{k}_t + \hat{v}_{t+1} + (b_3 - 1)\hat{v}_t) + (1 - \gamma(1+a))\hat{l}_{t+1} + (1+a)$$
$$\times \left[\gamma \ln(\overline{iv} + \mu/A^{1/\alpha}) - \gamma \ln(\bar{v}) - \ln(\rho) - (1-\gamma)\ln\left(\frac{\theta\bar{v}A^{1/\alpha}}{1+n}\right) - \sigma^2/2 \right]$$
$$+ \varepsilon_{t+1} \tag{25}$$

We will present some evidence in the next section that the rate of capacity utilization is related to labor growth, and that labor growth deviations from steady state are autocorrelated. In fact, we are making the following assumptions:

$$\hat{l}_{t+1} = \delta_1\hat{l}_t \text{ and } \hat{v}_{t+1} = \delta_2\hat{l}_{t+1} \tag{26}$$

Thus Eq. (25) becomes:

$$\hat{R}_{t+1} = \gamma(1+a)(b_1\hat{\theta}_t + b_2\hat{k}_t) + [1 - \gamma(1+a)[1 - \delta_2(1 + (b_3 - 1)/\delta_1)]] \cdot \hat{l}_{t+1}$$
$$+ (1+a)\left[\gamma \ln(\overline{iv} + \mu/A^{1/\alpha}) - \gamma \ln(\bar{v}) - \ln(\rho) - (1-\gamma)\ln\left(\frac{\theta\bar{v}A^{1/\alpha}}{1+n}\right) - \sigma^2/2 \right]$$
$$+ \varepsilon_{t+1} \tag{27}$$

Equation (27) is the fundamental result of the paper. It shows how stock return deviations from their long run mean are predicted by deviations of detrended capital stock, labor growth and technological shocks away from their long-term means. The rationale behind this equation is that stock return fluctuations around their mean are conditioned by fluctuations in real activity or the business cycle, around a growing trend. Our model essentially takes the stance that the sources of linearity in the relationship (19) come from a near steady state analysis, as well as the linearity of the technology in the long run. This equation is in a reduced form that lends itself easily to empirical testing.

4. EMPIRICAL RESULTS

A. Time Series Variables

All economic data series are obtained from Datastream International, spanning the period 1959–1998. All variables are quarterly, beginning first quarter 1959 to third quarter of 1998. The GDP (USGDP..D) and consumption expenditures (USCONEXPD) are deseasonalized real variables. Real investment is measured by real private non-residential investment (USNRINVD). Real wage is calculated by dividing nominal wage (USWAGSALB) by the CPI (USCP....F) normalized to be 100 in 1992. Labor is measured as total civilian population employed (USEMPTOTE). Capacity utilization in all industries (USOPERATE), is constructed for the missing years 1959–1967, by regressing capacity utilization over the period 1967–1998 onto the rate of employment and projecting that relationship backward over the missing years. Total stock returns were obtained using Datastream and Tradetools, for the S&P 500 and the dividend yield. Total annual real returns are calculated using the CPI index as deflator.

B. Construction of Capital Stocks and Technological Shocks

The capital stock is constructed using the permanent inventory approach. It is determined using quarterly data from 1959 to 1998. A feature of the model is that technological shocks influence depreciation. Thus there is a nested determination of shocks and capital stocks next period. The initial index of total factor productivity is derived using Baumol et al. (1986).[15] The subsequent indexes of technical shocks are obtained by using Thornqvist's formula given in Barro et al. (1994), which calculates discrete increments in the Solow residuals.[16] We select the depreciation parameter μ through a numerical procedure to get an average annual depreciation rate equal to 9.6%, over the sample period.[17]

C. The Production Function

Our regression is for the period 1966–1998. The production function is estimated with OLS. Here we follow a separate approach for constructing the capital stock. The capital stock is constructed based on the assumption that the depreciation rate is non-stochastic and *constant* at 9.6% per year.[18] In each period; the capital stock is weighted by the corresponding rate of capacity utilization. Our initial capital stock is arbitrarily chosen, thus the estimated stocks of capital will be unreliable for the first few quarters starting in 1959. However, as the stock is further accumulated and depreciated, future estimates become progressively more accurate. It was necessary to re-construct the sequence of random shocks, to avoid the problem of deriving shocks from time series of capital stocks, leading to serial correlation. We accomplish that by using a dual approach as in Barro (1998).[19] The series is then regressed on a linear time trend and detrended.

We use an iterative procedure where we estimate the parameters A and B, based on a chosen value for α (Table 1, Panel A). The second regression (Panel B) insures that the generated shock sequence indeed corresponds to the estimated residuals over the sample. This will be true only if, after substituting in the values for A and B, the coefficient on the exogenous variable is close to 1, and the constant term is close to zero. The parameter α is iteratively modified to achieve that outcome. The list of parameters of the production function is given in Table 4. From Table 1 we see that the adjusted $R^2 = 0.82$. The value for α is equal to 3.16, which implies increasing marginal returns to capital.[20] Even though the production function we chose conforms to the neoclassical theory of factor income distribution, the share of income going to capital rises, as the economy grows to the steady state, and then levels off. This is consistent with

Table 1. Production Function.

This regression is based on 128 quarterly observations for the period 1966–1998. The initial capital stock is chosen to be \$1 trillion (1992 dollars). The shocks θ_t are derived from a dual approach (Barro (1998)). The exponent α equals 3.16, and the depreciation rate equals a constant 2.4% per quarter.

Panel A:

$$\left[\frac{Y_t}{\theta_t}\right]^{\alpha} = A(v_t K_t)^{\varepsilon} + B$$

	Constant	$(v_t K_t)^{\alpha}$
Coefficient	4.71×10^{12}	0.03
T-values	(28.79)	(24.62)
Adjusted R^2	0.82	

Panel B:

$$\ln(Y_t^{\alpha}) = \rho_1 \ln[A(v_t K_t)^{\alpha} + B] + \rho_2 + u_t$$

where u_t is the residual. It should be true that $\exp(u_t) \approx \theta_t$ when the coefficient α is chosen appropriately so that $\rho_1 = 1$ and $\rho_2 = 0$. Here $\alpha = 3.16$.

	Constant	$\ln[A(v_t K_t)^{\alpha} + B]$
Coefficient	–0.39	1.01
T-values	(–1.1)	(84.01)
Adjusted R^2	0.98	

the labor productivity slowdown observed from the 1970s to the late 1980s (Baumol et al. (1991)).

D. Detrending the Variables

Recall that all our variables are detrended log deviations from a steady state path. In our case we detrend the variables using the average per-capita consumption growth rate for the period 1966–1997. We estimate it to be 1.23% annually. Deviations are defined with respect to the sample means.

E. Discussion of Results

Table 2 presents the results for our stock return regression corresponding to Eq. (27). In panel A, we are using 128 quarterly observations, from 1966 to 1997. The adjusted R^2 is 8%. In panels B, we use only second quarter observations at yearly intervals. The adjusted R^2 is 20%.[21] These results are consistent with the findings of Balvers, Cosimano and McDonald (1990), related to alternative

Table 2. Stock Returns, Productivity Shocks and other Economic Variables.

The rate of growth of the economy is 0.307% per quarter. The depreciation rate *averages* 2.4% per quarter.

$$\hat{R}_{t+1} = A_0 + A_1\hat{\theta}_t + A_2\hat{k}_t + A_3\hat{l}_{t+1} + \varepsilon_{t+1}$$

Panel A: This regression is based on 128 quarterly observations from the first quarter of 1966 to the first quarter of 1998.

	Constant	$\hat{\theta}_t$	$\hat{k}_t$	$\hat{l}_{t+1}$
Coefficient	0.57	197.60	24.26	–124.64
T-Values	(0.59)	(3.17)	(3.54)	(–0.83)
Adjusted R^2	0.08			

Panel B: This regression is based on 32 observations for the period 1966–1997. Each observation uses second quarter data at yearly intervals.

	Constant	$\hat{\theta}_t$	$\hat{k}_t$	$\hat{l}_{t+1}$
Coefficient	1.41×10^{-13}	575.84	122.06	–170.31
T-Values	(-5.5×10^{-14})	(2.25)	(3.22)	(–0.84)
Adjusted R^2	0.20			

return horizons. In the case of yearly intervals, we find that 20% of the volatility of stock returns around a steady state can be explained by macroeconomic variables such as volatility in capital stocks, technical shocks and labor growth fluctuations.

Looking at *t*-statistics, the only variables that are statistically significant are the detrended deviations of the capital stock and the Solow residual shocks. Labor growth deviations are not significant at the 95% confidence level. A possible interpretation for this result, is that the direct effects of employment growth are crowded out by the adjustments made to capacity utilization.[22] In fact, we find that rates of capacity utilization are correlated on a year to year basis with current labor growth, with an adjusted R^2 of 40% (see Table 3 panel B). As for labor growth, we find that it is autocorrelated on a quarterly basis with an adjusted R^2 of 27% (Table 3 Panel C). Initially, positive fluctuations in labor growth have a negative impact on stock returns, but because the rate of capacity utilization rises with the growth in labor, this tends to crowd out the first effect.[23]

We can also test our model using the predicted optimal growth rate for the economy given in Eq. (11). We derive the CRRA preference parameter that is consistent with having the theoretical optimal growth rate equal the sample average per capita consumption growth rate of 1.23% over the period 1966–1997. We find that the parameter value equals 2.84, which implies a fair degree of consumption smoothing behavior. The value of 0.965 for the discount rate is found by imposing the transversality condition, given the growth rate of the economy and the CRRA parameter. This is low compared to the range of estimates (0.988, 0.993) used in the standard RBC literature (King-Plosser & Rebelo, 1988; Ambler & Paquet, 1994).

Another interesting result is that the variance of the joint distribution of returns and marginal rates of substitution in consumption contributes for about 32% of the value of the expected growth rate of the economy.

The other important value derived from this exercise is the long-term interest rate found here to be equal to 7.75%, which is close to the annual mean return of 7.97%, over the sample period.[24] These results offer corroboration for our analysis.

5. CONCLUSION

Using an endogenous growth model we derived a theoretical relationship between the stock market returns deviations from long run expected value, expressed as a function of the deviations of aggregate macroeconomic variables

Table 3. Consumption Rate, Labor Growth and Capacity Utilization.

The rate of growth of the economy is 0.307% per quarter. The depreciation rate *averages* 2.4% per quarter.

Panel A: This regression is based on 32 observations at yearly intervals, from the second quarter of 1966 to the second quarter of 1997.

$$\hat{e}_t = a\hat{y}_t + b$$

	Constant	$\hat{y}_t$
Coefficient	7.19×10^{-3}	-10^{-5}
T-values	(3.87)	(–8.97)
Adjusted R^2	0.72	

Panel B: This regression is based on 32 observations for the period 1966–1997. Each observation is using second quarter data.

$$\hat{v}_{t+1} = \delta_2 \hat{l}_{t+1} + b$$

	Constant	$\hat{l}_{t+1}$
Coefficient	3×10^{-4}	1.95
T-values	(0.05)	(4.63)
Adjusted R^2	0.40	

Panel C: This regression is based on 128 quarterly observations for the period 1966–1997.

$$\hat{l}_{t+1} = \delta_1 \hat{l}_t + b$$

	Constant	$\hat{l}_t$
Coefficient	-1.56×10^{-6}	0.53
T-values	(-4×10^{-3})	(7.04)
Adjusted R^2	0.27	

Panel D: This regression is based on 32 observations for the period 1966–1997.

$$\ln(i_t v_t + \mu/A^{1/\alpha}) = \ln(\overline{iv} + \mu/A^{1/\alpha}) + b_1\hat{\theta}_t + b_2\hat{k}_t + b_3\hat{v}_t$$

	Constant	$\hat{\theta}_t$	$\hat{k}_t$	$\hat{v}_t$
Coefficient	–0.61	0.02	0.20	0.10
T-Values	(–331.62)	(0.13)	(5.71)	(1.81)
Adjusted R^2	0.66			

Table 4. Summary of Estimated and other Derived Parameters.

Avge yrly θ	A	B	Capacity Ut.	Quart. Dep.
4	0.03	4.71×10^{12}	82.24%	2.4%
α	ρ	n	σ^2	g
3.16	0.965	1.82%	2.22%	1.23%
δ_1	δ_2	μ	γ	Sample $\bar{R}$
0.079	1.95	0.154	2.84	7.97%
Derived $= \bar{R} = \ln(A^{1/\alpha}\hat{\nu}\theta) = 7.75\%$				

and technological shocks from their steady state values. The novelty of this model is that it shows the predictability of returns in the case when economic growth is endogenous, without appealing to outside exogenous progress. We also derived a closed form solution for the expected long run growth rate of the economy and the long-term expected stock market return, as functions of the underlying parameters of the economy.

The model implies that deviations of capital stocks and technological shocks and labor growth rate from their long run means, account for about 20% of the predictability of deviations of stock returns from long run trend. Labor growth does not play a significant role, because of a crowding out effect with capacity utilization.

The estimated parameters are consistent with the literature. We derive a value for the CRRA parameter well within the range of estimates in the literature. That value implies a fair degree of consumption smoothing behavior. In order to obtain a long run expected stock return close to the sample mean over the period 1966–1998, we are led to adopt a technology that has increasing marginal returns to capital. This assumption has some limitations, as it may lead to some indeterminacy of equilibria (Benhabib Farmer, 1994). Explicitly modeling externalities, might resolve that issue.

Thus, further research could encompass a look at alternate specifications for the way growth is embodied into the model. For instance, one possible way is to incorporate public goods or human capital, in the production function. Another interesting extension for this research is to use this model to examine the presence of consumption smoothing behavior along a sustained growth path in other economies.

NOTES

1. Note that labor is implicitly part of formulation (2) as K_t represents the capital/labor ratio. When the coefficient is greater than *1*, the production function has increasing marginal returns to capital. In our case though, the marginal productivity of capital is bounded above.
2. See Barro and Sala-i-Martin (1994).
3. See Solow (1957) and Paquet and Robidoux (1997).
4. This capital accumulation equation involves the term $\frac{L_t}{L_{t+1}}$ as all variables are normalized in terms of labor units in two subsequent periods.
5. Ambler and Paquet (1994) use stochastic depreciation in the context of a real business cycle model.
6. This approach is consistent with the use-factor method of depreciation used in accounting. Another example is when oil or mining corporations revise their estimate of the amount of recoverable units, as a result of further discoveries. This makes the depreciation rate a function of the rate of extraction. Note that there are instances in which technological progress calls for the use of a new generation of capital goods and the scrapping of the old equipment (for example the switch from analog to digital networks). This is referred to in the literature as the process of creative destruction. Our model does not contradict that process, as technological progress is embodied in the inputs here, and our technological shocks reflect short-term adjustments of productivity outside the scope of long-term productivity growth.
7. See for example Hansen-Singleton (1983).
8. We use a simple deterministic trend. Controversies abound on the complexity of the relationship between trend and cycle components. See Canova (1998). Our approach is similar to King, Plosser and Rebelo's (1988). In their real business cycle model, they define a stationary equilibrium for the detrended economy. They work from a certainty equivalence perspective. They posit a particular stochastic process for the random shocks and replace the sequence of random shocks by their conditional expectations.
9. Even though the probability of obtaining such a sequence is extremely small, it is still a useful concept as it implies the known stylized facts about actual economic time series. Nelson and Plosser (1982) have given evidence that macro time series have important stochastic trends. Our notion of steady state does not contradict these findings as it looks at expected trends.
10. In the case where $\alpha > 1$, we know that the technology has increasing marginal returns to capital. Thus the first order conditions might describe a minimum rather than a maximum. But in fact, a simple argument shows that this cannot be the case. The reason is that in our steady state, the marginal productivity of capital has reached its peak, and from the consumer's standpoint this maximizes the growth of consumption over time.
11. This condition is analogous to King-Plosser and Rebelo (1988).
12. Note that the variance σ^2 and the CRRA parameter are inversely related.
13. A method used by King-Plosser and Rebelo (1988). As the sequence of realizations of consumption and output levels are bounded in the appropriate sup-norm topology, paralleling an argument from Danthine-Donaldson (1981), we deduce that the

optimal policy $e(y_t)$ is continuous. We will assume that this optimal policy is actually differentiable.

14. The value $(1+\alpha)$ is the elasticity of consumption with respect to income.

15. The formula is $TFP = Y/[s_L L + (1 - s_L)K]$, where K represents the capital stock (in their notation), and s_L is the income share of labor.

16. The formula is

$$\ln\left(\frac{\theta_{t+1}}{\theta_t}\right) = \ln\left(\frac{Y_{t+1}}{Y_t}\right) - \left[\bar{s}_{Lt}\ln\left(\frac{L_{t+1}}{L_t}\right) + (1-\bar{s}_{Lt})\ln\left(\frac{K_{t+1}}{K_t}\right)\right]$$

where $\bar{s}_{Lt}$ is the average share of labor between period t and $t+1$.

17. This is within the range of usual values from 8.4% to 10% (see Ambler and Paquet 1994).

18. If we were to use a stochastic depreciation rate, as we did for the stock returns regression, the adjusted $R^2 = 0.61$ for the first regression (Panel A), and $R^2 = 0.77$ for the second regression (Panel B). The parameter is equal to 4.05. The implied long run expected stock return is 8%, and the CRRA parameter is 3.08.

19. The formula is

$$\ln\left(\frac{\theta_{t+1}}{\theta_t}\right) = \ln\left(\frac{Y_{t+1}}{Y_t}\right) - \left[\bar{s}_{Lt}\ln\left(\frac{W_{t+1}}{W_t}\right) + (1-\bar{s}_{Lt})\ln\left(\frac{R_{t+1}}{R_t}\right)\right]$$

where $\bar{s}_{Lt}$ is the average share of labor between period t and $t+1$ and W_t is the real wage.

20. Our model does not generate a large enough expected return for smaller αs. On the other hand, with the values shown for α and the average depreciation rate, we are able to generate a long-term expected return of 7.75%, which is close to the sample mean of 7.97%. Having increasing marginal returns to capital might lead to the indeterminacy of equilibria as explored in Benhabib and Farmer (1994). This means that our steady state might not be globally stable, when the economy is shocked away from it.

21. OLS Regressions for other quarters are slightly weaker, and still confirm the same significance levels for our variables.

22. Shawky and Peng (1995) incorporate technological shocks in a standard exogenous growth model. They express the expected return on capital assets as a function of relative growth in capital stocks, labor and Solow residuals. They find that labor growth is highly significant. Our analysis differs by focusing near the steady state. We find that labor growth does not play a significant role, due to the correlation between labor growth and capacity utilization.

23. Table 3 shows regressions for the processes hypothesized for the dynamics of the consumption rate, of labor growth and capacity utilization. We find an adjusted R^2 of 0.72 for the relationship (18) between the log of the rate of consumption and the log of the detrended real GDP (see Table 3, panel A). The relationship between effective capacity utilization and other variables expressed in (23) gives an adjusted R^2 of 0.66 (Panel D).

24. This in part is due to our assumptions made on the elasticity α of the production function. It is important to point out that all regression results are *insensitive* to the choice of the initial value for the capital stock in 1959.

REFERENCES

Ambler, S., & Paquet, A. (1994). Stochastic depreciation and the business cycle. *International Economic Review*, *35*(1), 101–116.

Balvers, R. J., Cosimano, T. F., & McDonald, B. (1990). Predicting stock returns in an efficient market. *Journal of Finance*, *45*, 1009–1128.

Barro, R. J. (1998). Notes on growth accounting. NBER Working Paper 6654.

Barro, R. J., & Sala-i Martin, X (1994). *Economic Growth*. McGraw-Hill.

Baumol, W. J., Batey Blackman, S. A., & Wolff, E. N (1991). *Productivity and American leadership*. The MIT Press.

Benhabib, J., & Farmer, R. E. A. (1994). Indeterminacy and increasing returns. *Journal of Economic Theory*, *63*, 19–41.

Breeden, D (1979). An intertemporal asset-pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, *7*, 265–296.

Brock, W. A., & Mirman, L. J. (1972). Optimal economic growth and uncertainty: The discounted case. *Journal of Economic Theory*, *4*, 479–513.

Campbell, J. Y., & Shiller, R. (1988). The dividend price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, *1*, 195–228.

Canova, F. (1998). Detrending and business cycle facts. *Journal of Monetary Economics*, *41*, 475–512.

Ceccetti, S. G., Lam, P., & Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review*, *80*, 398–418.

Chen, N. F., Roll R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, *59*(3), 383–404.

Cochrane, J. H. (1991). Production-based asset pricing and the link between returns and economic fluctuations. *Journal of Finance*, *46*, 209–231.

Danthine, J. P., & Donaldson, J. B. (1981). Stochastic properties of fast vs. slow growing economies. *Econometrica*, *49*(4), 1007–1033.

Fama, E. F. (1990). Stock returns, expected returns and real activity. *Journal of Finance*, *45*, 1089–1108.

Fama, E. F., & French, K. R. (1988a). Permanent and temporary components of stock prices. *Journal of Political Economy*, *96*, 246–273.

Fama, E. F., & French, K. R. (1988b). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*, 3–25.

Faugère, C. (1993). Essays on the dynamics of technical progress. Unpublished Ph.D. thesis, University of Rochester.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*, 3–29.

Hansen, L. P., & Singleton, K. J. (1983). Stochastic consumption, risk aversion and the temporal behavior of asset returns. *Journal of Political Economy*, *91*(2), 249–265.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, *17*, 357–390.

King, R. J., Plosser, R. F., & Rebelo, S. T. R. (1988). Production, growth and business cycle: The basic neoclassical model. *Journal of Monetary Economics*, *2/3*, 195–232.

Kydland, F., & Prescott, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, *50*, 1345–1370.

Lo, A. W., & McKinlay, C. A. (1988). Stock prices do not follow a random walk: Evidence from a new specification test. *Review of Financial Studies*, *1*, 41–66.

Lucas, R. (1978). Asset prices in an exchange economy. *Econometrica*, *46*, 1429–1444.

Merton, R. (1973). An intertemporal asset-pricing model. *Econometrica*, *41*, 867–887.

Paquet A., & Robidoux, B. (1997). Issues on the measurement of the Solow residuals and the testing of its exogeneity: a tale of two countries. Working paper No. 51, CREFE.

Restoy, G. B., & Rockinger, G. M. (1994). On stock market returns and returns on investment. *Journal of Finance*, *49*, 543–556.

Shawky, H., & Peng, Y. (1995). Expected Stock Returns, Real Business Activity and Consumption Smoothing. *International Review of Financial Analysis*, *4*(2/3), 143–154.

Solow, R. M. (1957). Technical change and aggregate production function. *The Review of Economics and Statistics*, *39*, 312–320.

A NOTE ON THE MARKOWITZ RISK MINIMIZATION AND THE SHARPE ANGLE MAXIMIZATION MODELS

Chin W. Yang, Ken Hung and Felicia A. Yang

ABSTRACT

The equivalence property between the angle-maximization portfolio technique and the Markowitz risk minimization model is proved. Via reciprocal and monotonic transformation, they can be made equivalent with or without different types of short sale. Since the Markowitz portfolio model is formulated in the standard convex quadratic programming, the equivalence property would enable us to apply the same well-known mathematic properties to the angle maximization model and enjoy the same convenient computational advantage of the quadratic program (e.g. Markowitz's critical line algorithm).

1. INTRODUCTION

Surprisingly, one of the earlier applications of quadratic programming is the Markowitz portfolio selection model (1952 and 1959) upon which modern investment theory is built.[1] Perhaps it is one of the least understood models in finance literature since his invention primarily falls within the domain of operations research (Markowitz, 1956). Nonetheless, the portfolio selection models have advanced beyond its prototype (see Sharpe, 1963, 1964; Lintner, 1965; Mossin, 1956; Ross, 1976; Markowitz & Perold, 1981; Markowitz, 1987). The purpose of this paper is to: (1) compare the angle-maximization

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 21–29.

ISBN: 0-7623-0887-7

technique as described in Elton and Gruber (1995) with the Markowitz quadratic programming models in order to propose a condition under which they can be made equivalent; and (2) examine the efficient frontiers of the commonly used portfolio selection model under different assumptions. Hence, we point out a common error prevalent in most investment theory texts.[2] According to efficient market hypothesis, security prices fully reflect all available information with zero lag. But in reality, there are imperfections and lack of foresight (Bernstein, 1999). As a result, security price is elusive and an accurate forecast begs serious scientific efforts. Furthermore, as pointed out by Miller (1999), mean and variance of stock returns possess rather different statistical characteristics. For instance, past data with smaller intervals many times provide reasonable estimates for portfolio risk (variance and covariance), while average return of past few years is a poor estimate of expected return. Despite that, econometric methodologies by Engle (1982) and start to provide a better understanding of error variance in the presence of clustered forecasting error. It is to be noted that the variance-based and mean-based models cannot be treated as identical as the duality theorem in linear programming.[3] The equivalence property enables us to choose one to which a robust estimate is more amenable.

2. THE EQUIVALENCE OF THE PORTFOLIO SELECTION MODELS

The essence of mean variance portfolio theory can be formulated most conveniently in the standard Markowitz portfolio selection model:

$$\underset{x_i}{\text{Minimize}} \quad v = \sum_{i \in I} s_{ii} x_i^2 + \sum_{i \in I} \sum_{\substack{j \in J \\ j \neq i}} x_i x_j s_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{i \in I} r_i x_i \geq k \tag{2}$$

$$\sum_{i \in I} x_i = 1 \tag{3}$$

$$x_i \geq 0 \qquad \forall i \in I \tag{4}$$

where v stands for the weighted variance and covariance of the portfolio returns; s_{ii} and s_{ij} are sample variance of the return of security i and covariance of returns between security i and j respectively; x_i denotes the proportion of investment on security i; r_i is the expected return of security i; k denotes the minimum expected portfolio return; and I, J are a set of positive integers (1, . . . , n). The properties of the minimization problem are well known (Markowitz, 1959), and the efficient computational algorithms are discussed by Tucker and Dafaro (1975), Pang (1980), Schrage (1986), and Markowitz (1956, 1987).

A variation of the Sharpe model (1964) used in most investment texts is formulated to maximize the slope or angle of the capital market line with a given risk-free rate or r_f:

$$\underset{x_i}{\text{Maximize}} \quad \theta = \sum_{i\in I} x_i(r_i - r_f) \Bigg/ \left(\sum_{i\in I} x_i^2 s_{ii} + \sum_{i\in I} \sum_{\substack{j\in J \\ j\neq i}} x_i x_j s_{ij} \right)^{1/2}$$

$$= u/v^{1/2} \tag{5}$$

$$\text{subject to} \quad \sum_{i\in I} x_i = 1 \tag{6}$$

$$x_i \geq 0 \qquad \forall i \in I \tag{7}$$

where: $u = \sum x_i(r_i - r_f)$.

Notice that these two models appear to be rather different on the surface. The maximization of a risk-adjusted portfolio mean return model is, in general, not equivalent to the minimization of the portfolio variance covariance problem. The identification of r_f is needed in order to solve for the optimum investment proportions whereas a minimum portfolio return of k must be specified to solve the Markowitz problem. If the value of k is too low, the constraint (Eq. (2)) will not be active, i.e. a rather uninteresting case in which the Markowitz efficient frontier is vertical occurs. Henceforth, we shall focus on the case of an active constraint: $\sum_{i\in I} r_i x_i = k$. The following proposition is stated regarding the equivalence of the two portfolio selection models.

Proposition 1: The angle maximization portfolio problem and the Markowitz risk minimization model can be made equivalent for a constant u: more specifically if r_f is independent of k.

Proof: Constraint (2) can be rewritten as

$$\sum_{i\in I} x_i(r_i - r_f) = k - r_f = u \tag{8}$$

in which k is normally predetermined by portfolio managers and r_f is usually approximated by a short-term bond (e.g. TB).

Hence, for a given level of u in Eq. (5), the objective function of the angle maximization model becomes $(k - r_f)/v^{1/2}$ where k is taken from the corresponding Markowitz portfolio model and r_f is a constant. It follows immediately from the property of the reciprocal transformation that maximizing $(k - r_f)/v^{1/2}$ is equivalent to minimizing $v^{1/2}$ for a constant $k - r_f$ of Eq. (5). Note that both models at the given level of $k - r_f$ contain the identical information on constraints since the information embedded in Eq. (2) is transplanted to Eq. (5). Furthermore, it is well known that the monotonic transformation of an objective function preserves the optimum solution. That is to say, for a given set of constraints, minimizing v gives rise to the identical set of optimum x's as that of minimizing $v^{1/2}$ or log v. This completes the proof.[4]

The proof of this equivalence is not simply a mathematic curiosity; it has some very important implications. Since the angle maximization model is relatively complicated, the equivalence property guarantees that it shares the same property as the original Markowitz model. In addition, the equivalence property applies to the model with two different types of short sales. The freedom for shortsell allows investors to benefit from stocks with negative expected return in an integrated portfolio approach (Jacobs et al., 1999). The first type, or the Black (1972) model, simply drops the nonnegativity constraint of Eqs (4) and (7). The other short sale possibility is to replace Eqs (3) and (6) with $\sum_{i \in I} |x_i| = 1$ (Lintner, 1965). To verify the equivalence of these portfolio models, we employ an excellent numerical example (Elton & Gruber 1995, Ch. 4) with three securities with $s_{11} = 0.0036$, $s_{22} = 0.0009$, $s_{33} = 0.0225$, $s_{12} = 0.0009$, $s_{23} = s_{13} = 0.0018$, $r_1 = 0.14$, $r_2 = 0.08$, $r_3 = 0.2$ and $r_f = 0.05$. By substituting these values into Eqs (5)–(7) via a nonlinear programming package,[5] we obtain $x = 0.778$, $x = 0.055$, $x = 0.167$ and $u = 0.097$. This is exactly the same as the answer derived by Elton and Gruber. We then substitute $k = u + r_f = 0.147$ along with the parameter values into Eqs (1)–(4). The identical solution is found as anticipated. Alternatively, we can substitute $u = k - r_f$ into Eq. (5) with other parameter values and again we obtain an identical solution. This relation is also verified in both cases of short sales. It is to be pointed out that the risk-free rate r_f is not explicitly considered in the Markowitz model. An optimal selection on a given efficient frontier requires the value of r_f in the final analysis. The risk-free rate r_f, among other things, represents an evolution of finance history from Markowitz to Sharpe. However, the equivalence property holds if $u = k - r_f$ is constant. If the minimum portfolio rate of return k is a function of r_f, then u is not constant. In the case when r_f is higher, it may not be unusual for portfolio managers to set a higher k value. Under this circumstance, the Markowitz minimization model is not equivalent to the Sharpe maximization model. As long as weighted sum of expected returns equals k, r_f and k can be related. That is the expected return has three components: risk-free return, unconditional expected excess return without any special information, and alpha or the difference between conditional and unconditional expected return (Grinold, 1999, p. 11). In this case, the two models may be somewhat similar, but are not equivalent as one is not the mathematical duality of the other. In the next section, we examine the shape of the efficient frontier using the same example by varying r_f from 0.005 to 0.075 (Table 1) in the case of the portfolio model without short sale.[6] We focus on solutions ignoring short sales in this paper because the shape of the efficient frontier can be readily examined.

3. THE EFFICIENT FRONTIER OF THE PORTFOLIO SELECTION MODELS

The equivalence property proven in the previous section leads immediately to the result that both models must have identical efficient frontiers since they have identical x's. The efficient frontier of the Markowitz portfolio selection model is normally seen to be a concave curve over $k - v^{1/2}$ as is drawn[7] in virtually every investment text except for Markowitz (1959, 1987). While it is an established "ritual" in chapters concerning portfolio analysis, it does not have to be concave over the $k - v^{1/2}$ space. An examination on the very example presented in Elton and Gruber (1987) indicates immediately that the efficient frontier is not necessarily concave even in the case where short sales are disallowed.[8] We prove this result in the following proposition:

Proposition 2: The efficient frontier of the portfolio selection models without short sale may not be concave over the $k - v^{1/2}$ space.

Proof: Markowitz proves that in the absence of a singular variance-covariance matrix the efficient frontier is concave or piecewise parabolic over the $k - v$ space (1959, 1987). However, the concavity of the efficient frontier curve over the $k - v$ space may not be preserved over the $k - v^{1/2}$ space. That is, $(d^2v/dk^2) > 0$ does not necessarily imply $(d^2v^{1/2}/dk^2) > 0$. This can be shown as below:

$$dv/dk = d[(v^{1/2})^2]/dk = 2v^{1/2}dv^{1/2}/dk \tag{9}$$

and

$$d^2v/dk^2 = d[(2v^{1/2} \bullet dv^{1/2})/dk]/dk = 2v^{1/2}(d^2v^{1/2}/dk^2) + 2(dv^{1/2}dk)^2 > 0 \tag{10}$$

Evidently, Eq. (9) holds via the composite function rule. As is proven by Markowitz (1959, 1987), the sign of Eq. (10) must be nonnegative (zero for a

Table 1. The Portfolio Selection Model Without Short Sale.

Risk-Free Rate r_f (%)	x_1	x_2	x_3	k	$v^{1/2}$
0.25	0.265	0.708	0.027	0.0991	0.0339
0.5	0.275	0.696	0.029	0.1	0.034
0.75	0.285	0.683	0.032	0.1009	0.0346
1.0	0.296	0.669	0.035	0.102	0.035
1.25	0.308	0.654	0.038	0.1031	0.0354
1.5	0.321	0.637	0.042	0.104	0.036
1.75	0.335	0.62	0.046	0.1056	0.0364
2.0	0.35	0.6	0.05	0.107	0.037
2.25	0.367	0.579	0.055	0.1086	0.0377
2.5	0.385	0.555	0.059	0.11	0.038
2.75	0.406	0.529	0.065	0.1122	0.039
3.0	0.429	0.5	0.071	0.114	0.04
3.25	0.454	0.468	0.078	0.1166	0.0415
3.5	0.483	0.431	0.086	0.119	0.043
3.75	0.515	0.39	0.095	0.1225	0.044
4.0	0.553	0.342	0.105	0.126	0.046
4.25	0.596	0.287	0.117	0.1298	0.048
4.5	0.646	0.223	0.131	0.134	0.051
4.75	0.706	0.147	0.147	0.14	0.054
5.0	0.778	0.056	0.167	0.147	0.058
5.25	0.82	0	0.18	0.1508	0.06066
5.5	0.817	0	0.183	0.151	0.0607
5.75	0.814	0	0.186	0.1511	0.0609
6.0	0.811	0	0.189	0.1513	0.061
6.25	0.808	0	0.192	0.1515	0.0611
6.5	0.805	0	0.195	0.15173	0.0613
6.75	0.8	0	0.2	0.152	0.0614
7.0	0.797	0	0.203	0.1522	0.0616
7.25	0.793	0	0.207	0.1525	0.0618
7.5	0.788	0	0.212	0.153	0.062

singular variance-covariance matrix). As a result, the sign of $d^2(v^{1/2})/dk^2$ is arbitrary since $v^{1/2}$ is positive. That is, the sign of its inverse or $dk^2/d^2(v^{1/2})$ (the curvature of the efficient frontier in most investment texts) can not be determined without additional qualifications.

It may be neither concave nor convex over the $k - v^{1/2}$ space using the Elton and Gruber example. While the concavity of the efficient frontier over the $k - v^{1/2}$ space is frequently drawn for the capital market equilibrium, it is generally not true. Most investment texts, however, have shown concave efficient frontiers over the $k - v^{1/2}$ space as the typical case. A nonconcave efficient frontier could occur if $dk^2/d^2(v^{1/2})$is positive or moderately positive (not its inverse): it cannot be significantly positive or condition (10) is violated. Such an efficient frontier suggest that $dk/d(v^{1/2})$increase at increasing rates: more return is needed for each marginal return/risk ratios.

4. CONCLUDING REMARKS

In this paper we, for the first time, prove the equivalence property between the angle-maximization portfolio technique and the Markowitz risk minimization model. Via reciprocal and monotonic transformation, they can be made equivalent with or without different types of short sale. Since the Markowitz portfolio model is formulated in the standard convex quadratic programming, the equivalence property would enable us to apply the same well-known mathematic properties to the angle maximization model and enjoy the same convenient computational advantage of the quadratic program (e.g. Markowitz's critical line algorithm). However, if k and r_f are functionally related in empirical practice, these two models are not duality of each other. Consequently, they are two different portfolio selection models with different orientations. Finally, the efficient frontier is not in general concave over the $k - v^{1/2}$ space especially for investment involving increasing marginal return/risk ratios.

NOTES

1. The application of the Markowitz portfolio model goes beyond stock markets (e.g. the wine market investment model by Labys et al., 1981).
2. It is worth mentioning that the most rigorous treatment of the portfolio selection problem is still found in Markowitz (1987).
3. Duality theorem in quadratic programming is rather different from that in linear programming which we do not intend to pursue in this paper.
4. The computer simulations are made to verify all the results in this paper, using the example by Elton and Gruber (1995). They are available upon request.
5. Our simulation is based on the package of GINO (Schrage et al., 1986), an efficient package for small-sized problems. Benninga (2000) used Excel to calculate and graph efficient frontiers.
6. Efficient frontiers with short sales are rather complicated and are not necessarily continuous. We defer discussion of the short sale conditions to a future paper.
7. Unlike other textbooks, Markowitz's texts correctly graph v (vertical axis) against k (horizontal axis). The concavity or piecewise parabolic property of the efficient frontier over the $k - v$ space has important implications: a potential multiple equilibria in the capital market.

8. The $k - v^{1/2}$ combination in Table 1 is not efficient frontier. We thank one of the referees for pointing this out.

REFERENCES

Benninga, S. (2000). *Financial Modeling*. Cambridge, MA: The MIT Press.

Bernstein, P. L. (1999). A New Look at the Efficient Market Hypothesis. *The Journal of Portfolio Management*, (Winter), 1–2.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *Journal of Business*, (July), 444–445.

Cragg, J. (1982). Estimation and Testing in Testing Time Series Regression Model with Heteroscedasticity. *Journal of Econometrics*, *20*, 135–157.

Elton, E. J., & Gruber, M. J. (1995). *Modern Portfolio Theory and Investment Analysis* (5th ed.). New York: John Wiley & Sons.

Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica*, *50*, 987–1008.

Grinold, R. C. (1999). Mean-Variance and Scenario-Based Approaches to Portfolio Selection. *The Journal of Portfolio Management*, (Winter), 10–21.

Jacobs, B. I., Levy K. N., & Starer D. (1999). Long-Short Portfolio Management: An Integrated Approach. *The Journal of Portfolio Management*, (Winter), 23–32.

Labys, W. C., Cohen, B. C., & Yang, C. W. (1981). The Rational Cavist. *European Review of Agricultural Economics*, *8*, 519–525.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, *47*(February), 13–37.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, *7*, 77–91.

Markowitz, H. (1956). The Optimization of a Quadratic Function Subject to Linear Constraints. *Naval Research Logistics quarterly*, *3*, 111–133.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc..

Markowitz, H. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Oxford, U.K.: Basil Blackwell.

Markowitz, H. (1991). Foundation of Portfolio Theory. *Journal of Finance*, *SLVI*(2), 469–477.

Markowitz, H., & Perold, A. F. (1981). Sparsity and Piecewise Linearity in Large Portfolio Optimization Problems. In: I. S. Duff (Ed.), *Sparse Matrices and Their Uses*, New York: Academic Press.

Markowitz, H. M. (1991) *Portfolio Selection: Efficient Diversification of Investments* (2nd ed.). Cambridge, MA: Basil Blackwell.

Markowitz, H. M. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Oxford: Basil Blackwell Ltd.

Markowitz, H. M., & Todd, P. (2000). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. With Chapter 13 by G. Peter Todd. First published in 1987 by Basil Blackwell. Revised reissue by Frank Fabozzi and Associaes, New Hope, PA.

Miller, M. H. (1999). The History of Finance. *The Journal of Portfolio Management*, (Summer), 95–101.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrika*, *34* (October), 768–783.

Pang, J. S. (1980). A New and Efficient Algorithm for a Class of Portfolio Selection Problem. *Operation Research*, *28*, 754–767.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, *13*, 341–360.

Schrange, L. (1986). *Linear, Integer and Quadratic Programming with LINDO* (3rd ed.). Palo Alto, CA: The Scientific Press.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Condition of Risk. *Journal of Finance*, (September), 425–442.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies*, (February), 65–86.

Tucker, J., & Dafaro, C. (1975). A Simple Algorithm for Stone's Version of the Portfolio Selection Problem. *Journal of Financial and Quantitative Analysis*, *10*(5), 559–570.

This Page Intentionally Left Blank

OPTIMAL HEDGE RATIOS AND TEMPORAL AGGREGATION OF COINTEGRATED SYSTEMS

Donald Lien and Karyl Leggio

ABSTRACT

This note considers optimal ratios for different lengths of hedging horizon when the highest frequency data is generated by a cointegrated system. It is found that, after reparameterization, a temporal aggregation of cointegrated systems remains a cointegrated system. This result provides a convenient method to estimate n-day hedge ratio for any integer n. The only remaining issue concerns the possible incorrect lag selections. Empirical results from ten futures contracts however indicate lag selections have no effect on the estimated hedge ratios.

1. INTRODUCTION

Many recent studies found that spot and futures prices contain a unit root whereas the basis (i.e. the difference between futures and spot prices) is stationary. Following Engle and Granger (1987), spot and futures returns should be described by a cointegrated system in which the lagged basis plays an important role. In contrast, the usual vector autoregression (VAR) model does not explicitly include the lagged basis term. Empirical research documented that the cointegrated system not only provides a better description of the behavior of spot and futures returns, it also produces a hedge ratio with

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 31–40.

ISBN: 0-7623-0887-7

better hedging performance than the ratio derived from the VAR model. For examples, see Chou, Denis and Lee (1993) and Ghosh (1993). Analytical comparison results are provided in Lien (1996).

In the absence of serial correlation of the error terms, the optimal hedge ratio derived from the VAR model is the same for different hedge horizons. This property is no longer valid when a cointegrated system is applied to generate the optimal hedge ratio. Thus, the n-day hedge ratio differs from the m-day hedge ratio whenever $m \neq n$. Suppose that the cointegrated system provides a valid description of the 1-day return series. Then an n-day hedge ratio can be calculated from this system. The calculation is however quite cumbersome.

An alternative method is to rely upon a cointegrated system for the n-day return series. For example, Chou, Denis and Lee (1993) assumed that, for every hedge horizon, there is a cointegrated system that provides a valid description of the data. They estimated hedge ratios for different hedge horizons and found the ratios vary across different horizons. The current paper validates this approach by showing that, after reparameterization, the aggregation of a cointegrating system produces a new cointegrating system. Thus, n-day hedge ratios can be derived more conveniently from the aggregated system.

In an empirical study, however, the one-day model may be incorrectly specified due to incorrect lag selections. The n-day hedge ratio calculated therefore is unreliable. Alternatively, one can select lags directly from the n-day model. Using data from ten futures contracts, we found that lag selections have no effect on the estimated hedge ratios.

2. OPTIMAL n-DAY HEDGE RATIO

Let s_t and f_t denote the logarithms of spot and futures prices at day t. The 1-day returns for spot and futures are calculated by $\Delta_1 s_t = s_t - s_{t-1}$ and $\Delta_1 f_t = f_t - f_{t-1}$, respectively. Similarly, the n-day returns for spot and futures are measured by $\Delta_n s_t = s_t - s_{t-n}$ and $\Delta_n f_t = f_t - f_{t-n}$, respectively. The basis at day t is defined as $z_t = f_t - s_t$, the difference between futures and spot prices. Fama and French (1987) documented the forecasting power of the basis for spot and futures prices. We assume that spot and futures prices are described by the following cointegrated system:

$$\Delta_1 s_t = -\alpha z_{t-1} + \varepsilon_{st}, \quad (1)$$

$$\Delta_1 f_t = \beta z_{t-1} + \varepsilon_{ft}; \quad (2)$$

where $\{\varepsilon_{st}, \varepsilon_{ft}\}$ is a sequence of white noise such that $Var(\varepsilon_{st}) = \sigma_{ss}$, $Var(\varepsilon_{ft}) = \sigma_{ff}$, and $Cov(\varepsilon_{st}, \varepsilon_{ft}) = \sigma_{sf}$. We expect both spot and futures prices to converge to the long-run equilibrium at which the spot price equals the futures price. Thus, both α and β are expected to be negative.

Consider a one-period (from t to t + n) framework. The length of the period is n days, $n \geq 1$. A hedger with a non-tradable spot position engages in futures trading to reduce risk exposure. Let $-h(n)$ denote his futures position. At the end of the period the return from the hedged portfolio is $\Delta_n s_{t+n} - h(n)\Delta_n f_{t+n}$. The optimal hedge ratio, $h^*(n)$ is chosen to minimize the conditional variance, $Var(\Delta_n s_{t+n} - h(n)\Delta_n f_{t+n} | I_t)$, where I_t is the information available at time t. It can be easily shown that

$$h^*(n) = Cov(\Delta_n s_{t+n}, \Delta_n f_{t+n} | I_t)/Var(\Delta_n f_{t+n} | I_t). \tag{3}$$

For the optimal 1-day hedge ratio, Eqs (1) and (2) lead directly to the following:

$$h^*(1) = \sigma_{sf}/\sigma_{ff}. \tag{4}$$

To derive the optimal n-day hedge ratio, we rewrite Eqs (1) and (2) as follows:

$$\begin{bmatrix} s_t \\ f_t \end{bmatrix} = \begin{bmatrix} 1+\alpha & -\alpha \\ -\beta & 1+\beta \end{bmatrix} \begin{bmatrix} s_{t-1} \\ f_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{st} \\ \varepsilon_{ft} \end{bmatrix} \tag{5}$$

By iteration, we have

$$\begin{bmatrix} s_t \\ f_t \end{bmatrix} = \begin{bmatrix} 1+\alpha & -\alpha \\ -\beta & 1+\beta \end{bmatrix}^n \begin{bmatrix} s_{t-n} \\ f_{t-n} \end{bmatrix} + \sum_{j=0}^{n-1} \begin{bmatrix} 1+\alpha & -\alpha \\ -\beta & 1+\beta \end{bmatrix}^j \begin{bmatrix} \varepsilon_{s,t-j} \\ \varepsilon_{f,t-j} \end{bmatrix}. \tag{6}$$

As a consequence, conditional variances and conditional covariance for spot and futures prices are determined as follows:

$$Var\begin{bmatrix} s_t \\ f_t \end{bmatrix} = \sum_{j=0}^{n-1} \begin{bmatrix} 1+\alpha & -\alpha \\ -\beta & 1+\beta \end{bmatrix}^j \begin{bmatrix} \sigma_{ss} & \sigma_{sf} \\ \sigma_{sf} & \sigma_{ff} \end{bmatrix} \begin{bmatrix} 1+\alpha & -\beta \\ -\alpha & 1+\beta \end{bmatrix}^j. \tag{7}$$

Denote the above matrix by M_n and the (i, j) element of the matrix by $M_n(i, j)$. Then the optimal n-day hedge ratio is $M_n(1, 2)/M_n(2, 2)$.

From Eq. (7), we conclude that, in general, the optimal hedge ratio varies as the hedging horizon changes. The direction is determined by all the parameters:

α, β, σ_{ss}, σ_{sf}, and σ_{ff}. For example, consider $n = 2$. After algebraic manipulation, the optimal 2-day hedge ratio is

$$h^*(2) = \frac{-\beta(1+\alpha)\sigma_{ss} + [1 + \alpha\beta + (1+\alpha)(1+\beta)]\sigma_{sf} - \alpha(1+\beta)\sigma_{ff}}{\beta^2\sigma_{ss} - 2\beta(1+\beta)\sigma_{sf} + [1 + (1+\beta)^2]\sigma_{ff}} \tag{8}$$

3. EFFICIENT FUTURES MARKET

If the futures market is efficient such that the past information cannot help to improve the forecast, then $\beta = 0$. By induction, we can derive the following:

$$\begin{bmatrix} 1+\alpha & -\alpha \\ 0 & 1 \end{bmatrix}^j = \begin{bmatrix} (1+\alpha)^j & 1-(1+\alpha)^j \\ 0 & 1 \end{bmatrix}. \tag{9}$$

After algebraic manipulations, we have

$$h^*(n) = \left\{ \sum_{j=0}^{n-1} (1+\alpha)^j \sigma_{sf} + [1-(1+\alpha)^j]\sigma_{ff} \right\} \Big/ (n\sigma_{ff}), \tag{10}$$

or

$$h^*(n) = h^*(1) + \left\{ \frac{(1+\alpha)^n - 1}{n\alpha} - 1 \right\} [h^*(1) - 1]. \tag{11}$$

Let d(n) denote the term inside the bracket. Because d(n) is always negative, $h^*(n) < h^*(1)$ if $h^*(1) > 1$ and vice versa. It can also be shown that d(n) is a decreasing function of n (see Appendix A). Consequently, $h^*(n)$ decreases with increasing n when $h^*(1) > 1$. If $h^*(1) < 1$, then $h^*(n)$ increases as n increases. Moreover, as n approaches infinity, $h^*(n)$ approaches one. The optimal hedge ratio can be characterized as follows. Suppose that the one-period optimal ratio is smaller than one. The optimal ratio increases toward one as the length of the hedge period increases. On the other hand, when the one-period hedge ratio is larger than one, the optimal ratio decreases towards one as the length of the hedge period increases. In either case, the unit hedge ratio is the optimal strategy provided that the length of the hedging period is sufficiently large.

4. HEDGE RATIOS CALCULATED FROM AGGREGATE MODELS

To describe the statistical behavior for n-day returns, we aggregate Eqs (1) and (2), respectively, to derive the following:

$$\Delta_n s_t = -\alpha \sum_{j=1}^{n} z_{t-j} + \sum_{k=0}^{n-1} \varepsilon_{s,t-k} = -\alpha z_{t-n} + u_{st}, \tag{12}$$

$$\Delta_n f_t = \beta \sum_{j=1}^{n} z_{t-j} + \sum_{k=0}^{n-1} \varepsilon_{f,t-k} = \beta z_{t-n} + u_{ft}, \tag{13}$$

where $u_{st} = \varepsilon_{st} + \sum_{j=1}^{n-1} (\varepsilon_{s,t-j} - \alpha z_{t-j})$ and $u_{ft} = \varepsilon_{ft} + \sum_{j=1}^{n-1} (\varepsilon_{f,t-j} - \beta z_{t-j})$. Although Eqs (12) and (13) appear to be similar to Eqs (1) and (2), the main difference lies in the fact that z_{t-n} is correlated with u_{st} and u_{ft}. For example, $\mathrm{Cov}(\Delta_n s_{t+n}, \Delta_n f_{t+n} | z_{t-n}) \neq \mathrm{Cov}(u_{st}, u_{ft})$ and $\mathrm{Var}(\Delta_n f_{t+n} | z_{t-n}) \neq \mathrm{Var}(u_{ft})$.

Let w_{st} and w_{ft} be the residual of regressing u_{st} and u_{ft} on z_{t-n}, respectively. Then Eqs (12) and (13) can be rewritten as:

$$\Delta_n s_t = -\alpha^* z_{n-t} + w_{st}, \tag{14}$$

$$\Delta_n f_t = \beta^* z_{n-t} + w_{ft}. \tag{15}$$

Following Myers and Thompson (1989), the optimal hedge ratio can be derived from the following regression model:

$$\Delta_n s_t = a(n) z_n + b(n) \Delta_n f_t + w_{nt}. \tag{16}$$

That is, $b(n) = \mathrm{Cov}(\Delta_n s_t, \Delta_n f_t | z_{t-n}) / \mathrm{Var}(\Delta_n f_t | z_{t-n})$ is the optimal n-day hedge ratio. By the least squares method, we have

$$\begin{bmatrix} a(n) \\ b(n) \end{bmatrix} = \begin{bmatrix} Var(z_{n-t}) & Cov(z_{n-t}, \Delta_n f_t) \\ Cov(z_{n-t}, \Delta_n f_t) & Var(\Delta_n f_t) \end{bmatrix}^{-1} \begin{bmatrix} Cov(z_{n-t}, \Delta_n s_t) \\ Cov(\Delta_n f_t, \Delta_n s_t) \end{bmatrix}. \tag{17}$$

Note that Eqs (1) and (2) imply that $\Delta_1 s_t = a(1) z_{t-1} + b(1) \Delta_1 f_t + w_{1t}$ such that w_{1t} is stochastically independent of z_{t-1} and $\Delta_1 f_t$. Consequently, the least squares method provides an unbiased estimator of the optimal hedge ratio. If we aggregate the equation over time, we derive $\Delta_n s_t = a(1) z_{t-n} + b(1) \Delta_n f_t + x_{nt}$ where x_{nt} is correlated with z_{t-n} and $\Delta_n f_t$. Consequently, the least squares estimators, a(n) and b(n), are respectively biased estimators of a(1) and b(1).

To proceed further, we subtract Eq. (1) from Eq. (2) to derive

$$z_t = (1 + \alpha + \beta)z_{t-1} + (\varepsilon_{ft} - \varepsilon_{st}). \tag{18}$$

Let $\gamma = 1 + \alpha + \beta$ and let $\varepsilon_t = \varepsilon_{ft} - \varepsilon_{st}$. Then

$$z_t = \sum_{j=0}^{\infty} \gamma^j \varepsilon_{t-j}. \tag{19}$$

Using the fact $\Delta_n f_t = \Delta_1 f_t + \Delta_1 f_{t-1} + \ldots + \Delta_1 f_{t-n+1}$, Eqs (2) and (19) lead to:

$$\Delta_n f_t = \beta \sum_{k=1}^{n} \sum_{j=0}^{\infty} \gamma^j \varepsilon_{t-j-k} + \sum_{k=0}^{n-1} \varepsilon_{f,t-k}. \tag{20}$$

A similar expression can be derived for the n-day spot return:

$$\Delta_n s_t = -\alpha \sum_{k=1}^{n} \sum_{j=0}^{\infty} \gamma^j \varepsilon_{t-j-k} + \sum_{k=0}^{n-1} \varepsilon_{s,t-k}, \tag{21}$$

Equations (19)–(21) can be applied to calculate every element of the two matrices contained in Eq. (17). Appendix B derives the required variances and covariances for $n = 2$. After algebraic manipulation, we have $b(2) = h^*(2)$. Similarly, in an efficient futures market, $b(n) = h^*(n)$.

5. AN EMPIRICAL STUDY

The above section demonstrates a more convenient method to estimate the optimal n-day hedge ratio via Eq. (16). In empirical studies, the lagged spot and futures returns may be incorporated into the regression equation. It is likely that model selection criteria such as AIC (Akaike Information Criteria) may prescribe different lags for different hedge horizon lengths. We conducted an empirical study to examine this issue and its effect on optimal n-day hedge ratios.

We consider ten futures contracts: three currency futures contracts (British Pound (BP), Deustchmark (DM) and Japanese Yen (JY)), five commodity futures contracts (soy bean oil (BO), wheat (KW), crude oil (CL), corn (C) and cotton (CT)) and two stock index futures contracts (NYSE composite (YX) and S&P 500 (SP)). Daily spot and nearby futures prices were obtained from Commodity System, Inc. The sample period extends from January 1988 to June 1998. Nearby futures prices were constructed with contract rollover occurring about one week before maturity in most cases. The trading volume was used as a criterion in deciding the actual rollover date. Following convention, both s_t and f_t were defined as logarithmic prices. When calculating the differences in futures prices at the rollover, it was ensured that the two prices pertained to the same futures contract.

Statistical properties of these spot and futures returns are described in Lien, Tse, and Tsui (2001). Herein we discuss only the augmented Dickey-Fuller (ADF) test results. In each market, the null hypothesis of the price series being non-stationary cannot be rejected at the 5% level, whereas the null hypothesis of a unit root for the return series is rejected at the 1% level. Moreover, the null hypothesis of a unit root for the basis series is rejected at the 1% level in each market.

We applied Eq. (16) to estimate optimal hedge ratios after incorporating lagged spot and futures returns as explanatory variables. An equal number of lags were adopted for both returns. The optimal number of lags was determined by AIC. If the one-day model is correctly specified, then the same number of lags should apply to any n-day model. Otherwise, AIC should be applied case by case. Both scenarios are considered in this empirical study. Table 1 summarizes the estimation results. There is little difference (if any at all) between the two hedge ratios. Although the one-day model may be incorrectly specified, the misspecification has no effect on the estimated optimal hedge ratios.

Table 1. Optimal n-Day Hedge Ratios.

	Length of the Hedge Horizon		
Contract	1 day	2 day	3 day
British Pound	0.9301	0.9426 (0.9425)	0.9469 (0.9480)
Deutsche Mark	0.9531	0.9550 (0.9551)	0.9580 (0.9588)
Japanese Yen	0.9458	0.9554 (0.9555)	0.9585 (0.9607)
Soybean Oil	0.9618	0.9657 (0.9674)	0.9650 (0.9671)
Corn	0.8302	0.8192 (0.8198)	0.8144 (0.8146)
Cotton	0.7200	0.7342 (0.7342)	0.7349 (0.7349)
Wheat	0.4609	0.6605 (0.6605)	0.7097 (0.7097)
Crude Oil	0.8869	0.9297 (0.9297)	0.9477 (0.9477)
NYSE Composite	0.8170	0.8816 (0.8812)	0.9114 (0.9116)
S&P 500	0.8826	0.9221 (0.9227)	0.9431 (0.9439)

* For the two-day and three-day cases, the first hedge ratio is calculated with the number of lags determined by the one-day model. The second ratio is determined by AIC for each hedge horizon length.

With the exception of corn and soybean oil markets, the optimal hedge ratios are always smaller than one and increase as the length of the hedge horizon increases. These results are consistent with the prediction under an efficient futures market. In the three currency markets, hedge ratios are very stable across different hedge horizon length. For NYSE composite and S&P 500, increases in optimal hedge ratios are noticeable. Commodity markets present conflicting results. In the corn market, the hedge ratio decreases as the hedge horizon length increases. On the other hand, the wheat market exhibits the greatest increment in hedge ratios. The one-day ratio is 0.4609 whereas the 3-day ratio is 0.7097, representing a 54% increase.

6. CONCLUDING REMARKS

In this note, we have shown that, after reparemeterization, a temporal aggregation of cointegrated systems remains a cointegrated system. This result provides a convenient method to estimate n-day hedge ratios. The only concern pertains to the possible misspecification due to incorrect lag selections. Empirical results from ten futures contracts indicate lag selections have no effect on the estimated hedge ratios.

ACKNOWLEDGMENTS

The authors wish to acknowledge Joseph Sicilian, De-Min Wu, and an anonymous referee for helpful comments and suggestions. Of course, the authors are solely responsible for any remaining errors.

REFERENCES

Chou, W. L., Denis, K. K. F., & Lee, C. F. (1996). Hedging with the Nikkei Index Futures: The Conventional versus the Error Correction Model. *Quarterly Review of Economics and Finance*, *36*, 495–505.

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

Fama, E. F., & French, K. R. (1987). Commodity Futures Prices: Some Evidence on Forecasting Power, Premiums, and the Theory of Storage. *Journal of Business*, *60*, 55–74.

Ghosh, A. (1993). Hedging with Stock Index Futures: Estimation and Forecasting with Error Correction Model. *Journal of Futures Markets*, *13*, 743–752.

Lien, D. (1996). The Effect of the Cointegration Relationship on Futures Hedging: A Note. *Journal of Futures Markets*, *16*, 773–780.

Lien, D., Tse, Y. K., & Tsui, A. K. C. (2001). Evaluating Hedging Performance of the Constant-Correlation GARCH Model. *Applied Financial Economics*, forthcoming.

Myers, R. J., & Thompson, S. R. (1989). Generalized Optimal Hedge Ratio Estimation. *American Journal of Agricultural Economics*, *71*, 858–868.

APPENDIX A

By definition, $d(n) = [(1+\alpha)^n - 1]/n\alpha$. To compare $d(n)$ with $d(n+1)$, consider the following:

$$\frac{(1+\alpha)^{n+1} - 1}{n+1} - \frac{(1+\alpha)^n - 1}{n} = \frac{(1+\alpha)^n(n\alpha - 1) + 1}{n(n+1)}. \tag{A1}$$

Note that

$$\frac{d[(1+\alpha)^n(n\alpha - 1)]}{d\alpha} = n(n+1)\alpha(1+\alpha)^{n-1}, \tag{A2}$$

which is negative whenever $-1 < \alpha < 0$. Thus, $(1+\alpha)^n(n\alpha - 1) > -1$ (the minimum value at $\alpha = 0$). Consequently, $d(n+1) < d(n)$.

APPENDIX B

When $n = 2$, we have the following:

$$\Delta_2 f_t = \varepsilon_{ft} + (1+\beta)\varepsilon_{f,t-1} - \beta\varepsilon_{s,t-1} + \beta(1+\gamma)\sum_{j=2}^{\infty} \gamma^{j-2}\varepsilon_{t-j}. \tag{B1}$$

After algebraic manipulation we derive

$$Var(\Delta_2 f_t) = \left(\frac{2\beta^2}{1-\gamma}\right)\sigma_{ss} + \left(2 + 2\beta + \frac{2\beta^2}{1-\gamma}\right)\sigma_{ff} - 2\left(\beta + \frac{2\beta^2}{1-\gamma}\right)\sigma_{sf}. \tag{B2}$$

Similarly,

$$\Delta_2 s_t = \varepsilon_{st} + (1+\alpha)\varepsilon_{s,t-1} - \alpha\varepsilon_{f,t-1} - \alpha(1+\gamma)\sum_{j=2}^{\infty} \gamma^{j-2}\varepsilon_{t-j}, \tag{B3}$$

and

$$Var(\Delta_2 s_t) = \left(\frac{2\alpha^2}{1-\gamma}\right)\sigma_{ff} + \left(2 + 2\alpha + \frac{2\alpha^2}{1-\gamma}\right)\sigma_{ss} - 2\left(\alpha + \frac{2\alpha^2}{1-\gamma}\right)\sigma_{sf}. \tag{B4}$$

Upon applying Eqs (19), (B1) and (B3), the following can be derived:

$$Var(z_{t-2}) = (\sigma_{ss} + \sigma_{ff} - 2\sigma_{sf})/(1-\gamma^2); \tag{B5}$$

$$Cov(\Delta_2 f_t, z_{t-2}) = \beta(\sigma_{ss} + \sigma_{ff} - 2\sigma_{sf})/(1-\gamma^2); \tag{B6}$$

$$Cov(\Delta_2 s_t, z_{t-2}) = -\alpha(\sigma_{ss} + \sigma_{ff} - 2\sigma_{sf})/(1-\gamma^2); \tag{B7}$$

$$Cov(\Delta_2 f_t, \Delta_2 s_t) = \left(-\alpha - \frac{2\alpha\beta}{1-\gamma}\right)\sigma_{ff} + \left(-\beta - \frac{2\alpha\beta}{1-\gamma}\right)\sigma_{ss} + \left(2 + \alpha + \beta + \frac{4\alpha\beta}{1-\gamma}\right)\sigma_{sf}. \tag{B8}$$

MARKET TIMING, SELECTIVITY, AND MUTUAL FUND PERFORMANCE

Cheng-Few Lee and Li Li

ABSTRACT

This paper tests various CAPM-based market-timing and selectivity models, we find that about 12% of the funds have a statistically significant Alpha with about 4% of the funds having a significantly positive Alpha, and 8% of the funds having a significantly negative Alpha. About 15% of funds show significant timing ability with about 9% funds having a significantly positive timing coefficient and 6% of the funds having a significantly negative timing coefficient. The Asset Allocation funds demonstrate the most timing ability and the Aggressive Growth funds demonstrate the least timing ability.

1. INTRODUCTION

We test several CAPM-based and APT-based market-timing and selectivity models in order to estimate time-varying risk and expected returns. These models assume that a manager has private information about firm-specific risk and/or the aggregate stock market movements, and that this information will lead a manager to revise his portfolio allocation or use derivatives instruments to capitalize on his superior ability to forecast market movements. We examine managers' market-timing ability by: (1) comparing the fund return and the market return directly; and (2) comparing market movements and the fund's beta over time.

Testing various CAPM-based market-timing and selectivity models, we find that there are about 12% of the funds have a statistically significant Alpha with

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 41–84.

ISBN: 0-7623-0887-7

about 4% of the funds having a significantly positive Alpha and 8% of the funds having a significantly negative Alpha. About 15% funds show significant timing ability with about 9% funds having a significantly positive timing coefficient and 6% of the funds having a significantly negative timing coefficient. The Asset Allocation funds demonstrate the most timing ability and the Aggressive Growth funds demonstrate the least timing ability. This empirical evidence is consistent with the reality that the Asset Allocation funds focus on forecasting the aggregate factor and that the Aggressive Growth funds concentrate on identifying under- or over-valued stocks. We also find a negative correlation between the timing coefficient and Alpha; that is, managers that are good at picking stocks are not good at timing the market, and vice versa. Testing of the APT-based market timing model, we find a higher percentage of the funds with significantly positive timing since it is easier for managers to correctly time a market that has more risk factors. The APT-based market timing model provides a portfolio manager with better means to assess and control the risk and the expected return of a portfolio than is available through the CAPM-based market timing models. This is especially true for the Asset Allocation funds and Equity Income funds, since these funds are more responsive to economic cycles.

Recent years have witnessed a rapid rise in the importance of mutual funds and other professionally managed funds as investment vehicles. There has been a corresponding flurry of research on the investment performance of mutual fund managers and research on techniques for measuring that performance. As a part of a market efficiency theory test, performance evaluation concentrates on investigating whether managed funds can outperform the market. Abnormal returns of managed funds are interpreted as reflecting managers' superior ability to outperform the market. Fama (1972) indicated that there are two ways to achieve such abnormal returns; they are superior security selection ability and superior market timing ability. Selection ability is the ability of funds managers to identify the potential winning securities. Market timing ability is the ability of funds managers to time market cycles and react accordingly; that is, portfolio managers increase the relative volatility of their portfolios in anticipation of a bull market and reduce its volatility prior to a bear market.

There are many ways to investigate the ability of fund managers to select securities with abnormally high returns. One consequence of a manager actively engaging in security selection is that the construction of the fund deviates from that of the market index. This result implies that the fund will exhibit some degree of diversifiable risk. The degree of a fund's nonmarket risk is therefore a proxy of sorts for the degree of security selection exhibited by its managers. CAPM and many elaborate versions of CAPM are designed specifically to detect the size of abnormal returns. Many empirical studies have been done in using these models. To examine the market timing ability of fund managers, one can compare the fund return and the market return directly. If the fund's manager does not engage in market timing, then the portfolio beta should be sticky over time. Thus, the portfolio should earn a fairly constant fraction of the market return over time (presuming the fund is well diversified). Otherwise, if the manager exhibited market timing, the relationship between the portfolio returns and the market returns will be nonlinear. Another way to assess timing performance is to compare market movements and the fund's beta over time. Managers must change their fund betas to enhance performance through market timing. Changing fund betas can be accomplished via security trading such as altering the bond-stock mix or transaction in derivatives such as shorting stock index futures contracts.

Several market timing and selectivity models have been developed along these two lines. We test the Treynor-Mazuy model (1966), the Jensen model (1972), and the Bhattacharya-Pfleiderer model (1983). Treynor and Mazuy (1966) developed the first quadratic market-timing model to examine market timing ability of mutual fund managers by comparing the fund return and the market return directly. Jensen (1972) and Bhattacharya and Pfleiderer (1983) elaborated on the Treynor-Mazuy model by checking some properties of the residual term in the regression. These models assume that the manager has private information about future market movements, and that this information will lead the manager to revise his portfolio allocation. They measured the market-timing ability by the correlation between the market timer's forecasted market returns and the actual market returns. Bhattacharya and Pfleiderer (1983) improved on the Jensen model by minimizing the variance of the forecast error; that is; they added the signal-to-noise ratio for adjustment. However, there are some problems with these models. (A) There are substantial estimation biases from using the error term to identify a manager's timing skill. It is impossible obtain proper coefficients from analyzing the residual term without a preliminary estimate for the unknown parameters. In addition, there are some inseparable error terms in the estimation. These reduce the power of the models. (B) There is a question about the sources and quality of managers' private information. (C) Managers' ability to utilize the private information is hard to measure and is associated with their risk preferences. We also test the Treynor-Mazuy Total Performance Measure that aggregates the effects of market timing and selectivity proposed by Grinblatt and Titman (1994). The Treynor-Mazuy total performance measure is specifically designed to pick up beta variations that are related linearly to the return of the benchmark portfolio.

A market timing advantage may instead show up through a different mechanism – using derivatives. A fund manager can avoid transaction costs of portfolio reallocation by using index futures and options. These derivatives instruments offer a manager a cheaper way of capitalizing on his superior ability to forecast market movements. We test the Henriksson-Merton model (1981) and the Connor-Korajczyk model (1991). Henriksson and Merton (1981) developed a market-timing model using a free put option on the market portfolio with its exercise price equal to the risk free rate. In their model, market-timing ability is measured by the correlation between the portfolio beta and the true market return, a manager adjusts his portfolio to a higher beta when the forecast is for an up market, and a lower beta when the market forecast is pessimistic. Connor and Korajczyk (1991) extended the Henriksson-Merton model in two ways: (1) They replaced the free put option with a costly replicated put option through dynamic trading; and (2) They replaced the CAPM-based Henriksson and Merton model with the APT-based Henriksson and Merton model. The results of testing the Henriksson-Merton model reveal the significant negative correlations between selectivity and timing parameters; that is, managers that are good at picking stocks are not good at timing the market, and vice versa. By incorporating the dynamic trading model and the asset beta nonlinearity model, the Connor-Korajczyk model reduced the significance of the negative correlation between selectivity and timing parameters. However, this negative correlation implies intuitively that some fund managers can predict market risk better than they can identify potential winning stocks; others can do the opposite; which is likely true in reality. By testing the model for funds with different investment objectives, we provide evidence that is consistent with the intuitive explanation for the negative correlations. The aggressive growth funds have more significant negative correlations than the asset allocation funds have, since a manager of an aggressive growth fund concentrates on identifying winning stocks such as high-tech stocks and bio-tech stocks, and a manager of a asset allocation fund would make effort to keep up with the market.

Testing various CAPM-based market-timing and selectivity models, we find that there are about 12% of the funds have a statistically significant Alpha with about 4% of the funds having a significantly positive Alpha and 8% of the funds having a significantly negative Alpha. About 15% funds show significant timing ability with about 9% funds having a significantly positive timing coefficient and 6% of the funds having a significantly negative timing coefficient. The Asset Allocation funds have the highest percentage of about 13% of the funds having a significantly positive timing coefficient; and have the lowest percentage of about 3% of the funds with a significantly negative timing coefficient. The Aggressive Growth funds have the lowest percentage of about 2% of the funds with a significantly positive timing coefficient; and have the highest percentage of about 10% of the funds with significantly negative timing coefficient. This empirical evidence is consistent with the reality that the Asset Allocation funds focus on forecasting the aggregate factor and that the Aggressive Growth funds concentrate on identifying under- or over-valued stocks.

We test two multi-factor market-timing models: the conditional Treynor-Mazuy model and the APT-based Henriksson-Merton model. The above CAPM-based market-timing and selectivity models interpret the problem of time-varying beta and expected market returns as fund managers' superior information on firm-specific level and/or the aggregate stock market level. However, the recently developed conditional CAPM, a modification of the CAPM proposed by Ferson and Schadt (1996), addresses the problem of the time-varying beta and expected market returns and the correlation between them by incorporating lagged information variables into the traditional CAPM. We find that the products of the future benchmark return and the lagged information variables capture the covariance between the conditional beta and the conditional expected market return given that market prices fully reflect readily available public information. This covariance is a major source of bias in the traditional, unconditional models. By incorporating the conditional model into the Treynor-Mazuy model, we obtain less biased estimates of timing and selectivity parameters.

We test the APT model using the macroeconomic-variable approach suggested by Chen, Roll and Ross (1986). The five macroeconomic shocks used in this study are: The monthly growth rate in the U.S. industrial production, the unanticipated change in default risk premium, the unanticipated change in the slope of the term structure of interest rates, the unanticipated inflation rate, and the unexpected change in the unemployment rate. We first estimate the statistical factors using maximum likelihood factor analysis, and we then rotate five macroeconomic shocks on the statistical factors. The primary advantages of this approach are: (1) By rotating the original macroeconomic shocks on the statistic factors, it enhances the interpretation of the macroeconomic factors; and (2) Rotated macroeconomic factors introduce additional information and link asset-pricing behavior to macroeconomic events. Motivated by an equilibrium version of the APT, we obtain the market residual factor as the sixth factor. This factor may be thought of as a proxy for otherwise omitted or incompletely specified factors – the part of the market index excess return that is not explained by the other five rotated factors. Mutual funds' abnormal returns, and their risk sensitivities to the five rotated macroeconomic shocks and the market residual factor, are then estimated from the time series analysis. Comparing to results from the CAPM-based market timing models, testing of the APT-based Henriksson-Merton model, we find a higher percentage of the funds with significantly positive timing since it is easier for managers to correctly time the market that has more risk factors. The APT-based market timing model provides the portfolio manager with better means to assess and control the risk and the expected return of a portfolio than is available through standard Markowitz-type portfolio analysis and the CAPM. This is especially true for the Asset Allocation funds and Equity Income funds since the APT-based Henriksson-Merton model captures deviations from the CAPM-based Henriksson-Merton model due to missing risk factors since the CAPM assumes that stock prices move together only because of common movement with the stock market. Managers' of these funds can develop separate forecasts of each macroeconomic variable and they can use their forecasts to design portfolios that provide the greatest return-to-risk performance based on the sensitivity of each stock to each macroeconomic variable, and the historical risk premiums associated with each macroeconomic variable.

However, the existing market timing models are not suitable for measuring a manager's timing ability when he uses index options and futures instead of reallocation. For example, if a manager believes that the market will go up, he will simply take a long position in index futures instead of revising his allocation, and the return realized on buying index futures will reflect his timing ability. The market-timing model using the mutual fund data alone will not measure timing correctly because of the use of derivatives in mutual fund management. It would be an interesting task to measure the market timing ability by checking their holdings of derivatives. There is problem with using index options: Mutual funds have very limited holdings of put and call options, most of them hold none at all. Furthermore, it is not easy to separate ability into two such dichotomous categories as selectivity and timing ability. To measure market timing ability, we have to develop more sophisticated models instead of using the CAPM-type models.

This paper is organized as follows. Section 2 decomposes market timing and selection ability theoretically. Section 3 and 4 presents various market-timing models. Section 5 describes the data. Section 6 tests the CAPM-based market timing models, the conditional market timing model, and the APT-based market timing model. Section 7 concludes the paper.

2. IDENTIFYING MARKET TIMING AND SELECTION ABILITY

Fama (1972) suggested that portfolio managers' forecasting skills could be partitioned into two distinct components: (1) Security Analysis or Micro-forecasting, that is, forecasts of price movements of selected individual stocks; and (2) Market timing or Macroforecasting, that is, forecasts of price movement of the general stock market as a whole. Microforecasting involves identifying under- or over-valued securities. Market timing refers to forecasts of movement of the market.

Within the specification of the CAPM, a microforecaster attempts to identify securities whose expected returns lie significantly off the securities market line. Specifically, the microforecaster would only forecast the nonsystematic or security-specific component of a security or a portfolio return. Following Jensen (1972), the excess return on a portfolio can be written as

$$\tilde{r}_{pt} = \beta_p \tilde{r}_{mt} + \tilde{\varepsilon}_{pt} \tag{1}$$

Where $\tilde{r}_{pt}$ and $\tilde{r}_{mt}$ are the excess rate of return on the pth portfolio and the market portfolio over the risk-free rate, respectively. β_p Measures the sensitivity of the portfolio return to the market return. $\tilde{\varepsilon}_{pt}$ is the random component of the portfolio return that is not explained by the systematic component of the portfolio return and has $E(\tilde{\varepsilon}_{pt}) = 0$.

If the portfolio manager is a superior forecaster, he will tend to select securities so that his portfolio will realize $\tilde{\varepsilon}_{pt} > 0$. In other word, his portfolio will earn more than the normal risk premium for its level of risk. Allowing for the possible existence of non-zero constant, we can rewrite Eq. (1) as follows:

$$\tilde{r}_{pt} = \alpha_p + \beta_p \tilde{r}_{mt} + \tilde{\upsilon}_{pt} \tag{2}$$

The new error term has $E(\tilde{\upsilon}_{pt}) = 0$. Thus if a portfolio manager has an ability to forecast security price, then $\alpha_p > 0$. A passive buy-and-hold strategy would be expected to have $\alpha_p = 0$. $\alpha_p < 0$ if a manger is doing worse than using a buy-and-hold strategy.

A market timer will attempt to capitalize on any expectation he may have regarding the behavior of the market in the next period. If a manager believes that he can forecast the market movement, he will adjust his portfolio risk level in anticipation of market movements. For example, if the manager correctly perceives that there is a high probability that the market return will rise in the next period, he will increase the relative volatility of his portfolios in anticipation of the bull market so that he will earn abnormal returns to the market portfolio. One the other hand, if the market return is expected to fall in the next period, he will reduce volatility of his portfolio prior to the bear market so that he can reduce the losses on the portfolio. Practically, a portfolio manager can adjust his portfolio risk by changing the asset mix such as the stocks versus money market securities in a common stock mutual fund, or switching from more risky to less risky securities (or vice versa) in an attempt to outguess the movement of the market such as adjusting the proportion of aggressive vs. defensive stocks. Therefore, we can allow for the existence of timing ability in Eq. (2) by permitting the sensitivity coefficient β_p to be stochastic. Market-timing ability will be present where β_p and $\tilde{r}_{mt}$ are positive correlated.

3. MARKET TIMING THROUGH PORTFOLIO REALLOCATION

Treynor-Mazuy Model (1966)

Treynor and Mazuy (1966) developed the first quadratic market-timing model to examine market timing ability of mutual fund managers by comparing the fund return and the market return directly. They argue that if the fund's manager does not engage in market timing, then the portfolio beta should be sticky over time. Thus, the portfolio should earn a fairly constant fraction of the market return over time (presuming the fund is well diversified). Otherwise, if the manager exhibited market timing, the relationship between the portfolio returns and the market returns will be nonlinear. Visually, the market returns and fund returns should plot on a straight line or if there is random noise should plot as a scatter of points about a straight line. On the other hand, if the manager exhibited superior market timing, then the points would plot above the normal line. There will be curvature to the scatter of points, indicating market timing ability.

Treynor and Mazuy (1966) examined the timing ability of mutual fund managers by testing for such curvature. Specifically, they fit the following quadratic to the data:

$$t_{pt} = \alpha_p + \beta_p(r_{mt}) + \gamma_p(r_{mt})^2 + \varepsilon_{pt} \tag{3}$$

Where r_{pt} is the excess return on the pth portfolio and r_{mt} is the excess return on the market portfolio. β_p measures the sensitivity of the portfolio return to the market return. If there is no market timing ability, there should be a linear relationship between r_{pt} and r_{mt}, so that coefficient γ_p should be statistically insignificant from zero. If there is superior market timing, γ_p will be positive. The addition of squared term in Eq. (3) will improve the empirical fit between

r_{pt} and r_{mt}. Hence, the coefficient γ_p can be regarded as an indicator of the manager's market timing ability.

Treynor and Mazuy report that only one of 37 mutual funds during 1953–1962 studied exhibited a significantly positive γ_p coefficient, which are less than the number expected to occur by random chance. Grinblatt and Titman (1988) and Cumby and Glen (1990) found large amount negative γ_p coefficients using the Treynor-Mazuy model.

Jensen Model (1972)

Jensen (1972) has developed a similar model to detect fund managers' selectivity and timing skills where the basis for evaluation is a comparison of the ex post return of a mutual fund with the market returns. It is assumed that the forecasted market return by a market timer and the actual return on the market have a joint normal distribution. Jensen showed that, under this assumption, a market timer's forecasting ability can be measured by the correlation between the market timer's forecasted return and the realized return on the market.

Jensen defines the market factor $\tilde{\pi}_t = \tilde{r}_{mt} - E(\tilde{r}_m)$ as the deviation of the market return from the expected market return at time t. Where $E(\tilde{r}_m)$ is the expected return of $\tilde{r}_{mt}$ unconditional upon any special information and on the market portfolio perceived by the market participants. Letting $E(\tilde{\pi}_t | \phi_t)$ represent the expected value of $\tilde{\pi}_t$ conditional upon ϕ_t, which is the information set available to the manager at the beginning of the period t. The fund manager receives information that is not available to others in the market about the market portfolio return in the next period. The manager's optimal forecast of the market factor is:

$$\tilde{\pi}_t^* = E(\tilde{\pi}_t | \phi_t) = \tilde{\pi}_t + \tilde{\upsilon}_t \qquad (4)$$

where, $\tilde{\upsilon}_t$ is a normally distributed and independent of $\tilde{\pi}_t$.

Assume that the portfolio is managed in the interest of a group of investors who have constant absolute risk aversion of an unknown degree. Given the objective of the manager and the assumption that the conditional distribution of $\tilde{\pi}_t^*$ is optimal, Jensen showed that:

$$\tilde{\beta}_{pt} = \beta_{pT} + \theta \tilde{\pi}_t^* \qquad (5)$$

where, β_{pT} is the "target beta" of the fund and θ measures the manager's response to information ϕ_t, that is, a manager's ability to reallocate based on his information. $\theta \tilde{\pi}_t^*$ reflects a fund manager's timing decision and indicates the risk-level deviation from the fund's target risk level depending on the optimal forecast of the market factor $\tilde{\pi}_t^*$ and the manager's response to the information ϕ_t.[1]

Then, Jensen rewrite the Eq. (2) as follows:

$$\tilde{r}_{pt} = \alpha_p + (\beta_{pT} + \theta(\tilde{\pi}_t + \tilde{\upsilon}_t)) \cdot [E(\tilde{r}_m) + \tilde{\pi}_t] + \tilde{\upsilon}_{pt} \tag{6}$$

Consider the regression of $\tilde{r}_{pt}$ on a constant, $\tilde{\pi}_t$, and $\tilde{\pi}_t^2$:

$$\tilde{r}_{pt} = \eta_0 + \eta_1 \tilde{\pi}_t + \eta_2 \tilde{\pi}_t^2 + \tilde{\upsilon}_{pt} \tag{7}$$

Jensen (1972)[2] claims that

$$\begin{aligned} p \lim \hat{\eta}_0 &= \alpha_p + \beta_{pT} E(\tilde{r}_m) + \theta(\rho^2 - 1)\sigma_\pi^2, \\ p \lim \hat{\eta}_1 &= \rho^2 \theta E(\tilde{r}_m) + \beta_{pT}, \\ p \lim \hat{\eta}_2 &= \theta. \end{aligned} \tag{8}$$

where, ρ is the correlation between the predicted $\tilde{\pi}_t$ and the realized of $\tilde{\pi}_t$, i.e. a measure of a manager's timing ability. A high ρ means that a manager has better timing ability. This system has more unknowns than number of equations. Jensen concludes that, under above structure, separate contributions of micro- and macro-forecasting can not be identified unless, for each period, the market timing forecast and $E(\tilde{r}_m)$ are known.

Bhattacharya-Pfleiderer Model (1983)

Bhattacharya and Pfleiderer (1983) extended Jensen's (1972) work and showed that one can use a simple regression technique to obtain accurate measures of timing and selection ability. They assume that the manager observes a signal $\tilde{\pi}_t + \tilde{\varepsilon}_t$ at the beginning of period t, where $\tilde{\varepsilon}_t$ is normally distributed and independent of $\tilde{\pi}_t$. They assume that the manager adjusts his forecasts to minimize the variance of the forecast error, whereas Jensen assumes that the manager uses the unadjusted forecast of the market factor in the timing decision.

According to Bhattacharya and Pfleiderer, The manager's optimal forecast of the market factor is

$$\tilde{\pi}_t^* = \psi(\tilde{\pi}_t + \tilde{\varepsilon}_t) \tag{9}$$

where, $\psi = \sigma_\pi^2/(\sigma_\pi^2 + \sigma_\varepsilon^2)$ is the adjustment factor, in other word, the signal-to-noise ratio, which is the coefficient of determination between the manager's forecast and the excess return on the market.

Using $\beta_{pT} = \theta E(\tilde{r}_m)$ and $\tilde{\pi}_t = \tilde{r}_{mt} - E(\tilde{r}_m)$ from Jensen (1972), they obtain the following equation:

$$\tilde{r}_{pt} = \alpha_p + \theta\{E(\tilde{r}_m) + \psi[\tilde{r}_{mt} - E(\tilde{r}_m) + \tilde{\varepsilon}_t]\} \cdot (\tilde{r}_{mt}) + \tilde{\upsilon}_{pt} \tag{10}$$

Rewriting Eq. (10),

$$\tilde{r}_{pt} = \alpha_p + \theta E(\tilde{r}_m)(1 - \psi)(\tilde{r}_{mt}) + \psi\theta(\tilde{r}_{mt})^2 + \theta\psi\tilde{\varepsilon}_t(\tilde{r}_{mt}) + \tilde{\upsilon}_{pt} \tag{11}$$

Rewriting Eq. (11),

$$\tilde{r}_{pt} = \eta_0 + \eta_1(\tilde{r}_{mt}) + \eta_2(\tilde{r}_{mt})^2 + \tilde{\omega}_t \tag{12}$$

$$p \lim \hat{\eta}_0 = \alpha_p$$

$$p \lim \hat{\eta}_1 = \theta E(\tilde{r}_m)(1 - \psi)$$

$$p \lim \hat{\eta}_2 = \theta\psi$$

Equation (12) is similar to the Treynor-Mazuy model in terms of observation variables. Using above regression, one can detect the existence of stock selection ability as revealed by α_p. The disturbance term is $\tilde{\omega}_t = \theta\psi\tilde{\varepsilon}_t(\tilde{r}_{mt}) + \tilde{\upsilon}_{pt}$, where $\theta\psi\tilde{\varepsilon}_t(\tilde{r}_{mt})$ contains the information needed to quantify the manager's timing ability.

Bhattacharya and Pfleiderer (1983) showed that α_p is an accurate measure of security selection ability. A manager having no security specific information will have $\alpha_p = 0$. An informed manager who acts appropriately and optimally will record a positive α_p. In this model, managers who have security specific information may also have information that permits them to time the market. This model is a refinement of the Treynor-Mazuy model, which focuses on the coefficient of the squared excess market return to detect timing skill. It is the first model to-date that analyzes the error term to identify a manager's forecasting skill. Such a refinement should make the model more powerful than previous ones.

Conditional Treynor-Mazuy Model

The above CAPM-based market-timing and selectivity models interpret the problem of time-varying beta and expected market returns as fund managers' superior information on firm-specific level and/or the aggregate stock market level. However, the recently developed conditional CAPM, a modification of the CAPM proposed by Ferson and Schadt (1996), addresses the problem of the time-varying beta and expected market returns and the correlation between them by incorporating lagged information variables into the traditional CAPM. They assume that market prices fully reflect readily available public information[3], which is measured by a vector of predetermined variables Z_t. In other words, they assume that markets are informational efficient in a version of the "semi-strong form" efficiency of Fama (1970). They assume a linear functional form for the conditional beta, given Z_t, of a managed portfolio:

$$\beta_{im}(Z_t) = b_{oi} + B_i' z_t \tag{13}$$

where, $z_t = Z_t - E(Z)$ is a vector of the deviations of Z_t from the unconditional means, and B_i is a vector with dimension equal to the dimension Z_t. The coefficient b_{oi} is an "average beta", i.e. and the unconditional mean of the conditional beta $E(\beta_{im}(Z_t))$. The elements of B_i measure the response of the conditional beta to the information variations Z_t. The time-variation of a managed portfolio beta can change as a function of Z_t because: (1) the betas of the underlying assets available to managers may change over time; (2) the weights of a passive strategy such as buy-and-hold will vary as relative values change; and (3) a manager can actively manipulate the portfolio weights by developing from a buy-and-hold strategy.

Incorporating the equation (13) into the traditional CAPM, they obtain the conditional model of a regression of a managed portfolio excess return on the market factor and the product of the market factor with the lagged information:

$$R_{it} - R_{ft} = \alpha_i + \delta_{1i}(R_{mt} - R_{ft}) + \delta_{2i}'(z_{t-1}(R_{mt} - R_{ft})) + \varepsilon_{it} \tag{14}$$

where,

$$E(\varepsilon_{it} | Z_{t-1}) = E(\varepsilon_{it}(R_{mt} - R_{ft}) | Z_{t-1}) = 0.$$

The vector of predetermined information variables z_{t-1} consists: (1) the lagged level of the one-month Treasury bill yield (TBILL), measured by monthly return of 3-month Treasury Bill; (2) the lagged dividend yield of the CRSP value-weighted NYSE and AMEX stock index (DY), measured by the differential of VW index include dividends and VW exclude dividends; (3) a lagged measure of the slope of the U.S. Government term structure (TERM), measured as the difference between the Long Term Treasury Bill and month return of 3-month Treasury Bill; (4) a lagged quality spread in the corporate bond market (DEFAULT), measured as the difference between the Lehman Brother High Yield corporate bond index and High Quality bond index, and (5). a dummy variable for the month of January.

We incorporate the conditional model into the Treynor-Mazuy model and write the conditional Treynor-Mazuy model as follows:

$$R_{it} - R_{ft} = \alpha_i + \delta_{1i}(R_{mt} - R_{ft}) + \gamma_p(R_{mt} - R_{ft})^2 + \delta_p'(z_{t-1}(R_{mt} - R_{ft})) + \varepsilon_{it} \tag{15}$$

The coefficient γ_p measures the sensitivity of the manager's beta to the private market-timing signal. The term $\delta_p' z_{t-1}(R_{mt} - R_{ft})$ controls for the public information effects, which would bias the coefficients in the original Treynor-Mazuy model. The new term in the conditional Treynor-Mazuy model captures the part of the squared term in the Treynor-Mazuy model that is attributes the public information variables. In the conditional model, the correlation of mutual fund betas with the future market return, which can be attributed to the public information, is not considered to reflect market timing ability.

4. MARKET TIMING BASED ON THE OPTION PRICING MODEL

Henriksson-Merton Model (1981)

Merton (1981) and Henriksson and Merton (1981) have developed a market-timing model based on the Option Pricing Model. In their model, a manager attempts to forecast when the market portfolio return will exceed the risk-free rate. They assumed that a manager receives a binary signal (high or low) each period that is correlated with the true outcome of the market return realization. The manager chooses one of two values for the portfolio beta (high or low) in response to the signal received. That is the manager adjusts the portfolio to a higher target beta when the forecast is for an up market, a lower target beta is used when the market forecast is pessimistic.

Henriksson and Merton (1981) prove that the following regression gives consistent estimates of timing ability:

$$r_{pt} = \alpha_p + \beta_p(r_{mt}) + \gamma_p put(r_{mt}) + \varepsilon_{pt} \tag{16}$$

where, $put(r_{mt}) = \max((-r_{mt}), 0)$ as the payoff to an option on the market portfolio with exercise price equal to the risk free asset. r_{pt} is the excess return on the pth portfolio and r_{mt} is the excess return on the market portfolio. They show that $\gamma_p > 0$ if and only if the investor possesses superior market timing ability and $\alpha_p > 0$ if the manager has selection ability.

Henriksson (1984) estimates the above equation for a sample of 116 mutual funds. He finds little evidence of superior timing ability. In fact, he notes that more of the funds have negative estimated gamma than positive ones. Furthermore, he provides some evidence that the estimated gamma is negatively correlated with the estimated alpha across funds. He conjectures that this may be due to errors-in-variables. Chang and Lewellen (1984) and Grinblatt and Titman (1988) estimate the Henriksson-Merton model and reach conclusion similar to those of Henriksson.

One weakness of the Henriksson-Merton model is that information is measured but there is no test of whether information is being used correctly. They assume that managers have a coarse information structure in which dichotomous signals are only predictive of the sign of the excess return of the market relative to the risk-free rate. In their model, the probability of receiving an “up” or a “down” signal in no way depends upon how far the market will be "up" or "down". The managers in this model are less sophisticated than those in Jensen (1972), where they do forecast how much better the superior investment will perform.

Connor-Korajczyk Model (1991)

Both Henriksson (1984) and Chang and Lewellen (1984) have found the large amount of significantly negative timing coefficients. These negative coefficients imply irrational behavior of fund managers. In order to produce a negative timing coefficient, a manager must possess superior information and employ it irrationally, i.e. raises market risk when he receives a signal that the market will fall and lowers market risk when he receives a signal that the market will rise. Both Chang and Lewellen (1984) and Henriksson (1984) provided the evidence that the estimate of gamma is negatively correlated with the estimate of alpha across funds. However, they conjectured that it might be due to error-in-variables. Jagannathan and Korajczyk (1986) argues that such results could arise from artificial market timing due to the differential leverage of the firms in the indices and those invested in by mutual funds. They theoretically and empirically demonstrate how to create a portfolio that would exhibit negative timing performance when no true timing exists.

Connor and Korajczyk (1991)[4] extended the original Henriksson-Merton model by incorporating a dynamic trading model and a asset beta nonlinearities model[5] within the Option Pricing Model framework. They view the funds as buying (or shoring) a replicated costly put option, whereas the Henriksson-Merton model views the funds as owing a free put option. The Connor and Korajczyk model is given as follows:

$$r_{pt} = \alpha_p^* + \beta_p(r_{mt}) + \gamma_p Nput(r_{mt}) + \varepsilon_{pt} \tag{17}$$

where, $Nput(r_{mt}) = put(r_{mt}) - (1 + R_f)P_0$ is the payoff to a market put minus the Treasury bill return necessary to pay the market price of the put. $P_0 = \dfrac{\alpha_p}{-(1 + R_f)\gamma_p}$ is present value of the put option on the market index, it is estimated from regression of the original Henriksson-Merton model. If the manager has no timing or selection ability, the Henriksson-Merton model predicts $\alpha_p^* = 0$, whereas the Connor and Korajczyk model concerning the prevalence of negative estimated gamma and the negative correlation between estimated gamma and alpha.predict $\alpha_p^* = -(1 + R_f)\gamma_p P_0$. This shows the negative relation between alpha and gamma and combined selectivity effect and timing effect. In the absence of selection ability, the cash flow from the costly put option makes the Connor and Korajczyk model distinct from the Henriksson-Merton model, since the cost of options will be reflected in the mean returns through the intercept.

This new instrument *Nput* is simply portfolio insurance. A fund that holds $Nput_0$ and the market portfolio is guaranteed the risk-free return minus the future value (at the risk free rate) of the put price. Under the null hypothesis of no selection and timing ability, even when the manager is trading option like securities, α_p^* should be zero. If the manager has true selection or timing ability, then α_p^* should be greater than zero. The extended Henriksson-Merton model allows for superior information based true selectivity and timing skill for managers as well as non-information based variation in the market risk of the portfolio caused by dynamic trading and asset beta nonlinearities.

The Connor-Korajczyk version of the Henriksson-Merton model has the advantage over the original Henriksson-Merton model in that it gives a consistent measure of performance when the mutual fund buy or sell costly options. However, it has several limitations. First, unlike the Henriksson-Merton model that allows timing skill and selectivity skill to be separately identified, the new model measures only the sum of timing an selectivity skills. Second, it only allows for one-month-ahead European options, whereas the return of the fund could correspond to an infinite variety of optionlike patterns.[6]

APT-based Henriksson-Merton Model

Connor and Korajczyk (1991) argued that the market index probably did not adequately reflect the significant γ coefficient, which might be better captured by the multiple factors of the APT. They extend the Henriksson-Merton model to a multi-factor model and they would expect the coefficient to disappear in the new model.

$$r_{pt} = \alpha_p + \beta_1 r_{f1t} + \cdots + \beta_k r_{fkt} + \gamma_1 put(r_{f1t}) + \cdots + \gamma_k put(r_{fkt}) + \varepsilon_{pt} \qquad (18)$$

where, $\gamma_i > 0$ if and only if the manager has superior information about factor i. The $P_0 = \dfrac{\alpha_p}{-(1+R_f)\gamma_p}$ term is included to measure selectivity and $\alpha_p > 0$ implies superior selectivity. They found that coefficient γ_i is little affected in moving from the original Henriksson-Merton model to their version of the extended model using the factors they have chosen for the APT.

5. DATA

The Fund Data

The data sample used in this study is a subset of the Morningstar Principia Plus database. The Morningstar database contains monthly data for 7901 open-ended mutual funds. The fund returns data reflect the reinvestment of dividends and capital gains and are net of expenses, excepting front-end load charges and exit fees.[7] We consider the monthly return of 1937 equity funds ranging from February 1978 to January 1997. The sample excludes sector funds, international funds and balanced funds, since these funds have different risk components and require additional factors to span the space covered by their investments. Any fund that did not survive until the end of 1996 is not included in this database. Therefore, the survivorship bias issue can not be addressed using this database. We sort 1937 equity funds into five groups according to their investment objectives, 118 funds are aggressive growth funds, 177 funds are asset allocation funds, 149 funds are equity income funds, 971 funds are growth funds, and 522 funds are growth income funds. During the time period being examined, there are, on average, 592 funds in existence each month, with a minimum of 212 funds and maximum of 1937 funds.

Choice of the Market Portfolio

We compare the CRSP value-weighted index (VW), the CRSP equally weighted index (EW), S&P500, and Wilshire 5000. Past research generate biased performance measures that relate to firm size (Banz, 1981; Reinganum, 1981), dividend yield (Litzenberger & Ramaswamy, 1979, 1982), and beta (Black, Jensen & Scholes, 1972). Grinblatt and Titman (1988) confirmed that the firm size, dividend yield, and beta-related biases with the EW index. In particular, funds that invest in large firms tend to exhibit negative performance with the EW index. Comparing the VW, S&P500 and Wilshire5000 by their mean and variance, Wilshire5000 is the most efficient market portfolio. We use the Wilshire5000 as the market portfolio in this research. In testing of the APT, we also use the Wilshire5000 as the market index. Furthermore, the risk-free rate is not real risk-free, since the standard deviation is positive.

Table 1. Mean-Standard Deviation of Equally-Weighted Portfolios and Market Indexes.

	Standard Deviation (monthly %)	Mean (monthly %)
CRSP VW	4.2458	1.2440
CRSP EW	4.9905	1.4221
SP500	4.1674	1.2675
Wilshire5000	4.2844	1.2845
T-bill	0.2286	0.5837
Aggressive Growth	5.5828	1.4607
Asset Allocation	2.6133	1.0223
Equity Income	3.1164	1.1430
Growth	4.3596	1.3235
Growth Income	3.6791	1.1737
All Funds	4.0189	1.2555

The Predetermined Information Variables

The conditional CAPM includes a vector of lagged information variables, z_{t-1}. We use the same variables used by Ferson and Schadt (1996). They are:

(1) The lagged level of the one-month Treasury bill yield (TBILL), measured by monthly return of the 3-month Treasury Bill,
(2) The lagged dividend yield of the CRSP value-weighted NYSE and AMEX stock index (DY), measured by the differential of the VW index include dividends and the VW exclude dividends,
(3) A lagged measure of the slope of the U.S. Government term structure (TERM), measured as the difference between monthly return of the Long Term Treasury Bill and month return of the 3-month Treasury Bill,
(4) A lagged quality spread in the corporate bond market (DEFAULT), measured as the difference between the Lehman Brother High Yield corporate bond index and the Lehman Brother High Quality bond index,
(5) A dummy variable for the month of January.

Macroeconomic Shocks Used in the APT Model

The macroeconomic variables model used in this study is based on the following five factors as used by Chen, Roll and Ross (1986). They are:

(1) Δ_{UIP} = Monthly growth rate in the U.S. industrial production. We obtain the monthly Industrial Production data from the Bloomberg Economics

Indexes. It can also be obtained from the Table 2.10 in the Federal Reserve Bulletin published by the Board of Governors of the Federal Reserve System.

$$\Delta_{UIP} = \ln(IP_t) - \ln(IP_{t-1})$$

(2) Δ_{URP} = Unanticipated changes in default risk premium, measured as the difference between monthly return of the Lehman Brothers High Yield Corporate Bond Index and monthly return of the Lehman Brothers High Quality Corporate Bond Index. Both are obtained from the Morningstar database.

(3) Δ_{UTS} = Unanticipated changes in the slope of the term structure of interest rates, measured as the difference between monthly return of the Lehman Brothers long-term Treasury bill rate and monthly return of the 3-month Treasury bill rate. Both are obtained from the Morningstar database.

(4) Δ_{UCPI} = Unanticipated inflation rate. The inflation rate is obtained from the Morningstar database.

$$\Delta_{UCPI} = R_{CPI,t} - E(R_{CPI,t}|_{t-1})$$

Where $R_{CPI,t} = CPI_t/CPI_{t-1}$, is monthly percentage increase in the consumer price index.

(5) Δ_{UEMP} = Unexpected changes of the unemployment rate. The seasonal-adjusted civilian unemployment rate $R_{unemp,t}$ is obtained from the Bureau of Labor Statistics.

$$\Delta_{UEMP,t} = \Delta UNE_t - E(\Delta UNE_t|_{t-1})$$

Where, $\Delta UNE_t = (R_{unemp,t} - R_{unemp,t-1})$ is the change in the civilian unemployment rate.

6. EMPIRICAL RESULTS

Treynor-Mazuy Model

We test the following Treynor-Mazuy model.

$$R_{pt} - R_{ft} = \alpha_p + \beta_p(R_{mt} - R_{ft}) + \gamma_p(R_{mt} - R_{ft})^2 + \varepsilon_{pt} \tag{19}$$

where,

R_{pt} = the rate of return of the pth fund in month t

$Nput$ = the rate of return of the risk-free asset in month t

R_{mt} = the rate of return of the market portfolio in month t

Table 2. Average Measures from the Treynor-Mazuy Model.

Fund Objectives	No. of Funds	Alpha	Beta	Gamma	Adj. R^2	DW
Aggressive Growth	118	0.0465	1.2811	–4.4608	0.6252	1.7533
		(0.037)	(12.41)	(–0.8458)		
Asset Allocation	177	–0.0858	0.5388	0.3965	0.6884	1.6931
		(–0.4059)	(11.34)	(0.6641)		
Equity Income	149	0.1220	0.7175	–0.2765	0.7547	1.7950
		(0.4639)	(15.56)	(–0.3202)		
Growth	971	–0.0729	1.0155	–0.7709	0.7479	1.8477
		(–0.370)	(15.44)	(–0.168)		
Growth Income	522	–0.0264	0.8852	0.5046	0.8454	1.8520
		(–0.3032)	(22.2)	(0.4522)		
Total	1937	–0.0393	0.9301	–0.5073	0.7618	1.8249
		(–0.2663)	(16.71)	(–0.1148)		

β_p measures the sensitivity of the portfolio return to the market return.

γ_p can be regarded as an indicator of the manager's market timing ability.

Table 2 reports the average performance and the average timing coefficient estimated from the Treynor-Mazuy model. The Aggressive Growth funds have positive average Alpha of 56 basis points annually instead of a negative average Alpha. For the rest of funds, sign of average Alpha are same as those from the CAPM and its functional form models. The entire sample underperforms the market by 47.2 basis points annually with the worst performance for the Asset Allocation funds (–1.03% per year) and best performance for the Equity Income funds (1.46% per year). The beta increase monotonically as we move along the risk spectrum from the Asset Allocation funds (0.54), the Equity Income funds (0.72), the Growth Income funds (0.89), the Growth funds (1.02), to the Aggressive Growth funds (1.28).

Table 3 gives the number of funds with a significantly positive and a significantly negative Alpha at 5% and 10% (in parentheses) significance levels estimated from the Treynor-Mazuy model for the entire sample and each of the five groups. There are 3.8% of the funds (74 funds) with a significantly positive Alpha at the 5% significance level and 6% (117 funds) funds with a significantly positive Alpha at the 10% significance level. However, there are more funds with a significantly negative Alpha. There are 7.7% of the funds (149 funds) with a significantly negative Alpha at the 5% significance level and 12.7% (246 funds) funds with a significantly negative Alpha at the 10%

Table 3. Funds with Statistically Significant Alpha and Gamma Estimated from the Treynor-Mazuy Model.

Fund Objectives	No. of Funds	Statistically Significant Alpha: Positive Alpha	Statistically Significant Alpha: Negative Alpha	Statistically Significant Alpha: Total Percentage	Statistically Significant Gamma: Positive Gamma	Statistically Significant Gamma: Negative Gamma	Statistically Significant Gamma: Total Percentage
Aggressive Growth	118	0.8% (4.3%)	1.7% (2.5%)	2.5% (6.8%)	1.7% (3.4%)	9.3% (25.4%)	11% (28.8%)
Asset Allocation	177	4.5% (6.8%)	9.1% (15.2%)	13.6% (22%)	12.4% (14.7%)	3.4% (5.1%)	15.8% (19.8%)
Equity Income	149	6.7% (9.4%)	0.7% (2.7%)	7.4% (12.1%)	4% (4%)	6.1% (12.8%)	10.1% (16.8%)
Growth	971	3.6% (5.2%)	8.6% (14.9%)	12.2% (20.1%)	9.1% (12.6%)	6.1% (9.2%)	15.2% (21.8)
Growth Income	522	3.8% (6.9%)	9% (12.8%)	12.8% (19.7%)	8.8% (18.2%)	7.1% (9.6%)	15.9% (27.8%)
Total	1937	3.8% (5.2%)	7.7% (12.7%)	11.5% (18.7%)	8.5% (13.1%)	6.3% (10.2%)	14.8% (23.3%)

significance level. Appendix A.2. lists funds with both a significantly positive Alpha and a significantly negative Alpha at 5% and 10% significance levels.

For market timing parameter, there are 8.5% funds (164 funds) with a significantly positive timing coefficient at the 5% significance level, and 13.1% funds (253 funds) with a significantly positive market timing coefficient at the 10% significance level. However, there are 6.3% funds (122 funds) with a significantly negative timing coefficient at the 5% significance level, and there are 10.2% funds (197 funds) with a significantly negative timing coefficient at the 10% significance level. Among the five groups, the Asset Allocation funds have the highest percentage (12.4% at the 5% significance level and 14.7% at the 10% significance level) funds with a significantly positive timing coefficient and the lowest percentage (3.4% at the 5% significance level and 5.1% at the 10% significance level) funds with a significantly negative timing coefficient. The Aggressive Growth funds have the lowest percentage funds (1.7% at the 5% significance level and 3.4% at the 10% significance level) with a significantly positive timing coefficient and the highest percentage funds (9.3% at the 5% significance level and 25.4% at the 10% significance level) with a significantly negative timing coefficient. This empirical evidence is consistent with what in reality that the Asset Allocation funds focus on forecasting the aggregate factor and the Aggressive Growth funds concentrate on identifying under- or over value stocks.

Treynor-Mazuy Total Performance Measure

Grinblatt and Titman (1994) proposed the Treynor and Mazuy (1966) total performance measure that aggregate the effects of market timing and selectivity. The Treynor-Mazuy total performance measure is defined as:

$$TM = \alpha_p + \gamma_p Var(R_{mt} - R_{ft}) \tag{20}$$

The standard t-statistic can be used to test whether it is significantly different from zero, when conditioned on the excess returns of the portfolios in the benchmark. The test statistic is $TM/s(TM)$, which has a t-distribution with T-K-1 degrees of freedom if there are T returns and K benchmark portfolios. Where $s(TM)$ is the standard error of the Treynor-Mazuy total performance measure.

To computer this standard error, we need the following:

(1) Computed $s(e)$, the standard error of the regression from the excess return regression used to compute the Jensen Measure for the portfolio being evaluated.
(2) Computed the variance-covariance matrix of the three coefficients in the quadratic regression, conditional on the benchmark excess return. This is $V = s^2(e)(X'X)^{-1}$. Where X is the $\mathrm{T} \times 3$ matrix of regressors in the Treynor-Mazuy quadratic regression.
(3) Computed $s(TM)$ as the square root of $q'Vq$, where the 1×3 row vectors $q' = (10Var(R_{mt} - R_{ft}))$, and $Var(R_{mt} - R_{ft})$ is the variance of the benchmark portfolio's excess return.

Table 4 reports the results estimated from the Treynor-Mazuy total performance measure. The results are similar to that estimated from the Treynor-Mazuy model. While the absolute value of average total performance are slightly reduced for most of the funds, since for most of funds, selectivity Alpha is negatively correlated to market timing Gamma and the cross-term in the Eq. (21) actually reduced the total performance for most of funds. The number of funds with a statistically significant total performance is similar to the number of funds with a statistically significant Alpha in Table 13. There are 3.9% of the funds (76 funds) with a significantly positive Alpha at the 5% significance level and 6.4% of the funds (124 funds) with a significantly positive Alpha at the 10% significance level. However, there are more funds with a significantly negative Alpha. There are 7.2% of the funds (140 funds) with a significantly

Table 4. Summary Results Estimated from the Treynor-Mazuy Total Performance Measure.

Fund Objectives	No. of Funds	Alpha	Gamma	Total Performance	Statistically Significant Total Performance: Positive	Negative	Total Percentage
Aggressive	118	0.0465	–4.4608	0.0422	0.8%	0.8%	1.7%
Growth		(0.037)	(–0.8458)	(0.0290)	(5.1%)	(2.5%)	(7.6%)
Asset	177	–0.0858	0.3965	–0.0854	5.65%	6.2%	11.9%
Allocation		(–0.4059)	(0.6641)	(–0.4137)	(8.5%)	(14.1%)	(22.6%)
Equity	149	0.1220	–0.2765	0.1218	7.4%	0	7.4%
Income		(0.4639)	(–0.3202)	(0.4711)	(12.1%)	(3.3%)	(15.4%)
Growth	971	–0.0729	–0.7709	–0.0735	3.4%	9%	12.4%
		(–0.370)	(–0.168)	(–0.3106)	(4.9%)	(11.5%)	(20.5%)
Growth	522	–0.0264	0.5046	–0.0260	4%	7.9%	11.9%
Income		(–0.3032)	(0.4522)	(–0.3011)	(7.1%)	(11.5%)	(18.6%)
Total	1937	–0.0393	–0.5073	–0.0398	3.9%	7.2%	11.1%
		(–0.2663)	(–0.1148)	(–0.2366)	(6.4%)	(12.6%)	(19.0%)

positive Alpha at the 5% significance level and 12.6% of the funds (244 funds) with significantly negative Alpha at the 10% significance level. Comparing to the results from the Treynor-Mazuy model, the number of funds with a statistically significant Alpha is slightly more.

Conditional Treynor-Mazuy Model

We test the following conditional Treynor-Mazuy model:

$$\begin{aligned} R_{pt} - R_{mt} = {} & \alpha_p + \beta_p(R_{mt} - R_{ft}) + \gamma_p(R_{mt} - R_{ft})^2 + \delta_{tbill}TBILL(R_{mt} - R_{ft}) \\ & + \delta_{dy}DY(R_{mt} - R_{ft}) + \delta_{term}TERM(R_{mt} - R_{ft}) \\ & + \delta_{default}DEFAULT(R_{mt} - R_{ft}) + dummy + \varepsilon_{pt} \end{aligned} \tag{21}$$

where,

γ_p measures the sensitivity of the manager's beta to the private market-timing signal.

TBILL is the lagged level of the one-month Treasury bill yield (TBILL), measured by monthly return of the 3-month Treasury Bill, *DY* is the lagged dividend yield of the CRSP value-weighted NYSE and AMEX stock index (DY), measured by the differential of the VW index include dividends and the VW exclude dividends,

TERM is the lagged measure of the slope of the U.S. Government term structure (TERM), measured as the difference between monthly return of the Long Term Treasury Bill and month return of the 3-month Treasury Bill,

DEFAULT is the lagged quality spread in the corporate bond market (DEFAULT), measured as the difference between the Lehman Brother High Yield corporate bond index and the Lehman Brother High Quality bond index,

dummy is the dummy variable for the month of January.

Table 5 reports average monthly performance estimated from the conditional Treynor-Mazuy model. Comparing to the results estimated from the unconditional Treynor-Mazuy model in Table 2, the average Alpha estimated from the conditional Treynor-Mazuy model have less absolute value for the entire sample and each of the five groups. The entire sample underperforms the market portfolio by 34.8 basis points for the conditional model, underperforms by 10.68 basis points for the unconditional model. Furthermore, the Asset Allocation and the Equity Income groups have higher percentage of significantly positive market timing abilities, which is consistent with common

Table 5. Average Monthly Performance Estimated from the Conditional Treynor-Mazuy Model.

Fund Objectives	No. of Funds	Alpha	Beta	Gamma	Adj. R^2
Aggressive Growth	118	0.0236	1.3661	–0.0227	0.6801
		(0.0315)	(3.8814)	(–0.6386)	
Asset Allocation	177	–0.0557	0.5326	0.0012	0.7502
		(–0.4064)	(2.8762)	(0.3403)	
Equity Income	149	0.0711	0.7897	–0.0041	0.8072
		(0.4467)	(4.4673)	(–0.4870)	
Growth	971	–0.065	1.0404	–0.0023	0.7964
		(–0.4391)	(4.5041)	(–0.0466)	
Growth Income	522	0.015	0.9282	0.0045	0.8786
		(0.4099)	(6.3406)	(0.3823)	
Total	1937	–0.02901	0.9684	–0.0015	0.8094
		(–0.3297)	(4.8465)	(–0.0802)	

understanding. The average betas estimated for the conditional Treynor-Mazuy model are slightly larger than that for the unconditional Treynor-Mazuy model. Like the unconditional Treynor-Mazuy model, the beta increase monotonically as we move along the risk spectrum from the Asset Allocation funds (0.53), the Equity Income funds (0.79), the Growth Income funds (0.93), the Growth funds (1.04), to the Aggressive Growth funds (1.37).

Table 6 reports funds with statistically significant Alpha at 5% and 10% (in parentheses) levels estimated from the conditional Treynor-Mazuy model for the entire sample and each of five groups. Comparing to the results estimated from the unconditional Treynor-Mazuy model in Table 3, the number of funds with a significantly positive Alpha and market timing coefficients are more for the conditional model. It is the evidence that the conditional model capture the time-varying beta and expected market returns.

Bhattacharya-Pfleiderer Model

Without Correction for Heteroskedasticity

We test the following regression using OLS procedure:

$$R_{pt} - R_{ft} = \eta_0 + \eta_1(R_{mt} - R_{ft}) + \eta_2(R_{mt} - R_{ft})^2 + \omega_t \tag{22}$$

Table 6. Funds with Statistically Significant Alpha and Gamma Estimated from the Conditional Treynor-Mazuy Model.

		Statistically Significant Alpha			Statistically Significant Gamma		
Fund Objectives	No. of Funds	Positive Alpha	Negative Alpha	Total Percentage	Positive Gamma	Negative Gamma	Total Percentage
Aggressive Growth	118	1.7% (4.2%)	0.8% (2.6%)	2.5% (6.8%)	1.7% (4.2%)	8.5% (24.6%)	10.2% (28.8%)
Asset Allocation	177	8.5% (13%)	5.1% (6.2%)	13.6% (19.2%)	14.1% (16.9%)	2.3% (4%)	16.4% (20.9%)
Equity Income	149	7.4% (10%)	0 (2.8%)	7.4% (12.8%)	4% (4%)	4.7% (11.4%)	8.7% (15.4%)
Growth	971	4.8% (6.5%)	8% (13.5%)	12.8% (20%)	10% (14.3%)	5.4% (8.7%)	15.4% (23%)
Growth Income	522	3.6% (7.1%)	7.3% (12.6%)	10.9% (19.7%)	9.6% (19.3%)	5.9% (9.6%)	15.5% (28.9%)
Total	1937	4.9% (7.4%)	6.5% (11.1%)	11.4% (18.5%)	9.3% (14.5%)	5.4% (9.7%)	14.7% (24.2%)

We get $\hat{\eta}_0 = \alpha_p$, $\hat{\eta}_1 = \theta E(R_m - R_f)(1-\psi)$, and $\hat{\eta}_2 = \theta\psi$ from the regression, and the residual term ω_{pt}. The α_p reflects managers' security selection ability. Then, we use the residual term ω_{pt} to test the second regression:

$$\omega_t^2 = \theta^2\psi^2\sigma_\varepsilon^2(R_{mt} - R_{ft})^2 + \zeta_t \tag{23}$$

where,

$$\zeta_t = \theta^2\psi^2(R_{mt} - R_{ft})^2(\varepsilon_t^2 - \sigma_\varepsilon^2) + \upsilon_{pt}^2 + 2\theta\psi(R_{mt} - R_{ft})\varepsilon_t\upsilon_{pt} \tag{24}$$

We get the coefficient of the squared term in Eq. (23). Coupled with coefficient of the squared term in Eq. (22), we obtain σ_ε^2. *We estimate* σ_π^2 using Merton (1980) technique of estimating the variance of $\tilde{\pi}_t$ from the available time series of realized return on the market under the assumption that $\tilde{\pi}_t$ follows a stationary Wiener process. The advantage of this estimator is that the variance can be estimated without knowing or even having an estimate of, the mean. It saves one degree of freedom. A reasonable estimate of σ_π^2 was derived as follows:

$$\hat{\sigma}_\pi^2 = \left(\sum_{n=1}^{N} (\ln(1 + \tilde{R}_{mt}))^2\right) \Big/ N \tag{25}$$

Because the estimator for σ_π^2 is not taken around the sample mean, it will be biased. However, for large sample, the difference between the second central and noncentral moments is trivial. In this case, σ_π^2 as estimated in the Eq. (25) is 0.00200489, while the sample variance of $\tilde{\pi}_t$ is 0.002. The timing ability is obtained by $\rho = \sqrt{\sigma_\pi^2/(\sigma_\pi^2 + \sigma_\varepsilon^2)}$.

The results of the average monthly performance and average market timing parameter for entire funds sample and each of the five funds groups are as same as the results from the Treynor-Mazuy model, since the Eq. (22) is simply the Treynor-Mazuy model. The significance of the timing ability is also as same as that in the Treynor-Mazuy model. The average ρ for the entire sample is 0.3164, the five groups have very similar the average ρ. However, there exists cross-sectional heteroskedasticity[8] for residual terms. In the following, we report the more efficient estimators with correction for the heteroskedasticity problem.

With Correction for Heteroskedasticity

In order to obtain efficient estimators, we test the Eq. (22) using the full information maximum likelihood estimation for correction of heteroskedasticity. Lee and Rahman (1990) used the two-step FGLS (Feasible Generalized Least Square) procedure for correction of heteroskedasticity. The FGLS procedure assumes that a residual covariance matrix is a block diagonal. The full information Maximum likelihood estimation is superior to the FGLS procedure in the way that the MLE does not assume a block diagonal of a residual covariance matrix. It provides the efficient estimators when a residual covariance matrix is unknown and a non block diagonal. If a residual covariance matrix is a block diagonal, the FGLS and MLE will give the same estimators.

To estimate the unknown parameters in Eq. (22), we use the full information maximum likelihood method to perform the mixed effect modeling involves two random variables $\tilde{\varepsilon}_t$ and $\tilde{\upsilon}_{pt}$. Note, the error terms $\theta\psi\tilde{\varepsilon}_t E(R_m - R_f)$ and $\tilde{\upsilon}_{pt}$ may be inseparable without prior information regarding to their distribution properties. If no prior information is available, the variance of these two terms, i.e. $\theta^2\psi^2\sigma^2_{\tilde{\varepsilon}_t}E(R_m - R_f)^2$ and $\sigma^2_{\tilde{\upsilon}_{pt}}$ cannot be estimated separately. Hence, the estimates of the coefficients in Eq. (22) will be the same estimates that are obtained by treating these two error terms as one total error term.

Let MLE_{linear} and MLE_{square} represent the maximum likelihood estimate of the coefficients of the linear and square term, respectively, then the maximum likelihood estimate of ψ can be obtained by

$$\hat{\psi} = \frac{E(R_m - R_f)MLE_{linear}}{E(R_m - R_f)MLE_{linera} + MLE_{square}} \tag{26}$$

Rewrite Eq. (26) as

$$\hat{\psi} = \frac{E(R_m - R_f)MLE_{linear}}{E(R_m - R_f)MLE_{linera} + MLE_{square}} = \frac{E(R_m - R_f)}{E(R_m - R_f) + \dfrac{MLE_{linera}}{MLE_{square}}} \tag{27}$$

Table 7 shows average performance estimated from the Bhattacharya and Pfleiderer model with correction for heteroskedasticity. The average monthly Alpha is similar to those estimated without correction of heteroskedasticity, but with smaller absolute value. The entire sample underperforms the market by 11.5 basis points annually with the worst performance for the Asset Allocation funds (–75 basis points per year) and best performance for the Equity Income funds (1.36% per year). The beta increase monotonically as we move along the risk spectrum from the Asset Allocation funds (0.54), the Equity Income funds (0.72), the Growth Income funds (0.89), the Growth funds (1.02), to the Aggressive Growth funds (1.28). The average Rho is increased for the entire sample and each of the five groups comparing to the estimators without correction for heteroelasticity.

Table 8 reports the number of a significantly positive and a significantly negative Alpha at 5% and 10% (in parentheses) significance levels estimated

Table 7. Average Measures Estimated from the Bhattacharya-Pfleiderer Model with Correction for Heteroskedasticity.

Fund Objectives	No. of Funds	Alpha	Beta	Gamma	Rho	Adj. R^2	DW
Aggressive Growth	118	0.0339	1.2835	–4.5029	0.4662	0.634	1.740
		(0.1334)	(12.419)	(–0.967)			
Asset Allocation	177	–0.0622	0.5409	0.4906	0.4843	0.691	1.704
		(–0.439)	(11.461)	(0.8454)			
Equity Income	149	0.1137	0.7170	–0.1917	0.3970	0.749	1.809
		(0.4263)	(15.582)	(–0.159)			
Growth	971	–0.0334	1.0164	–0.7407	0.4977	0.768	1.848
		(–0.431)	(15.501)	(–0.021)			
Growth Income	522	–0.012	0.8857	0.5414	0.4826	0.871	1.859
		(–0.332)	(22.243)	(0.5323)			
Total	1937	–0.006	0.931	–0.4696	0.4834	0.776	1.828
		(–0.321)	(16.767)	(–0.159)			

Table 8. Funds with Statistically Significant Alpha and Gamma Estimated from the Bhattacharya-Pfleiderer Model with Correction for Heteroskedasticity.

		Statistically Significant Alpha			Statistically Significant Gamma		
Fund Objectives	No. of Funds	Positive Alpha	Negative Alpha	Total Percentage	Positive Gamma	Negative Gamma	Total Percentage
Aggressive	118	1.7%	0.8%	2.5%	1.7%	8.5%	10.2%
Growth		(5.1%)	(2.5%)	(7.6%)	(3.4%)	(25.4%)	(28.8%)
Asset	177	4.5%	8.5%	13%	14.1%	2.8%	16.9%
Allocation		(7.3%)	(14.1%)	(21.4%)	(16.9%)	(5.1%)	(22%)
Equity Income	149	7.4%	0.7%	8.1%	5.4%	5.4%	10.8%
		(10.7%)	(2.7%)	(13.4%)	(6.7%)	(10%)	(16.7%)
Growth	971	4%	7.8%	11.8%	9.2%	5.9%	15.1%
		(6.2%)	(14.2%)	(20.4%)	(12.8%)	(9%)	(21.8%)
Growth Income	522	4.2%	8.2%	12.4%	10%	6.9%	16.9%
		(7.9%)	(12.5%)	(20.4%)	(19%)	(9.2%)	(28.2%)
Total	1937	4.2%	7%	11.2%	9.1%	6%	15.1%
		(7%)	(12.1%)	(19.1%)	(13.8%)	(9.8%)	(23.6%)

from the Bhattacharya and Pfleiderer model with correction for heteroskedasticity for the entire sample and each of the five groups. There are 4.2% of the funds (82 funds) with a significantly positive Alpha at the 5% significance level and 7% of the funds (136 funds) with a significantly positive Alpha at the 10% significance level. However, there are more funds with a significantly negative Alpha. There are 7% of the funds (136 funds) with a significantly negative Alpha at the 5% significance level and 12.1% of the funds (235 funds) with a significantly negative Alpha at the 10% significance level. Comparing to the results without correction for heteroskedasticity, the number of funds with a significantly positive are increased slightly and the number of funds with a significantly negative are decreased slightly. Overall, the total percentage of funds with a statistically significant Alpha remained the same.

For market timing parameter, there are 9.1% of the funds (176 funds) with a significantly positive timing coefficient at the 5% significance level, and 13.8% of the funds (267 funds) with a significantly positive timing coefficient at the 10% significance level. However, there are 6% of the funds (116unds) with a significantly negative timing coefficient at the 5% significance level, and there are 9.8% of the funds (189 funds) with a significantly negative timing coefficient at the 10% significance level. Among the five groups, the Asset Allocation funds have the highest percentage (13% at the 5% significance level and 15.3% at the 10% significance level) funds with a significantly positive timing coefficient and the lowest percentage (2.3% at the 5% significance level and 5.1% at the 10% significance level) funds with a significantly negative timing coefficient. The Aggressive Growth funds have the lowest percentage funds (1.7% at the 5% significance level and 3.4% at the 10% significance level) with a significantly positive timing coefficient and the highest percentage funds (9.3% at the 5% significance level and 25.4% at the 10% significance level) with a significantly negative timing coefficient. This empirical evidence is consistent with the Treynor-Mazuy model.

The Aggressive Growth funds and the Growth funds have the largest variation of the estimated selectivity parameter, and the Asset Allocation funds have the most uniform selectivity. A manager earns abnormal returns when he has superior selectivity and market timing ability.

Henriksson-Merton Model

We test the following Henriksson-Merton model:

$$R_{pt} - R_{ft} = \alpha_p + \beta_p(R_{mt} - R_{ft}) + \gamma_p put(-(R_{mt} - R_{ft}), 0) + \varepsilon_{pt} \tag{28}$$

where, $put(-(R_{mt} - R_{ft}), 0) = \max(-(R_{mt} - R_{ft}), 0)$ as the payoff to an option on the market portfolio with exercise price equal to the risk free asset.

Table 9. Average Measures Estimated from the Henriksson-Merton Model.

Fund Objectives	No. of Funds	Alpha	Beta	Gamma	Adj. R^2	DW
Aggressive Growth	118	0.1968	1.0553	–0.4471	0.6159	1.7386
		(0.125)	(6.403)	(–0.608)		
Asset Allocation	177	–0.0996	0.5603	0.0360	0.6846	1.6874
		(–0.406)	(6.675)	(0.436)		
Equity Income	149	0.2141	0.6749	–0.0992	0.7548	1.7882
		(0.659)	(8.290)	(–0.56)		
Growth	971	–0.0757	0.9874	–0.0534	0.7449	1.8427
		(–0.357)	(8.803)	(–0.120)		
Growth Income	522	–0.0288	0.9060	0.0342	0.8443	1.8485
		(–0.300)	(12.85)	(0.337)		
Total	1937	–0.0263	0.9066	–0.0491	0.7591	1.8196
		(–0.238)	(9.513)	(–0.111)		

Table 9 reports the average monthly performance and market timing parameter estimated from the Henriksson-Merton model. Comparing to results from the Treynor-Mazuy model, the estimated average Alpha are similar to those estimated from the Treynor-Mazuy model except for the Aggressive Growth funds and the Equity Income funds. The average Alpha is larger than to those estimated from the Treynor-Mazuy model for the Aggressive Growth funds and the Equity Income funds. The entire sample underperforms the market by 31.6 basis points per year, the Equity Income funds still have the highest average Alpha of 2.57% per year. The beta increase monotonically as we move along the risk spectrum from the Asset Allocation funds (0.56), the Equity Income funds (0.67), the Growth Income funds (0.91), the Growth funds (0.99), to the Aggressive Growth funds (1.06). However, beta coefficients are smaller than those from the Treynor-Mazuy model.

Table 10 reports the number of a significantly positive and a significantly negative Alpha at 5% and 10% (in parentheses) significance levels estimated from the Henriksson-Merton model for the entire sample and each of the five groups. There are 4.2% of the funds (82 funds) with a significantly positive Alpha at the 5% significance level and 6.9% of the funds (134 funds) with a significantly positive Alpha at the 10% significance level. However, there are more funds with a significantly negative Alpha. There are 8.3% of the funds (161 funds) with a significantly positive Alpha at the 5% significance level and 13% of the funds (252) with a significantly negative Alpha at the 10% significance level. Comparing to the results from the Treynor-Mazuy model in

Table 10. No. of Funds with a Statistically Significant Alpha and Gamma Estimated from the Henriksson-Merton Model.

Fund Objectives	No. of Funds	Statistically Significant Alpha: Positive	Negative	Total Percentage	Statistically Significant Gamma: Positive	Negative	Total Percentage
Aggressive Growth	118	3.4% (6.8%)	0.8% (2.5%)	4.2% (9.3%)	0.8% (2.5%)	6.8% (11.9%)	7.6% (14.4%)
Asset Allocation	177	2.8% (4%)	10.8% (18.6%)	13.6% (22.6%)	10.7% (11.9%)	2.3% (4.5%)	13% (16.4%)
Equity Income	149	6.7% (12.8%)	1.3% (1.3%)	8% (14.1%)	1.3% (2%)	8.1% (10.1%)	9.4% (14.1%)
Growth	971	3.8% (6.4%)	9.7% (15.4%)	13.5% (21.8%)	8.4% (11.6%)	4.5% (7%)	12.9% (18.6%)
Growth Income	522	5% (7.3%)	8.6% (12.2%)	13.6% (19.5%)	8.2% (13.4%)	5.6% (8.1%)	13.8% (21.5%)
Total	1937	4.2% (6.9%)	8.3% (13%)	12.5% (19.9%)	7.5% (10.8%)	5% (7.8%)	12.5% (18.6%)

Table 3, the number of funds with a statistically significant (both positive and negative) Alpha from the Henriksson-Merton model in Table 10 is more. Therefore, the total percentage of funds with a statistically significant Alpha is increased.

For market timing ability, there are 7.5% funds (146 funds) with a significantly positive timing coefficient at the 5% significance level, and 10.8% funds (210 funds) with a significantly positive timing coefficient at the 10% significance level. However, there are 5% funds (97 funds) with a significantly negative timing coefficient at the 5% significance level, and there are 7.7% funds (150 funds) with a significantly negative timing coefficient at the 10% significance level. Among the five groups, the Asset Allocation funds have the highest percentage (10.7% at the 5% significance level and 11.9% at the 10% significance level) funds with a significantly positive timing coefficient and the lowest percentage (2.3% at the 5% significance level and 4.5% at the 10% significance level) funds with a significantly negative timing coefficient. The Aggressive Growth funds have the lowest percentage funds (0.8% at the 5% significance level and 2.5% at the 10% significance level) with a significantly positive timing coefficient and the second highest percentage funds (6.8% at the 5% significance level and 11.9% at the 10% significance level) with a significantly negative timing coefficient. The Equity Income funds have the highest percentage funds (8.1% at the 5% significance level and 12.1% at the 10% significance level) with a significantly negative timing coefficient.

In general, negative correlation between selectivity and market timing parameters for the entire funds sample and each of the five funds groups.

Connor-Korajczyk Model

We test the following Connor-Korajczyk model:

$$R_{pt} - R_{ft} = \alpha_p^* + \beta_p(R_{mt} - R_{ft}) + \gamma_p Nput + \varepsilon_{pt} \tag{29}$$

The Connor-Korajczyk model measures the sum of timing and selectivity skills by redefine a negative correlation between timing and selectivity as $\alpha_p^* = -(1 + R_f)\gamma_p P_0$, which is predicted by the dynamic trading model and the asset beta nonlinearities model. While, the Henriksson-Merton model predicts $\alpha_p^* = 0$. The replicated put option is priced as $Nput = put - (1 + R_f)P_0$, which is the payoff to a market put minus the Treasury bill return necessary to pay the market price of the put.

Table 11 reports the average monthly performance and the average timing coefficient estimated from the Connor-Korajczyk model. The average Alpha is quite different from those from the Henriksson-Merton model. The Connor-Korajczyk model mixes the selectivity and market timing ability with the replicated put option on the market portfolio. The Alpha reflects the sum performance from selectivity and market timing ability. The Aggressive

Table 11. Average Measures Estimated from the Connor-Korajczyk Model.

Fund Objectives	No. of Funds	Alpha	Beta	Gamma (nput)	Adj. R^2	DW
Aggressive Growth	118	–0.3512	1.3563	0.8698	0.6397	1.8119
		(–0.527)	(12.869)	(1.5080)		
Asset Allocation	177	–0.0442	0.5274	–0.0756	0.6826	1.7326
		(–0.164)	(10.75)	(–0.448)		
Equity Income	149	0.0896	0.7067	–0.0999	0.7522	1.8590
		(0.3367)	(15.301)	(–0.514)		
Growth	971	–0.1412	1.0312	0.1579	0.7475	1.8796
		(–0.438)	(15.352)	(0.2397)		
Growth Income	522	0.0095	0.8759	–0.1406	0.8436	1.8987
		(0.0829)	(21.967)	(–0.837)		
Total	1937	–0.0868	0.9382	0.0797	0.7613	1.8656
		(–0.263)	(16.559)	(0.0941)		

Growth funds have negative average Alpha (–4.2% annually). While, the Equity Income funds still have the highest average Alpha (1.08% per year). On average, the entire sample underperforms the market by 1% per year. The beta increase monotonically as we move along the risk spectrum from the Asset Allocation funds (0.53), the Equity Income funds (0.71), the Growth Income funds (0.88), the Growth funds (1.03), to the Aggressive Growth funds (1.36).

Table 12 reports the number of a significantly positive and a significantly negative Alpha at 5% and 10% (in parentheses) significance levels estimated from the Connor-Korajczyk model for the entire sample and each of the five groups. There are 3.3% of the funds (63 funds) with a significantly positive Alpha at the 5% significance level and 4.9% of the funds (95 funds) with a significantly positive Alpha at the 10% significance level. However, there are more funds with a significantly negative Alpha. There are 6.6% of the funds (127 funds) with a significantly positive Alpha at the 5% significance level and 11.8% of the funds (228 funds) with a significantly negative Alpha at the 10% significance level.

For market timing ability, there are 4.2% funds (81 funds) with a significantly positive timing coefficient at the 5% significance level, and 8.2% funds (158 funds) with a significantly positive timing coefficient at the 10% significance level. However, there are 4.2% funds (81 funds) with a

Table 12. Funds with Statistically Significant Alpha and Gamma Estimated from the Connor-Korajczyk Model.

Fund Objectives	No. of Funds	Statistically Significant Alpha Positive	Negative	Total Percentage	Statistically Significant Gamma Positive	Negative	Total Percentage
Aggressive Growth	118	0 (0)	3.4% (6.8%)	3.4% (6.8%)	26.3% (41.5%)	0 (0)	26.3% (41.5%)
Asset Allocation	177	6.8% (9.6%)	4.5% (7.9%)	11.3% (17.5%)	0 (1.7%)	3.4% (5.6%)	3.4% (7.3%)
Equity Income	149	4% (6%)	1.3% (2%)	5.3% (8%)	0 (0)	4% (5.4%)	4% (5.4%)
Growth	971	3.2% (4.8%)	8.5% (16.2%)	11.7% (21%)	5.1% (10.7%)	2.8% (4.6%)	7.9% (15.3%)
Growth Income	522	2.7% (4.2%)	5.7% (8.8%)	8.4% (13%)	0 (0.4%)	8% (14.7%)	8% (15.1%)
Total	1937	3.3% (4.9%)	6.6% (11.8%)	9.9% (16.7%)	4.2% (8.2%)	4.2% (7.2%)	8.4% (15.4%)

significantly negative timing coefficient at the 5% significance level, and there are 7.2% funds (140 funds) with a significantly negative timing coefficient at the 10% significance level. Among the five groups, the Aggressive Growth funds have the highest percentage funds (26.3% at the 5% significance level and 41.5% at the 10% significance level) with a significantly positive timing coefficient and the lowest percentage funds (zero) with a significantly negative timing coefficient. The Asset Allocation funds and the Equity Income funds have the lowest percentage funds with a significantly positive timing coefficient and the modest percentage funds with a significantly negative timing coefficient. The Growth Income funds have the highest percentage funds (8% at the 5% significance level and 14.8% at the 10% significance level) with a significantly negative timing coefficient. The results are quite different from other market timing models. One explanation is that this model mixes the selectivity and market timing ability with the replicated put option on the market portfolio.

APT-Based Henriksson-Merton Model

We test the APT model using the macroeconomic-variable approach suggested by Chen, Roll and Ross (1986). The five macroeconomic shocks used in this study are: The monthly growth rate in the U.S. industrial production, the unanticipated change in default risk premium, the unanticipated change in the slope of the term structure of interest rates, the unanticipated inflation rate, and the unexpected change in the unemployment rate. Chen, Roll and Ross (1986) suggest that these macroeconomic variables are highly correlated with the statistical factors that come out of factor analysis. Therefore, these variables can then be correlated with stock returns to come up with a model of expected returns, with firm-specific betas calculated relative to each variable. Maximum likelihood factor analysis is used to extract five statistical factors from the 40 constructed portfolios using the sample of stocks that consists of all stocks traded on the New York Security Exchange (NYSE) and American Stock Exchange (AMEX) from the monthly CRSP tape with no missing observations from February 1977 to December 1996 (total 239 months). To investigate the 'size effect' observed by Banz (1981) and to control the errors-in-variables problem that arises from the cross-sectional regressions, the entire sample is grouped into 40 portfolios (18 portfolios of 47 stocks and 22 portfolios of 46 stocks) that are constructed according to their year-end market capitalization value. Therefore, portfolio 1 represents the portfolio of the smallest 47 firms for each of twenty years; portfolio 2 represents the portfolio of the second smallest 47 firms for each of twenty years; and so forth (see Li Li & C. F. Lee, 1998).

Estimation of the Inflation Shock and the Unemployment Shock

To estimate the inflation shock, we estimate an autoregressive moving average model for $R_{CPI,t} = CPI_t / CPI_{t-1}$ and regress $R_{CPI,t}$ over $R_{CPI,t-1}$ and $R_{CPI,t-2}$. Then we use the differential between the actual $R_{CPI,t}$ and the predicted $R_{CPI,t}$ from this model as the Δ_{UCPI}. To estimate the unemployment shock, we estimate an autoregressive moving average model for $\Delta UNE_t = (R_{unemp,t} - R_{unemp,t-1})$ and regress the ΔUNE_t over ΔUNE_{t-1}, ΔUNE_{t-2}, ΔUNE_{t-3} and ΔUNE_{t-4}. The fitted time series models for $R_{CPI,t}$ and ΔUNE_t are reported in Table 13.

Rotation of the Statistical Factors

We adopt a rotation technique that rotates the macroeconomic shocks over the statistical factors. It is a linear combination of the statistical factors rotated so

Table 13. Fitted time series models $R_{CPOI,t}$ and ΔUNE_t
Monthly data from February 1977 to December 1996. 239 observations

	Intercept	$R_{CPI,t-1}$	$R_{CPI,t-2}$			Adj. R^2
$R_{CPI,t}$	0.1133 (4.1254)	0.5518 (8.5926)	0.1690 (2.6451)			0.4572

	Intercept	ΔUNE_{t-1}	ΔUNE_{t-2}	ΔUNE_{t-3}	ΔUNE_{t-4}	Adj. R^2
ΔUNE_t	–0.0038 (–0.3329)	–0.0672 (–1.055)	0.1712 (2.7436)	0.2171 (3.4814)	0.1865 (2.9272)	0.1066

Table 14. Correlation Matrix of Five Macroeconomic Shocks and Excess Return of Market Indexes.

	Δ_{UCPI}	Δ_{URP}	Δ_{UTS}	Δ_{UEMP}	Δ_{UIP}	Wsh-TBill	VW-TBill	EW-TBill
Δ_{UCPI}	1							
Δ_{URP}	0.0449	1						
Δ_{UTS}	0.2123	–0.1031	1					
Δ_{UEMP}	–0.0723	0.0296	–0.3130	1				
Δ_{UIP}	–0.1162	–0.0351	–0.2608	0.4173	1			
Wsh-TBill	0.1277	0.4199	0.3518	–0.1336	–0.1049	1		
VW-TBill	0.1270	0.4195	0.3608	–0.1375	–0.1108	0.9994	1	
EW-TBill	0.0221	0.6090	0.1729	–0.0682	–0.0859	0.8578	0.8592	1

as to mimic the movements in a macroeconomic shock. Each macroeconomic shock is regressed against five statistical factors. Then, we use the predicted series from these regressions as the new factors. Table 15 reports the rotation regression of the five macroeconomic shocks on the five statistical factors. It also reports that the excess returns of three market indexes regress on the five statistical factors. It shows that the R-squares from the VW and Wilshire5000 are very high (more than 96% of the variation can be explained) and they are almost the same. The R-square from the unemployment rate is very low (0.0061).

Motivated by an equilibrium version of the APT, we obtain the market residual factor as the sixth factor. This factor may be thought of as a proxy for otherwise omitted or incompletely specified factors – the part of the market index excess return that is not explained by the other five rotated factors. To obtain this variable, we first run a time series regression of the Wilshire5000 index on the five rotated factors and use the residual from this regression as the sixth factor. If the model is a reasonable return generating process, we would expect the first five variables to be related to the market in a sensible manner, and we would expect returns on individual mutual funds to be related to the six variables in a sensible manner.

Table 15. Regressions of the Macroeconomic Shocks and Excess Returns of Market Indexes on the Statistical Factors.

	Intercept	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Adj. R^2
Δ_{UCPI}	–5.87E-09	0.0009	0.0047	–0.0048	0.0023	–0.0058	0.0228
	(–3.6E-07)	(1.554)	(2.3205)	(–1.147)	(0.4514)	(–1.1049)	
Δ_{URP}	0.1308	0.0266	–0.0623	0.0399	–0.0357	–0.012	0.3944
	(1.8133)	(10.2657)	(–6.883)	(2.1194)	(–1.5856)	(–0.5119)	
Δ_{UTS}	0.2593	0.0358	0.1544	–0.0525	–0.0124	–0.013	0.2255
	(1.3978)	(5.3861)	(6.6355)	(–1.0945)	(–0.214)	(–0.2171)	
Δ_{UEMP}	–2.0E-08	–0.0008	–0.0019	–0.0018	–0.0013	–0.0016	0.0061
	(–1.8E-06)	(–1.9781)	(–1.3634)	(–0.5997)	(–0.3707)	(–0.4367)	
Δ_{UIP}	0.2215	–0.0031	–0.008	0.0009	0.0053	0.0354	0.0295
	(5.0261)	(–1.9669)	(–1.4392)	(0.0739)	(0.3836)	(2.4805)	
Wsh-TBill	0.6831	0.1459	0.1327	0.0715	0.0327	0.0666	0.9653
	(13.1537)	(78.454)	(20.3674)	(5.2749)	(2.0218)	(3.9619)	
VW-TBill	0.6596	0.1449	0.1314	0.0694	0.0319	0.0666	0.9694
	(13.6532)	(83.7558)	(21.6745)	(5.5089)	(2.1164)	(4.2571)	
EW-TBill	0.8376	0.168	–0.143	0.0671	–0.0214	0.017	0.9312
	(9.8416)	(55.1106)	(–13.397)	(3.0195)	(–0.8066)	(0.6159)	

We calculate the correlation matrix of the five rotated macroeconomics factors and excess returns of three market indexes that is given in Table 16.

The primary advantages of this approach are: (1) By rotating the original macroeconomic shocks on the statistic factors, it enhances the interpretation of the macroeconomic factors; (2) Rotated macroeconomic factors introduce additional information and link asset-pricing behavior to macroeconomic events.

Testing of the APT-based Henriksson-Merton Model

We test the APT-based Henriksson-Merton model as the follows:

$$R_{pt} - R_{ft} = a_{pt} + \beta_{UTS}\, UN\Delta_{UTS,t} + \beta_{URP}\, UN\Delta_{URP,t} + \beta_{UIP}\, UN\Delta_{UIP,t} + \beta_{UCPI}\, UN\Delta_{UCPI,t} + \beta_{UEMP}\, UN\Delta_{UEMP,t} + \beta_{mktres} MKTRES_t + \gamma put(R_{mt} - R_{ft}) + e_{pt} \quad (30)$$

where, $put(R_{mt} - R_{ft}) = \max(-(R_{mt} - R_{ft}), 0)$ is a European put option on the market index Wilshire5000. $UN\Delta_{UCPI}$, $UN\Delta_{URP}$, $UN\Delta_{UTS}$, $UN\Delta_{UEMP}$, and $UN\Delta_{UIP}$ are rotated macroeconomic shocks on the statistical factors that extracted using maximum likelihood factor analysis.

Table 17 reports the average monthly performance, average risk sensitivities to macroeconomic shocks, and market timing parameter estimated from the APT-based Henriksson-Merton model for the entire sample funds and each of five fund groups.

Comparing to results from the CAPM-based market timing models, testing of the APT-based Henriksson-Merton model, we find a higher percentage of the funds with significantly positive timing since it is easier for managers to correctly time a market that has more risk factors. We find that the fund betas of the term structure risk, the default risk, and the inflation risk are positive and increase monotonically along the risk spectrum from the Asset Allocation funds, Equity Income funds, Growth Income funds, Growth funds, and Aggressive Growth funds. A fund with an inflation beta greater (less) than one has more (less) inflation sensitivity than the market index. The Aggressive Growth funds, Growth funds, and Growth Income funds have more inflation sensitivity than the market index since these funds have average inflation betas greater than one. All funds are significantly positive correlated with unanticipated changes in default risk premium and unanticipated changes in the slope of the term structure of interest rates. The unemployment betas are negative for all the funds and the Growth Income funds have the highest absolute beta. The market residual beta is significantly positive for the Aggressive Growth, Growth, and Growth Income funds, that is, five macroeconomic factors do not completely explain mutual returns since the beta coefficient of the market residual factor is statistical significant for these funds.

Table 16. Correlation Matrix of the Rotated Macrofactors and Excess Return of Market Indexes.

	$UN\Delta_{UCPI}$	$UN\Delta_{URP}$	$UN\Delta_{UTS}$	$UN\Delta_{UEMP}$	$UN\Delta_{UIP}$	Mktres	Wsh-TBill	VW-TBill	EW-TBill
$UN\Delta_{UCPI}$	1								
$UN\Delta_{URP}$	–0.0633	1							
$UN\Delta_{UTS}$	0.8988	0.0713	1						
$UN\Delta_{UEMP}$	–0.6344	–0.3540	–0.8616	1					
$UN\Delta_{UIP}$	–0.7965	–0.2712	–0.6910	0.5153	1				
Mktres	4.2E-16	4.4E-15	7.6E-16	–2.1E-15	4.5E-15	1			
Wsh-TBill	0.5942	0.6410	0.7718	–0.8966	–0.5960	0.1844	1		
VW-TBill	0.5953	0.6428	0.7732	–0.8979	–0.5970	0.1704	0.9994	1	
EW-TBill	0.2621	0.8941	0.4037	–0.6189	–0.4259	0.1211	0.8578	0.8592	1

Table 17. Average Measures Estimated from the APT-based Henriksson-Merton Model.

Fund Objective	No. of Funds	Alpha	$UN\Delta_{UTS}$	$UN\Delta_{URP}$	$UN\Delta_{UCPI}$	$UN\Delta_{UIP}$	$UN\Delta_{UEMP}$	Mktres	Gamma	Adj. R^2	DW
Aggressive	118	–0.1321	0.949	2.2351	1.9396	1.9356	–1.4102	1.039	–0.1848	0.7708	1.7771
Growth		(–1.332)	(2.4696)	(5.5379)	(2.358)	(1.4971)	(–1.1549)	(3.7810)	(–0.519)		
Asset	177	–0.0297	0.7481	0.6561	0.4056	0.3613	–1.0187	0.1949	0.0816	0.8105	1.8138
Allocation		(–0.571)	(3.9637)	(3.1773)	(0.3199)	(0.6495)	(–2.4551)	(1.6872)	(0.4788)		
Equity	149	–0.0014	0.8329	0.8677	0.5993	0.5664	–1.2117	0.0665	–0.0504	0.8573	1.8275
Income		(–0.085)	(2.0729)	(4.5292)	(0.946)	(0.9668)	(–2.8074)	(0.6061)	(–0.519)		
Growth	971	–0.1019	0.8798	1.5973	1.1747	1.4781	–1.6677	0.6008	0.0675	0.8451	1.9005
		(–1.149)	(3.0775)	(5.3494)	(2.0152)	(1.3859)	(–2.6050)	(3.3336)	(0.4183)		
Growth	522	–0.0486	0.8936	1.1008	0.9348	0.9907	–1.7928	0.2918	0.0561	0.9041	1.8616
Income		(–0.992)	(3.6325)	(6.7268)	(1.6375)	(1.7526)	(–5.1135)	(3.4331)	(0.1763)		
All funds	1937	–0.0751	0.8721	1.3605	1.0421	1.2032	–1.5926	0.4658	0.0415	0.8544	1.8692
		(–0.984)	(3.1937)	(5.4738)	(1.6989)	(1.3930)	(–3.1974)	(3.0283)	(0.2640)		

For funds in these three groups, the number of funds with statistically significant Alpha estimated from the APT-based Henriksson-Merton model is more than those estimated from the CAPM-based Henriksson-Merton model. For the Asset Allocation funds and Equity Income funds, the number of funds with statistically significant Alpha estimated from the APT-based Henriksson-Merton model is less than those from the CAPM-based Henriksson-Merton model. The Equity Income funds and Asset Allocation funds are more responsive to these macroeconomic shocks and economic cycles; the Aggressive Growth funds, Growth funds and Growth Income funds are less responsive to these macroeconomic shocks. The Equity Income funds have traditionally favored well-established companies in high-yielding sectors. The Asset Allocation funds utilize three assets: stocks, bonds, and cash. As markets fluctuate, most funds actively shift weightings among these classes. The Aggressive Growth funds usually favor small- and midsize companies in rapidly growing sectors that are less responsive to economic cycles. The Growth funds are more volatile than both Growth Income and Equity Income funds; they tend to be less risky than the Aggressive Growth funds. The Growth Income funds generally split the difference between Growth funds and Equity Income funds. Among all equity funds, these funds look most like the S&P500; the majority has large-cap portfolios with marketlike sector weightings. The APT-based market timing model provides the portfolio manager with better means to assess and control the risk and the expected return of a portfolio than is available through the CAPM. Portfolio managers can develop separate forecasts of each macroeconomic variable and they can use their forecasts to design portfolios that provide the greatest return-to-risk performance based on the sensitivity of each stock to each macroeconomic variable, and the historical risk premiums associated with each macroeconomic variable.

Table 18 reports funds with statistically significant Alpha estimated from the APT-based Henriksson-Merton model. Comparing to the CAPM-based Henriksson-Merton model, we find higher percentage funds with significantly positive timing since it is easier for managers to correctly time the market that has more risk factors. The number of funds with a significantly positive Gamma is more than the number of funds with a significantly negative Gamma.

7. CONCLUSION

Testing various CAPM-based market-timing and selectivity models, we find that about 12% of the funds have a statistically significant Alpha with about 4%

Table 18. Funds with Statistically Significant Alpha and Gamma Estimated from the APT-based Henriksson-Merton Model.

		Statistically Significant Alpha			Statistically Significant Gamma		
Fund Objectives	No. of Funds	Positive Alpha	Negative Alpha	Total Percentage	Positive Gamma	Negative Gamma	Total Percentage
Aggressive Growth	118	1.7% (4.2%)	4.2% (9.3%)	5.9% (13.5%)	9.3% (11.9%)	0.8% (1.7%)	10.1% (12.6%)
Asset Allocation	177	1.7% (3.4%)	7.9% (15.3%)	9.6% (18.7%)	13% (15.8%)	1.7% (2.3%)	14.7% (18.1%)
Equity Income	149	3.4% (7.4%)	0.7% (1.3%)	4.1% (8.7%)	1.3% (4%)	14.1% (17.4%)	15.4% (21.4%)
Growth	971	3.5% (6.1%)	11.3% (19.6%)	14.8% (25.7%)	7.7% (13.1%)	1.8% (3.4%)	9.5% (16.5%)
Growth Income	522	2.9% (5.7%)	10% (13.6%)	12.9% (19.3%)	7.7% (11.9%)	4.6% (7.3%)	12.3% (19.2%)
All Funds	1937	3% (5.7%)	9.4% (15.5%)	12.4% (21.2%)	7.8% (12.2%)	3.4% (5.3%)	11.2% (17.5%)

of the funds having a significantly positive Alpha and 8% of the funds having a significantly negative Alpha. About 15% of funds show significant timing ability with about 9% funds having a significantly positive timing coefficient and 6% of the funds having a significantly negative timing coefficient. The Asset Allocation funds demonstrate the most timing ability and the Aggressive Growth funds demonstrate the least timing ability. This empirical evidence is consistent with the reality that the Asset Allocation funds focus on forecasting the aggregate factor and that the Aggressive Growth funds concentrate on identifying under- or over-valued stocks. We also find a negative correlation between the timing coefficient and Alpha; that is, managers that are good at picking stocks are not good at timing the market, and vice versa. Testing of the APT-based market timing model, we find a higher percentage of the funds with significantly positive timing since it is easier for managers to correctly time a market that has more risk factors. The APT-based market timing model provides a portfolio manager with better means to assess and control the risk and the expected return of a portfolio than is available through the CAPM-based market timing models. This is especially true for the Asset Allocation funds and Equity Income funds, since these funds are more responsive to economic cycles.

NOTES

1. Hattacharya and Pfleiderer (1983) has shown that if the investors in whose interest the fund is being managed have a coefficient of absolute risk aversion equal to α, then $\theta = 1/[\alpha \operatorname{var}(\tilde{\pi}_t | \phi_t)]$, and $\beta_{pT} = \theta E(\tilde{R}_m)$.

2. In deriving the probability limits of $\hat{\eta}_0$, $\hat{\eta}_1$, and $\hat{\eta}_2$, Jensen (1972) implicitly assumed independence between $\tilde{\pi}$ and $\tilde{\upsilon}_t$. This is not true because $\tilde{\pi}_t^*$ and $\tilde{\pi}_b$ are joint normally distributed. It is possible to write: $\tilde{\pi} = d_0^* + d_1^* \tilde{\pi}_t^* + \tilde{\upsilon}_t^*$, where $\tilde{\upsilon}_t^*$ is normally distributed and independent of $\tilde{\pi}_t^*$. If $\tilde{\pi}_t^*$ is the optimal forecast, then, $d_0^* = 0$, and $d_1^* = 1$. However, if we write $\tilde{\pi}_t^* = d_0 + d_1 \tilde{\pi}_1 + \tilde{\upsilon}_t$. We cannot have $d_0 = 0$ and $d_1 = 1$, and $\tilde{\upsilon}_t$ independent of $\tilde{\pi}_t$.

3. Ferson and Schadt (1996) assume that market prices fully reflect readily available, public information. They hypothesize that managers may use this information to determine their portfolio strategies. The use of public information should not imply abnormal performance, under semi-strong form market efficiency, because investors can replicate on their own any strategy that depends on public information. The argument that investors can replicate or undo managers' trades that are based on public information assumes that investors can infer the trades. It also ignores any cost advantages in trading that funds may have over investors and assumes that managers do not waste resources by churning their clients portfolios at cost.

4. Connor and Korajczyk (1991) reestimated the Henriksson-Merton model using risk-sorted portfolios of mutual funds approach. Connor and Korajczyk (1991) group the mutual funds into five risk classes: income, stability-growth-income, growth-income, growth, maximum capital gain, according to the Weisenberger Investment Survey classifications. Comparing to Chang and Lewellen (1984) and Henriksson (1984), in which they use individual mutual fund returns in their empirical studies, the risk-sorted approach has advantage of lower residual variance in the regression, therefore, is more precise estimates of parameters. An obvious potential disadvantage is the masking of interesting cross-sectional differences in funds of the same type. Empirically, this data reduction technique significantly strengthens the finds of the two earlier studies

5. The dynamic trading model argues that it could replicate a put option following a dynamic trading strategy. If fund managers trade more frequently than we observe returns, then their dynamic trading decisions (without any superior information) can create false evidence of timing. With continuous trading, they can perfectly replicate a put without any special information. The asset beta nonlinearities model argues that the underlying return process driving assets could have beta non-linearities arising from leverage effects or from other sources that give rise to a putlike structure to returns.

6. Glosten and Jagannathan (1988) developed a contingent-claim approach to performance evaluation for portfolio managers who trade a variety of options on optionlike securities.

7. AIMR (American Investment Management Research) prefers that performance results be presented gross (before deduction) of management fees. This is because a manager's fee schedule is usually scaled to size of assets. Therefore, performance results after deduction of an average management fee will not be representative of results for a portfolio that is much larger or much smaller than the size of the portfolio represented by the average fee. AIMR feels it is more representative to show results

before the deduction of management fees and to provide a fee schedule that represents the fee that would actually be paid by the prospective client. In addition, because fees are sometimes negotiable, presenting performance gross of fees shows the manager's expertise in managing assets without the impact of negotiating skills on the part of the manager or manager's clients.

8. Heteroskedasticity affects the size of the standard error of the regression coefficient, thereby biasing hypothesis-test results. The effect on standard error will depend on the exact manner in which the heteroskedasticity was formed. The GLS procedure can be used when a residual covariance matrix is known. The FGLS and MLE are used when a residual covariance matrix is unknown.

REFERENCES

Admati, A. R., Bhattacharya, S., & Pfleiderer, P. (1986). On Timing and Selectivity. *Journal of Finance*, *41*, 715–730.

Berry, M. A., Burmeister, E., & McElroy, M. B. (1988). Sorting out Risks Using Known APT Factors. *Financial Analyst Journal*, *44*(2), 29–42.

Bhattacharya, S., & Pfleiderer, P. (1983). A Note on Performance Evaluation. Technical Report 714. Graduate School of Business, Stanford University, CA.

Black, F., Jensen, M., & Scholes, M. (1972). The Capital Asset Pricing Model: Some Empirical Tests. In: M. C. Jensen (Ed.), *Studies in the Theory of Capital Markets*. New York. Praeger.

Blume, M. (1971). On the Assessment of Risk. *Journal of Finance*, 1–10.

Bodurtha, J. N. Jr., & Nelson, C. M. (1991). Testing the CAPM with Time-Varying Risks and Returns. *Journal of Finance*, *46*(4), 1485–1506.

Breen, W., Jagannathan, R., & Ofer, A. R. (1986). Correcting for Heteroskedasticity in Tests for Market Timing Ability. *Journal of Business*, *59*(4), Part 1, 585–598.

Brown, S. J., & Goetzmann, W. N. (1995). Attrition and Mutual Fund Performance. *Journal of Finance*, *50*, 679–698.

Burmeister, E., & McElroy, M. B. (1991). The Residual Market Factor, the APT, and Mean-Variance Efficiency. *Review of Quantitative Finance and Accounting*, *1*(1), 27–50.

Burmeister, E., & McElroy, M. B. (1988). Joint Estimation of Factor Sensitivities and Risk Premia for the Arbitrage Pricing Theory. *Journal of Finance*, 721–275.

Chang, E. C., & Lewellen, W. G. (1984). Market Timing and Mutual Fund Investment Performance. *Journal of Business*, *57*(1), Part 1, 57–72.

Chen, C. R., & Stockum, S. (1986). Selectivity, Market Timing, and Random Beta Behavior of Mutual Funds: A Generalized Model. *Journal of Financial Research*, *9*(1), 87–96.

Chen, N. F., Roll, R. R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *Journal of Business*, *59*, 383–403.

Connor, G., & Korajczyk, R. A. (1988). Risk And Return in an Equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics*, *21*(2), 255–290.

Connor, G., & Korajczyk, R. (1991). The Attributes, Behavior, and Performance of U.S. Mutual Funds. *Review of Quantitative Finance and Accounting*, *1*, 5–26.

Copeland, T. E., & Weston, J. F. (1988). *Financial Theory and Corporate Policy* (3rd ed.). Addison-Wesley.

Cornell, B. (1979). Asymmetric Information and Portfolio Performance. *Journal of Financial Economics*, *7*, 381–390.

Elton, E. J., & Gruber, M. J. (1995). *Modern Portfolio Theory and Investment Analysis* (5th ed.). John Wiley & Sons.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1995). Fundamental Economic Variables, Expected Returns, and Bond Fund Performance. *Journal of Finance*, *50*, 1229–1256.

Evans, M. D. D. (1994). Expected Returns, Time-Varying Risk and Risk Premia. *Journal of Finance*, *49*, 655–680.

Fama, E. F. (1972). Components of Investment Performance. *Journal of Finance*, *27*, 551–567.

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy*, *38*, 607–636.

Ferson, W. E., & Schadt, R. W. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. *Journal of Finance*, *51*, 425–461.

Grant, D. (1978). Market Timing and Portfolio Management. *Journal of Finance*, *33*(4), 1119–1131.

Greene, W. H. (1990). *Econometric Analysis*. New York: Macmillan.

Grinblatt, M., & Titman, S. (1994). A Study of Mutual Fund Returns and Performance Evaluation Techniques. *Journal of Financial and Quantitative Analysis*, *29*, 419–444.

Grossman, S., & Stiglitz, J. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, *70*, 393–408.

Gruber, M. J. (1996). Another Puzzle: The Growth in Actively Managed Mutual Funds. *Journal of Finance*, *51*, 783–810.

Haugen, R. A. (1997). *Modern Investment Theory* (4th ed.). Prentice-Hall.

Henriksson, R. D. (1984). Market Timing and Mutual Fund Performance: An Empirical Investigation. *Journal of Business*, *57*(1), Part 1, 73–96.

Henriksson, R. D., & Merton, R. C. (1981). On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business*, *54*(4), 513–534.

Ippolito, R. A. (1989). Efficiency with Costly Information: A Study of Mutual Fund Performance 1965–1984. *Quarterly Journal of Economics*, *104*, 1–23.

Jagannathan, R., & Korajczyk, R. A. (1986). Assessing the Market Timing Performance of Managed Portfolios. *Journal of Business*, *59*(2), Part 1, 217–236.

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945–1964. *Journal of Finance*, *23*, 389–416.

Jensen, M. C. (1969). Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios. *Journal of Business*, *42*, 167–247.

Jensen, M. C. (1972). Optimal Utilization of Market Forecasts and the Evaluation of Investment Performance. In: Szego & Shell (Eds), *Mathematical Methods in Investment and Finance*. Amsterdam: North Holland/American Elsevier.

Kon, S. J., & Jen, F. C. (1979). The Investment Performance of Mutual Funds: An Empirical Investigation of Timing, Selectivity, and Market Efficiency. *Journal of Business*, *52*(2), 263–290.

Lee, C-f., & Rahman, S. (1990). Market Timing, Selectivity, and Mutual Fund Performance: An Empirical Investigation. *Journal of Business*, *63*(2), 261–278.

Levy, H., & Sarnat, M. (1983). *Portfolio and Investment Selection: Theory and Practice*. Prentice-Hall.

Li, L., & Lee, C-f. (1998). Mutual Fund Performance Evaluation: the CAPM vs. Multi-factor Models. Working paper. Rutgers University.

Malkiel, B. G. (1995). Returns from Investing in Equity Mutual Funds 1971–1991. *Journal of Finance*, *50*, 549–572.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, *7*(1), 77–91.
Markowitz, H. (1970). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley.
Markowitz, H., & Perold, A. (1981). Portfolio Analysis with Factors and Scenarios. *Journal of Finance*, *36*(14), 871–877.
Merton, R. C. (1981). On Market Timing and Investment Performance I. An Equilibrium Theory of Value for Market Forecasts. *Journal of Business*, *54*(3), 363–406.
Miller M. H., & Scholes, M. (1972). Rates of Return in Relation to Risk: A Re-Examination of Some Recent Findings. In: M. C. Jensen (Ed.), *Studies in the Theory of Capital Markets*. Praeger.
Pfeifer, P. E. (1985). Market Timing and Risk Reduction. *Journal of Financial and Quantitative Analysis*, *20*(4), 451–460.
Reilly, F. K. (1994). *Investment Analysis and Portfolio Management* (4th ed.). Dryden Press.
Roll, R. (1980). Performance Evaluation and Benchmark Errors. *Journal of Portfolio Management*, *6*(4), 5–12.
Roll, R. (1981). Performance Evaluation and Benchmark Errors (II). *Journal of Portfolio Management*, *7*(2), 17–22.
Ross, S. A. (1976). The Arbitrage Pricing Theory of Capital Asset Pricing. *Journal of Economic Theory*, *13*, 341–360.
Roll, R., & Ross, S. (1980). An Empirical Investigation of the Arbitrage Pricing Theory. *Journal of Finance*, *35*, 1073–1103.
Sharpe, W. F. (1963). A simplified Model for Portfolio Analysis. *Management Science*, *9*(2), 277–293.
Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, *19*(3), 425–442.
Sharpe, W. F. (1975). Are Gains Likely From Market Timing. *Financial Analyst Journal*, *31*(2), 60–69.
Sharpe, W. F. (1975). Adjusting for Risk in Portfolio Performance Measurement. *Journal of Portfolio Management*, (Winter), 29–34.
Sharpe, W. F. (1984). Factor Models, CAPMs and the APT. *Journal of Portfolio Management*, *11*, 21–25.
Sharpe, W. F. (1995). The Styles and Performance of Large Seasoned U.S. Mutual Funds, 1985–1994. http://gsb-www.stanford.edu/~wfsharpe/ls100.html
Tobin. J. (1965). The Theory of Portfolio Selection. In: F. Hain & F. Breechling (Eds), *The Theory of Interest Rates* (pp. 3–51). London: MacMillan.
Treynor, J., & Black, F. (1973). How to Use Security Analysis to Improve Portfolio Selection. *Journal of Business*, *45*(1), 66–86.
Treynor, J., & Mazuy, M. (1966). Can Mutual Funds Outguess the Market? *Harvard Business Review*, (July/August), 131–136.
Tucker, A. L., Becker, K. G., Isimbabi, M. J., & Ogden, J. P. (1994). *Contemporary Portfolio Theory and Risk Management*. West Publishing Company.

SOURCES OF TIME-VARYING RISK PREMIA IN THE TERM STRUCTURE

John Elder

ABSTRACT

This paper investigates the extent to which three observable macroeconomic factors can explain the time-varying risk premia in the short-end of the term structure. We employ an empirical model that is motivated by a dynamic asset pricing model with time-varying risk premia and time-invariant reward-to-volatility measures. We find that, in our model, two factors explain up to 65% of the temporal variation in Treasury bill returns, with the short-end of the term structure responding significantly to contemporaneous innovations the funds rate and shifts (or twists) in the yield curve. Our primary new findings are that a factor based on shifts in the yield curve may explain the time-variation in risk premia at the very short end of the term structure, and that a factor based on innovations in the federal funds rate may be weakly linked to the time-varying risk premia over the post-1966 sample, when the federal funds market first began to function as a major source of bank liquidity. This latter result is somewhat sensitive to the sample period.

1. INTRODUCTION

Early empirical investigations of the link between asset returns and macroeconomic factors identified several factors, such as yield spreads and unexpected changes in output, that are empirically informative about asset

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 85–108.

ISBN: 0-7623-0887-7

returns. These early investigations were based on either simple linear regressions, as in Chen, Roll and Ross (1985), or an arbitrage pricing theory (APT) with static risk premia, as in McElroy and Burmeister (1988). More recently, several authors, such as Thorbecke (1997), Patelis (1997) and Madura and Schnusenberg (2000), have found that innovations to monetary policy are also empirically informative about asset returns within these static frameworks. Most financial data, however, display evidence of either time-vary betas or time-varying risk premia, and there exists a large and fruitful empirical literature in this vein.[1] It seems desirable and relevant, therefore, to assess the relationship between asset returns, monetary policy and other relevant macroeconomic factors within a framework that accommodates time-varying risk-premia.

More specifically, this paper investigates the extent to which three observable macroeconomic factors, including a factor based on innovations in monetary policy, can explain the *time-variation in the risk premia* in the short-end of the term structure. Our empirical model is motivated by a modern, dynamic asset pricing model that can be viewed as a multi-factor intertemporal capital asset pricing model (ICAPM) with conditionally heteroskedastic factors, although similar asset pricing equations can also be derived under a dynamic "no-arbitrage" framework or the "stochastic discount factor" framework, as detailed by Cochrane (2001). In our model, the risk premia are time-varying, while the reward-to-volatility parameters are not. Our empirical model is similar to those estimated by, for example, Engle Ng and Rothschild (1990), King, Sentana and Wadhwani (1994) and Flannery, Hameed and Harjes (1997).

We posit three macroeconomic factors which have previously been found to be informative for bond returns in static frameworks. Two of the macro-economic factors are based on bond portfolios and have been widely used since they were initially proposed by Chen, Roll and Ross (1986). The first bond factor is the return on long-term government bonds minus the return on short-term governments, which reflects the risk premia associated with long-term bonds. The second bond factor is the spread between long-term corporate bonds and long-term government bonds, which reflects the default premium associated with corporate bonds.

The third factor we posit is based on the federal funds rate. Innovations in the federal funds rate, as the primary tool of monetary policy, may reflect the risk of future shifts in the term structure (through a mechanism such as the expectations hypothesis), or the risk associated with macroeconomic fundamentals, such as output and inflation, which in turn affect the conditional distribution of asset returns. Previous authors, including Thorbecke (1997), have found that innovations in the funds rate are priced in the context of a static APT. The component of the funds rate relevant for our analysis is the portion of the contemporaneous realization that is unanticipated, or orthogonal to the contemporaneous information set. We construct these innovations by filtering the funds rate in a manner consistent with the large and recent VAR literature on the identification of monetary policy shocks, such as Bernanke and Gertler (1995) and Christiano, Eichenbaum and Evans (1996).

Our empirical results indicate that the yield curve and funds rate factors explain nearly 65% of the variation in T-bill returns, and that T-bill returns respond significantly to innovations in these factors. We also find some evidence that the volatility of the yield spread and the volatility of funds rate innovations explain the time-varying risk premia in T-bills over the sample beginning in 1966, when the federal funds market began to function as a major source of bank liquidity. This suggests that the risk premia at the short end of the term structure has tended to rise during periods when the expected volatility of these macroeconomic factors is high.

The paper is organized as follows: Section 2 describes the asset-pricing framework that motivates the empirical specification; Section 3 describes construction of the macroeconomics factors; Section 4 presents the data and summary statistics; Section 5 presents the empirical results and reconciles the results with earlier findings. Section 6 concludes.

2. THEORETICAL AND EMPIRICAL MODEL

This section provides some background on the asset pricing framework that forms the basis for the empirical specification. We derive the pricing equation from a discrete time version of Merton's (1973) intertemporal CAPM, which is probably the framework most familiar and intuitive to readers. Similar asset pricing restrictions can be derived directly under more general and rigorous frameworks, such as a "no-arbitrage" framework of King, Sentana and Wadhwani (1994) and the "stochastic discount factor" framework of Cochrane (1996). We derive and estimate a fully specified empirical model, which avoids the concerns raised by Kan and Zhou (1999).

The standard equilibrium pricing equation implied by the discrete time version of the intertemporal CAPM is

$$\mathrm{E}_{t-1}(r_{i,t} - r_{0,t}) = \frac{\mathrm{cov}_{t-1}(r_{i,t}, r_{m,t})}{\mathrm{var}_{t-1}(r_{m,t})} \mathrm{E}_{t-1}(r_{m,t} - r_{0,t}) \tag{2-1}$$

where E_{t-1}, cov_{t-1} and var_{t-1} represent moments conditional on the information set available at the end of $t-1$; $r_{i,t}$ is the return on asset i at time

t; $r_{0,t}$ is the return on a zero-beta portfolio; and $r_{m,t}$ is the stochastic return on a benchmark portfolio, not necessarily observable, that is perfectly correlated with the "stochastic discount factor", as detailed in, for example, Cochrane (2001) and Campbell (2000). As noted by Campbell (2000), if the economy has a representative agent with well defined utility, then the stochastic discount factor is the discounted rate of change in the marginal utility of consumption. Otherwise, the existence of a unique stochastic factor can be inferred from a no-arbitrage framework with complete markets.

We also assume realized asset returns are a linear function K factors, or state variables, that are conditionally mean zero and possibly heteroskedastic

$$r_{i,t} - r_{0,t} = E_{t-1}(r_{i,t} - r_{0.t}) + \sum_{k=1}^{K} b_{i,k} f_{k,t} + \varepsilon_{i,t} \quad \text{for } i = 1, 2, 3, \cdots, N \text{ and } i = m \tag{2-2}$$

where $f_{k,t}$ denotes factor k at time t and $b_{i,k}$ denotes the loading on factor k for asset i. The idiosyncratic error $\varepsilon_{i,t}$ is assumed conditionally mean zero and conditionally orthogonal to the factors, although we allow for correlation of idiosyncratic risk across assets, which would be likely in the event that a relevant factor is inadvertently omitted. We also assume for expositional simplicity that the factors are conditionally orthogonal or have constant conditional covariances. In practice, many empirical applications do not explicitly model the correlation between the factors.[2] These conditions can be stated more precisely as

$$\begin{aligned} &\mathrm{E}_{\mathrm{t}-1}(f_{k,t}) = 0 \qquad (2\text{-}3)\\ &\mathrm{cov}_{\mathrm{t}-1}(f_{j,t}, f_{k,t}) = \sigma_{j,k}(t) \ or\ 0\ for\ j \neq k\\ &\mathrm{var}_{\mathrm{t}-1}(f_{k,t}) = \sigma^2_{k,t}\\ &\mathrm{E}_{\mathrm{t}-1}(\varepsilon_{i,t}) = \mathrm{E}_{\mathrm{t}-1}(\varepsilon_{i,t} \mid f_{1,t}, f_{2,t}, \cdots, f_{k,t}, \varepsilon_{m,t}) = 0\\ &\mathrm{var}_{\mathrm{t}-1}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Psi} \end{aligned}$$

Note that (2-2) is required to hold for the benchmark portfolio $r_{m,t}$, which is perfectly correlated with the stochastic discount factor. This suggests that plausible macroeconomic factors should either be good proxies for the growth in marginal utility, or state variables that describe the conditional distribution of future asset returns.

The equilibrium condition that relates asset risk premia (or ex-ante returns) to the factors is derived by combining the stochastic linear process generating realized returns (2-2) with the ex-ante pricing restrictions implied by the intertemporal CAPM (2-1). This gives

$$E_{t-1}(r_{i,t} - r_{0,t}) = c_i + \sum_{k=1}^{K} \xi_{i,k}\sigma_k^2(t) \tag{2-4}$$

where

$$c_i = \delta \sum_{j=1}^{K} \sum_{\substack{k=1 \\ k \neq j}}^{K} b_{i,j} b_{m,k} \sigma_{j,k}$$

$$\delta = E_{t-1}(r_{m,t} - r_{0,t})/\mathrm{var}_{t-1}(r_{m,t})$$

and

$$\xi_{i,k} = \delta b_{i,k} b_{m,k}.$$

The constant c_i captures the constant conditional covariance terms, which equals zero if the factors are conditionally orthogonal or if there is only a single factor. The parameter δ is the additional return in the benchmark portfolio required by investors with respect to a one-unit increase in its variance, which is known as the price of market covariance risk. For empirical implementations, the usual assumption, as in for example Harvey (1989), Engle, Ng and Rothschild (1990) and Harvey (1991), is that the price of market covariance risk is time invariant. The parameter $\xi_{i,k}$ summarizes the relationship between the conditional volatility of the kth state variable and the required return for asset i.

One intuition behind such dynamic equilibrium pricing equations is that they represent a sort of combination of the intertemporal CAPM and a factor-based model such as the static APT. Under the assumptions required for the intertemporal CAPM, the appropriate measure of an asset's risk is the conditional covariance with the market portfolio, while under the static APT the appropriate measure of risk is an asset's exposure to the constant price of risk associated with the state variables. In this dynamic factor model, the appropriate measure of risk is instead an asset's exposure to the conditional volatility of each of the factors.

Substituting the equilibrium condition (2-4) into (2-2) yields the equation to be estimated

$$r_{i,t} - r_{0,t} = c_i + \sum_{k=1}^{K} \xi_{i,k}\sigma_{k,t}^2 + \sum_{k=1}^{K} b_{i,k} f_{k,t} + \varepsilon_{i,t} \tag{2-5}$$

Note that this derivation implies several cross-equation parametric restrictions. For example, if the asset pricing model were literally correct, the ratio $\xi_{i,k}/b_{i,k}$ would be constant across assets (i.e. $\xi_{i,k}/b_{i,k} = \delta_{j,k}/b_{j,k} = \delta b_{m,k}$) in the absence of sampling error, and, in a single factor model, the market risk-return ratio δ would not be identified independently of the market beta $b_{m,1}$. These restrictions, however, are partially the result of the particular framework used to motivate the pricing equation, and closely related pricing equations may be derived by employing the framework of Campbell (2000) (c.f. Eqs (7) and (8)), or by employing the no-arbitrage framework of King, Sentana and Wadwhani (1994).[3] It is important to note that we are interested primarily in the empirical relationship between asset returns measured macroeconomic factors, rather than in testing the parametric restrictions implied by a particular model. Therefore, we do not impose these restrictions on the data. Rather, we estimate $\xi_{i,k}$ as a free parameter, sacrificing N − 1 degrees of freedom in a single factor model. This approach is not entirely novel. For example, Flannery, Hameed and Harjes (1997) and Engle, Ng and Rothschild (1990) also estimate the unrestricted versions of Eq. (2-5). Scruggs (1998) estimates an empirical model closely related to (2-5).

3. EMPIRICAL SPECIFICATION FOR MACROECONOMIC FACTORS

To estimate the empirical model we also need to specify the state variables and the stochastic processes governing the evolution of their conditional means and variances. This yields estimates of the factor innovations $f_{k,t}$ and conditional variances $\sigma_{k,t}$ in Eq. (2-5). To be consistent with the theoretical model, the state variables should proxy for the growth in marginal utility (i.e. investor sentiment), or should reflect the future conditional distribution of asset returns. We propose three candidate factors: (1) the excess return on long-term government bonds, denoted *TERM*; (2) the difference between the returns on long-term corporate and long-term government bonds, denoted *DFLT*; and (3) a measure of innovations in monetary policy based on the federal funds rate, denoted *FUNDS*. The two bond factors were proposed by Chen, Roll and Ross (1986) and have been widely applied since. For example, McElroy and Burmeister (1988) find these factors to be priced in the context of a static APT. Fama and French (1993) find that these two factors have explanatory power in linear regressions of intermediate and long term government bond portfolios.

The raw data for *TERM* is the difference between the returns on 20-year government bonds and 1-month Treasury bill, and it reflects risk premia associated with long term bonds, which may be related to inflation expectations and liquidity premiums. The raw data for *DFLT* is the difference between the return on 20-year corporate bonds and 20-year government bonds as published in the CRSP Indices File. The *DFLT* factor reflects the risk premia associated with possible default of corporate bonds, and may therefore be related to expectations about future real economic activity. Additional descriptions of the data sources are given in Table 1.

We assume that each of the bond factors have a stationary, invertible finite-order ARIMA representation, where the innovations $f_{k,t}$ follow a univariate GARCH(p, q) process

$$\mathrm{var}(f_{k,t}) = \sigma^2_{k,t} = \alpha_0 + \sum_{j=1}^{p} \beta_j \sigma^2_{k,t-j} + \sum_{i=1}^{q} \alpha_i f^2_{k,t-1}. \tag{3-1}$$

In practice, we fit a low-order autoregressive GARCH process to each of the bond factors, with the autoregressive order selected by the Schwartz criterion, and the order of the GARCH process determined by sequential Lagrange-multiplier tests. This yields estimates of the innovations $f_{k,t}$ and conditional variances $\sigma_{k,t}$ of the bond factors.

The funds rate factor (*FUNDS*) is a measure of innovations in monetary policy and is based on orthogonalized innovations in the federal funds rate.[4] As noted by Thorbecke (1997) and others, *FUNDS* may also be related to expectations about future real economic activity. We filter the conditional mean of the funds rate so that it is comparable to the policy innovations identified by

Table 1. Macroeconomic Factors.

Factor (1) is based on Bernanke and Gertler (1995), and is similar to those specified in Strongin (1995) and Christiano, Eichenbaum and Evans (1996). The lag-length was determined by sequential likelihood-ratio tests with small sample correction under an initial alternative of 12 lags.

Factor	Description
FUNDS	Orthogonalized shocks to the federal funds rate from 9 lag MGARCH-VAR of price, output, commodity prices and funds. Price is measured by the consumer price index less shelter, output is measured by industrial production and commodity prices are measured by an index of sensitive commodity prices. Data are from CITIBASE.
DFLT	Percent return on AAA long-term corporate bond minus the return on long-term Treasury bond. Data are from the CRSP Indices File.
TERM	Percent return on long-term Treasury bond minus the return on one-month Treasury bill. Data are from the CRSP Indices File.

the substantial empirical literature on the dynamic effects of monetary policy, such as Bernanke and Gertler (1995) and Christiano, Eichenbaum and Evans (1996). The identification scheme employed by these authors is usually a vector autoregressive (VAR) framework with a Cholesky decomposition. We construct our policy innovations similarly, although, as is discussed in detail below, we relax the usual assumption of homoskedasticity by estimating a multivariate VAR with GARCH. Our benchmark VAR-filter is based on to Bernanke and Gertler (1995) and consists of 9 lags of prices as measured by the CPI less shelter, industrial production, an index of sensitive commodity prices and the funds rate. The lag-length is determined by sequential likelihood-ratio tests with small sample correction with an initial alternative of 12 lags.

Note that the GARCH specification of the conditional variance for the funds rate factor is at least consistent with the stylized facts associated with the Federal Reserve's policy making behavior. The policy making committee at the Fed usually meets at six-week intervals to update policy targets in response to changing macroeconomic conditions. Their tendency to smooth changes in policy targets results in the targets being updated relatively frequently, and in small increments, during the transition phases of the business cycle. Invariably, financial markets speculate most intensely, and are most uncertain, about impending policy innovations during these transition phases, suggesting that a symmetric and clustered measure of funds rate volatility, such as that produced by GARCH, may be a reasonable specification.

More precisely, we estimate an N-variable VAR of the form

$$\mathbf{B}\mathbf{y}_t = \mathbf{C} + \mathbf{\Gamma}_1\mathbf{y}_{t-1} + \mathbf{\Gamma}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Gamma}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t. \tag{3-2}$$

where $\dim(\mathbf{B}) = \dim(\mathbf{\Gamma}_i) = (N \times N)$, and

$$\mathbf{z}_t \sim \text{iid } N(\mathbf{0}, \mathbf{I})$$

$$\boldsymbol{\varepsilon}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{z}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_t) \text{ conditional on the information set at } t-1.$$

The matrix $\mathbf{\Sigma}_t$ is the covariance matrix of the *structural disturbances*, and $\mathbf{B}$ is the linear operator mapping the forecast errors to the N orthogonal disturbances from the primitive data generating process (i.e. $\mathbf{B}[\mathbf{y}_t - E_{t-1}(\mathbf{y}_t)] = \boldsymbol{\varepsilon}_t$). The usual identifying restrictions in a just-identified VAR are that $\mathbf{B}$ is lower triangular.[5] with the diagonal elements normalized to one, and that the structural disturbances are contemporaneously uncorrelated (i.e. $\mathbf{\Sigma}_t$ is diagonal). By estimating the structural parameters of equation (3-2) directly, we can exploit the diagonality of $\mathbf{\Sigma}_t$ to choose a parsimonious parameterization for the multivariate GARCH variance function. We therefore do not require the other common multivariate GARCH specifications which are detailed in Engle and Kroner (1995). Rather, we need only to specify the variance function as

$$\mathrm{diag}(\boldsymbol{\Sigma}_t) = \mathbf{C}_v + \mathbf{F}_1 \mathrm{diag}(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1}) + \mathbf{G}_1 \mathrm{diag}(\boldsymbol{\Sigma}_{t-1}) \tag{3-3}$$

where $\mathrm{diag}(\cdot)$ is the function that extracts the diagonal of a square matrix, and $\mathbf{F}_1$ and $\mathbf{G}_1$ are diagonal $N \times N$ matrices and $\mathbf{C}_v$ an $N \times 1$ vector. This specification imposes the zero covariance restriction of identified VARs, and it models the conditional variance of each variable as a function of its own past innovations and variances.

The only other required modification to the usual homoskedastic VAR is to transform the conditional mean of the non-stationary variables, so we first difference the log of the CPI and the log of industrial production. Bernanke and Blinder (1992) and Strongin (1995), among others, indicate that their measures of policy innovations are robust to this transformation. Note also that if the true VAR is with these variables differenced, then the conditional mean parameters are estimated consistently in either levels or first differences. The usual interpretation of the conditional mean of the VAR is therefore not altered. The policy factor is then the Nth element of $\boldsymbol{\varepsilon}_t$. The conditional variance of the policy factor, or policy volatility, is the $\{N, N\}$ element of $\boldsymbol{\Sigma}_t$. In our specification, the funds rate is ordered last in the four-variable VAR.

Given the data $r_t, f_{1,t}, f_{2,t}, \cdots, f_{K,t}$ and some distributional assumptions, the parameters of Eqs (2-5) and (3-1) or (3-3) are then estimated by the following procedure. First, fit a GARCH model (or a VAR with GARCH) with the appropriate conditional mean to the macroeconomic factor. This yields an estimate of the conditional variance of the factor $\hat{\sigma}^2_{k,t}$ and of innovations in the factor $\hat{f}_{k,t}$. The second step is to estimate (2-5) by regressing the asset return data on the factor and the conditional variance of the factor. Joint estimation of the factors and the asset pricing model was avoided because of the large number of parameters involved and the complexity of likelihood function.

Note that the regressors in the second stage, $\hat{f}_{k,t}$ and $\hat{\sigma}^2_{k,t}$, are "generated", the properties of which are analyzed by Pagan (1984), Murphy and Topel (1985) and Pagan and Ullah (1988). The theory of two step maximum likelihood indicates that the coefficients on $\hat{f}_{k,t}$ and $\hat{\sigma}^2_{k,t}$ are consistent and asymptotically Gaussian. If there is no measurement error in $\hat{f}_{k,t}$ and $\hat{\sigma}^2_{k,t}$, then the usual standard errors are also consistent. If there is measurement error in $\hat{f}_{k,t}$ and $\hat{\sigma}^2_{k,t}$ then the usual standard errors are consistent only under the null hypotheses that the true coefficients are zero. It is therefore valid asymptotically to use the usual standard errors for the test that the coefficients are zero. Under the alternative hypothesis that the coefficients on $\hat{f}_{k,t}$ and $\hat{\sigma}^2_{k,t}$ are not zero, the usual standard errors are not consistent, but Murphy and Topel (1985) derive a procedure for calculating standard errors that are consistent, which is described in the Appendix. We therefore calcuate both the Murphy-Topel standard errors and the maximum likelihood standard errors for the single factor models.

4. DATA AND SUMMARY STATISTICS

The class of assets we consider represent the short-end of the term structure. These are the monthly holding-period excess returns on T-bills maturing 2, 4, 6, 8, 10 and 12-month T-bills. This class of assets is similar to that examined by Engle, Ng, and Rothschild (1990), except our sample period begins several years prior, in 1966:01, when the federal funds market began to function as a major source of bank liquidity (see e.g. Meulendyke, 1989). Beginning the sample 1966 is an important component of our analysis, both for the empirical results and because it coincides with structural features associated with monetary policy. The returns are calculated from monthly yields on discount bonds published by McCulloch and Kwon (1993). McCulloch and Kwon's yield data are derived from asset prices on the afternoon of the last business day of each month. The total return on January's 6-month Treasury bill, for example, is the percentage change in the price from December 31 to January 31 of a bill maturing on June 30. Excess returns are calculated by subtracting the return on 1-month T-bills. The sample terminates in 1991:02, with the end of the McCulloch and Kwon data set.

The return and squared return series are plotted in Figs 1 and 2. The T-bill returns, not surprisingly, appear highly correlated, although the magnitude of volatility increases substantially with maturity. For example, the vertical axis for the squared 12-month T-bill returns is about 75 times greater than the scale for the squared return on 2-month T-bills. The pair-wise correlations, reported in Tables 2 and 3, are above 0.73 in levels and above 0.79 in squares. The standard deviation varies from a low of 0.06 for 2-month T-bills to 0.64 for 12-month T-bills, compared with means of 0.04 to 0.10. All the T-bill returns and factors are most volatile during the 1979:10–1982:10 time period, when the Federal Reserve allowed the funds rate to fluctuate over a much larger band.

The T-bill return series is most highly correlated with the *TERM* factor – about 0.70 in both levels and squares, suggesting that *TERM* might prove to be an informative factor. The correlation of both the *FUNDS* and *DFLT* factors with the T-bill returns is negative, as economic theory might suggest. The correlation with *FUNDS* is fairly high in absolute value, about 0.40, and somewhat lower in squares, about 0.25. The correlation of T-bills with *DFLT* is relatively low in both levels and squares.

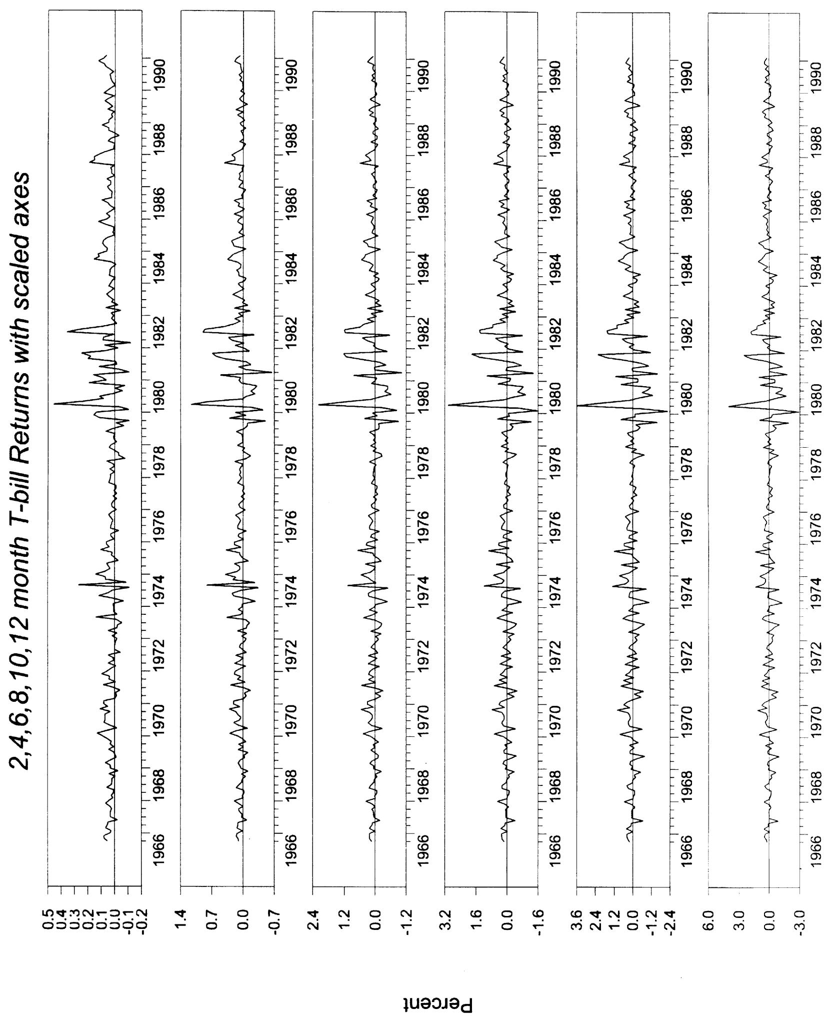


Fig. 1. Monthly Excess Returns.

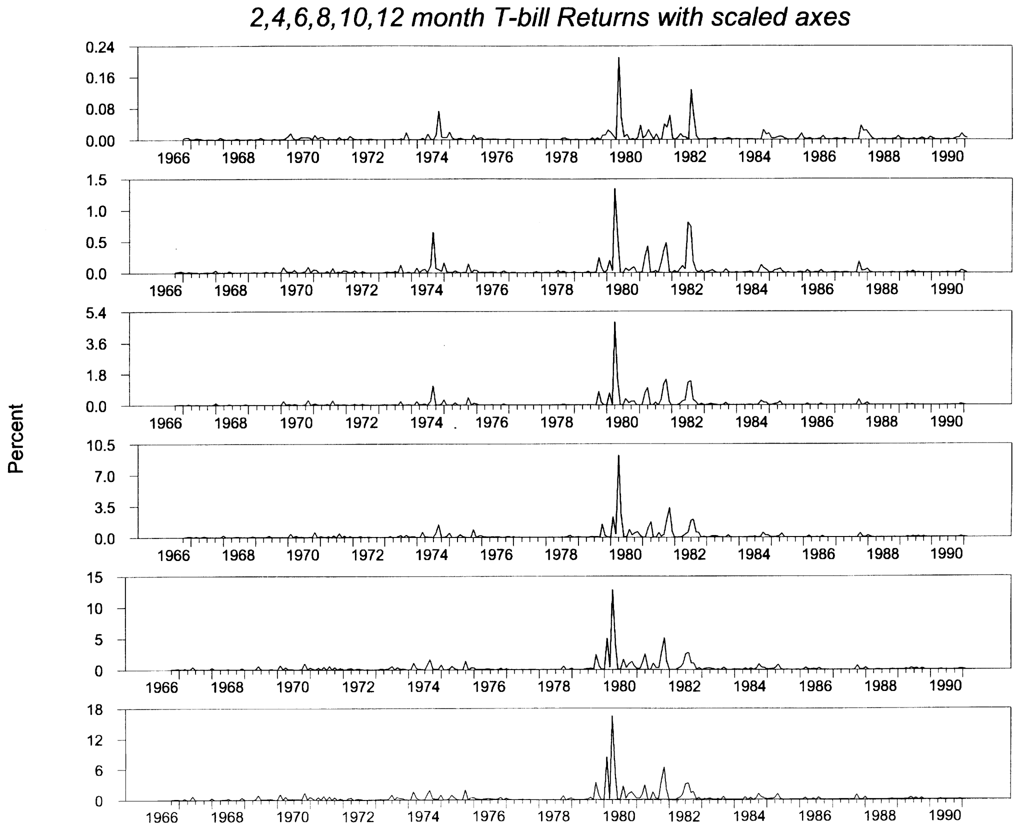


Fig. 2. Squared Monthly Excess Returns.

Table 2. Correlation Matrix for Excess Returns and Factors 1966:10–1991:2.

	2-mo	4-mo	6-mo	8-mo	10-mo	12-mo	FUNDS	DFLT	TERM
2-mo	1.00								
4-mo	0.86	1.00							
6-mo	0.82	0.97	1.00						
8-mo	0.79	0.94	0.99	1.00					
10-mo	0.76	0.92	0.97	0.99	1.00				
12-mo	0.73	0.89	0.95	0.98	1.00	1.00			
FUNDS	–0.32	–0.43	–0.42	–0.41	–0.41	–0.40	1.00		
DFLT	–0.10	–0.10	–0.11	–0.11	–0.11	–0.11	–0.11	1.00	
TERM	0.45	0.60	0.68	0.73	0.74	0.76	–0.17	–0.43	1.00

Table 3. Correlation Matrix for Squared Excess Returns and Factors 1966:10–1991:2.

	2-mo	4-mo	6-mo	8-mo	10-mo	12-mo	FUNDS	DFLT	TERM
2-mo	1.00								
4-mo	0.92	1.00							
6-mo	0.90	0.95	1.00						
8-mo	0.86	0.91	0.99	1.00					
10-mo	0.82	0.88	0.96	0.99	1.00				
12-mo	0.79	0.84	0.93	0.97	0.99	1.00			
FUNDS	0.21	0.31	0.27	0.26	0.27	0.26	1.00		
DFLT	0.01	0.04	0.04	0.04	0.04	0.04	0.01	1.00	
TERM	0.56	0.57	0.67	0.71	0.71	0.71	0.09	0.24	1.00

Summary statistics for the excess return series and the three factors are reported in Table 4. The unconditional mean of each return series is significantly different from 0, and there is significant serial correlation in both levels. The unconditional distributions exhibit kurtosis in excess of Gaussian, which is consistent with univariate ARCH. The Ljung-Box statistic for 12th order serial correlation in each squared return series indicates significant ARCH effects with relatively long memory in the conditional variance, which suggests that a low-order GARCH model may be a parsimonious description of the return data.

Table 4. Summary Statistics for Excess Returns 1966:10–1991:02.

Return data are monthly percentage excess returns for T-bills. ** indicates significance the 0.01 level. The test statistics are for the test H_0:{test stat = 0}, and the measure of kurtosis is normalized so that a Gaussian distribution would have a kurtosis of 0.

T-bills	Mean	Min	Max	Std Error	Skew	Kurtosis	Q(12)	Mean of Squared	Q(12) of Squared
2-mo	0.04**	–0.12	0.46	0.06	1.90**	8.92**	33.28**	*0.01***	*45.45***
4-mo	0.07**	–0.65	1.26	0.18	1.57**	8.86**	32.39**	*0.04***	*78.10***
6-mo	0.09**	–1.02	2.19	0.30	1.67**	11.01**	36.64**	*0.10***	*63.87***
8-mo	0.09**	–1.53	3.04	0.43	1.34**	10.40**	37.12**	*0.19***	*71.60***
10-mo	0.09**	–2.25	3.58	0.54	0.91**	8.57**	34.04**	*0.30***	*86.44***
12-mo	0.10**	–2.91	4.06	0.64	0.63**	7.67**	31.74**	*0.41***	*89.71***
FUNDS	0.01	–5.73	2.95	0.67	–1.59**	21.56**	68.05**	*0.45***	*91.01***
DFLT	0.04	–5.14	4.56	1.35	–0.20	1.96**	46.54**	*1.83***	*94.16***
TERM	0.05	–9.26	13.97	3.22	0.53**	1.86**	21.91**	*10.32***	*55.57***

5. RESULTS

As described above, the factors are generated by fitting a ARMA GARCH(p, q) model to the appropriate raw factor series. The lag order for the conditional mean of the *DFLT* and *TERM* factors is selected by the Akaike Information Criterion. The conditional mean for the FUNDS factor is specified as described previously. Since the conditional mean parameters of the factors are not of particular interest, we do not report them here.

The lag order in the GARCH conditional variance was selected by sequential likelihood ratio and Lagrange multiplier tests, although information criteria would select similar lag orders. For the funds rate factor over this sample, this procedure suggests that a VAR with a GARCH(1,1) model for the funds rate provides the best fit. For the *DFLT* and *TERM* factors, this procedure suggests that a GARCH(1,1) model also provides the best fit. The estimates for the variance function parameters are reported in Table 5. There is evidently considerably long memory in the conditional variance for each of the GARCH models, as the coefficient estimates sum to nearly 1.

Engle, Ng and Rothschild (1990) found some support for the hypothesis that the time-varying risk premium at the short end of the term structure may be driven by a single common factor, so we first estimate the empirical asset pricing model with each factor individually. These results are reported in Table 6.

Table 5. Variance Function Parameter Estimates for Factors 1966:10–1991:2.

These are maximum likelihood coefficient estimates from the GARCH variance function $\sigma_{k,t} = \alpha_0 + \sum_{j=1}^{p} \beta_j \sigma^2_{k,t-j} + \sum_{i=1}^{q} \alpha_i f^2_{k,t-1}$ for each of the macroeconomic factors with asymptotic t-statistics in parentheses. The reported LM-tests are p-values for LM-tests of: omitted GARCH-M; omitted GARCH($p+1$, q) and GARCH(p, $q+1$). Conditional mean autoregressive parameters are not reported. The order of autoregression was selected by the Schwartz criterion.

Factor		Factor 1 FUNDS GARCH(1,1)			Factor 2 DFLT GARCH (1,1)			Factor 3 TERM GARCH (1,1)	
$\hat{\alpha}_{i,0}$		0.020			0.075			0.424	
		(2.36)			(1.25)			(1.35)	
$\hat{\alpha}_{i,1}$		0.662			0.240			0.127	
		(2.31)			(4.16)			(2.61)	
$\hat{\beta}_{i,1}$		0.316			0.742			0.830	
		(1.17)			(11.47)			(14.57)	
LM	G-M	G-(2,1)	G-(1,2)	G-M	G-(2,1)	G-(1,2)	G-M	G-(2,1)	G-(1,2)
Test	0.87	0.26	0.23	0.63	0.40	0.27	0.09	0.44	0.42

Consider first the response of the T-bills to innovations in the factors, given by the estimated coefficients $\hat{b}_{i,k}$. The response of T-bill returns to innovations in *FUNDS* and *TERM*, i.e. $\hat{b}_{i,1}$ and $\hat{b}_{i,3}$, are highly significant, according to the both "asymptotic" t-statistics and the Murphey-Topel t-statistics. For the *TERM* factor, the coefficients are all positive and less than one, with the magnitude increasing for longer maturities. For the *FUNDS* factor, the coefficients are all negative and less than one, also with the magnitude increasing for longer maturities. That is, a positive innovation to *FUNDS*, which is measured as the average of daily observations, tends to be followed by falling end-of-month T-bill prices, which implies a negative holding period return. Since a decrease in the price of a pure-discount bond implies an increase its yield, this finding is consistent with the expectations hypothesis dominating at the short end of the term structure, and coincides, for example, with the findings of Cook and Hahn (1979). The coefficients on *TERM* and *FUNDS* are also economically meaningful. A one standard error shock to each factor implies substantial movements in T-bills returns relative to their mean. For example, a one-standard error shock to *TERM*, tends to increase the return on 6-month T-bills by $3.22 * 0.064 = 0.21$ compared to mean monthly return of only 0.06. A one

Table 6. Coefficients for Single-Factor Models 1966:10–1991:2.

These are maximum likelihood coefficient estimates for the single-factor model $r_{i,t} - r_{0,t} = c_i + \xi_{i,k}\hat{\sigma}^2_{k,t} + b_{i,k}\hat{f}_{k,t} + \varepsilon_{i,t}$, for each of the macroeconomic factors. "Asymptotic" t-statistics and Murphey-Topel (1985) t-statistics are in parentheses. $\bar{R}^2$ is the centered adjusted R-squared.

	Factor 1 FUNDS				Factor 2 DFLT				Factor 3 TERM			
	$\bar{R}^2$	$\hat{c}_i$	$\hat{\xi}_{t,1}$	$\hat{b}_{i,1}$	$\bar{R}^2$	$\hat{c}_i$	$\hat{\xi}_{i,2}$	$\hat{b}_{i,2}$	$\bar{R}^2$	$\hat{c}_i$	$\hat{\xi}_{i,3}$	$\hat{b}_{i,3}$
TB2	0.16	0.034	0.011	–0.035	0.02	0.033	0.003	–0.006	0.23	0.019	0.002	0.009
		(9.62)	(4.91)	(–6.69)		(6.04)	(1.27)	(–2.04)		(3.18)	(3.74)	(8.63)
			(5.16)	(–7.41)			(0.91)	(–1.93)			(3.53)	(13.01)
TB4	0.20	0.063	0.016	–0.123	0.01	0.054	0.008	–0.014	0.36	0.040	0.003	0.034
		(6.34)	(2.69)	(–8.52)		(3.41)	(1.25)	(–1.71)		(2.58)	(2.00)	(12.64)
			(2.83)	(–10.68)			(0.84)	(–1.77)			(1.91)	(18.11)
TB6	0.19	0.076	0.030	–0.200	0.01	0.062	0.013	–0.027	0.46	0.053	0.003	0.064
		(4.64)	(3.04)	(–8.36)		(2.41)	(1.25)	(–2.03)		(2.24)	(1.56)	(15.84)
			(3.05)	(–10.45)			(0.81)	(–2.05)			(1.49)	(23.86)
TB8	0.19	0.074	0.041	–0.277	0.01	0.058	0.016	–0.039	0.52	0.051	0.003	0.095
		(3.22)	(2.90)	(–8.16)		(1.58)	(1.13)	(–2.07)		(1.62)	(1.29)	(17.77)
			(2.76)	(–10.28)			(0.73)	(–2.12)			(1.28)	(26.63)
TB10	0.18	0.075	0.049	–0.345	0.01	0.052	0.021	–0.046	0.55	0.044	0.004	0.123
		(2.55)	(2.72)	(–8.02)		(1.12)	(1.15)	(–1.95)		(1.13)	(1.33)	(18.74)
			(2.43)	(–10.07)			(0.75)	(–1.99)			(1.35)	(28.23)
TB12	0.17	0.079	0.056	–0.399	0.01	0.049	0.026	–0.054	0.57	0.037	0.01	0.148
		(2.24)	(2.65)	(–7.76)		(0.89)	(1.18)	(–1.91)		(0.81)	(1.48)	(19.49)
			(2.26)	(–9.67)			(0.79)	(–1.93)			(1.47)	(29.29)

standard error shock to *FUNDS* decreases the return on 6-month T-bills by 0.67*(–0.200) = –0.092,also greater (in absolute value) than its mean monthly return.

Innovations in *DFLT* apparently, however, have little effect on T-bill returns. The coefficients on innovations in *DFLT*, $\hat{b}_{i,2}$, are only marginally significant at the usual levels, and the adjusted R-squared statistics are extremely low – less than 2%. This suggests, somewhat surprisingly given the findings of Fama and French (1993), that the *DFLT* factor is not particularly useful for explaining the temporal variation in the short-end of the term structure.

With regard to the type of risk that is priced, the point estimates for the coefficients on the conditional volatilities of the factors, $\hat{\xi}_{i,k}$, imply that the volatility of *DFLT* is not priced, although there is evidence that the volatility of *FUNDS* is priced, and some evidence that *TERM* is priced at the very short end of the term structure. For example, *TERM* volatility appears to be priced in 2- and 4-month T-bills, while the volatility of *FUNDS* is priced significantly in each of the assets. Both of these factors explain a substantial portion of T-bill returns, with the adjusted R-squared as high as 20% in the *FUNDS* equations and as high as 57% the *TERM* equations.

Since both *FUNDS* and *TERM* have substantial explanatory power for the asset returns, we re-estimate the asset pricing equation with both factors. These results are reported in Table 7.[6] In the two-factor model, innovations in *TERM* and *FUNDS* are still highly significant. The coefficients are somewhat smaller in magnitude, but still economically significant. The volatility of *TERM* and the volatility of *FUNDS* retain their statistical significance in explaining the time-varying risk premia of T-bills, and together the two factors explain a remarkable percentage of the variation in the T-bills – up to 65%.

Discussion

These empirical results provide evidence that the volatilities of common bond factors are not dynamically priced at the short end of the term structure, but that *FUNDS* volatility may be macroeconomic factor whose risk is dynamically priced over the post-1966 sample. The results also provide some evidence that both *TERM* and *FUNDS* are empirically informative for T-bill returns. The estimated price of *FUNDS* risk is unambiguously positive, which would be expected if investors are risk averse. (In contrast, Thorbecke (1997) estimates static risk premium for *FUNDS* that is negative, and several authors, including Campbell (1987) have estimated market risk premiums that are negative.) The intuition for *FUNDS* volatility being priced is straightforward. *FUNDS* volatility is the conditional variance of the one-month ahead forecast error in

Table 7. Coefficients for Multi-Factor Model 1966:10–1991:2.

These are maximum likelihood coefficient estimates for the two-factor model $r_{i,t} - r_{0,t} = c_i + \sum_{k=1}^{2} \xi_{i,k}\hat{\sigma}^2_{k,t} + \sum_{k=1}^{2} b_{i,k}\hat{f}_{k,t} + \varepsilon_{i,t}$. "Asymptotic t-statistics" are in parentheses. $\bar{R}^2$ is the centered adjusted R-squared.

			Factor 1 FUNDS		Factor 3 TERM	
TB2	0.33	0.024	0.008	–0.028	0.001	0.008
		(4.11)	(3.42)	(–5.73)	(2.03)	(7.95)
TB4	0.48	0.047	0.008	–0.097	0.001	0.030
		(3.19)	(1.56)	(–8.08)	(1.31)	(12.31)
TB6	0.57	0.068	0.020	–0.151	0.001	0.058
		(3.07)	(2.50)	(–8.43)	(0.49)	(15.86)
TB8	0.62	0.071	0.028	–0.204	0.001	0.088
		(2.41)	(2.52)	(–8.52)	(0.21)	(18.00)
TB10	0.64	0.066	0.030	–0.250	0.001	0.115
		(1.82)	(2.26)	(–8.49)	(0.36)	(19.04)
TB12	0.65	0.061	0.033	–0.282	0.002	0.139
		(1.43)	(2.11)	(–8.22)	(0.56)	(19.77)

the funds rate. It can therefore be interpreted as a measure of uncertainty regarding next month's realization of the funds rate. When investors are unsure about the next month's realization, i.e. when the conditional volatility of *FUNDS* is high, they may require additional compensation to hold T-bills.

The factors and the conditional variance of the factors are plotted in Fig. 3. Visual inspection suggests that the temporal variation in *FUNDS* and *TERM* tends to coincide with the T-bill returns plotted in Fig. 1. Inspection also suggests that the conditional volatilities of *FUNDS*, and possibly *TERM*, tend to coincide with squared excess T-bill returns plotted in Fig. 2. In each case, the volatilities tend to be greatest during the 1979:10–1982:10 time period, with additional increases in volatility around the 1970 and 1974 recessions.

This empirical result that *FUNDS* volatility may be priced is somewhat surprising given the results of Elder (2001), which did not find evidence that *FUNDS* volatility was priced, although in a single factor framework over an alternative sample period. That is, Elder (2001) utilized the data sample initially studied by Engle, Ng and Rothschild (1990). The sample studied in this paper, however, begins in 1966, which coincides with the period when the federal funds market first began to function as a major source of bank liquidity,

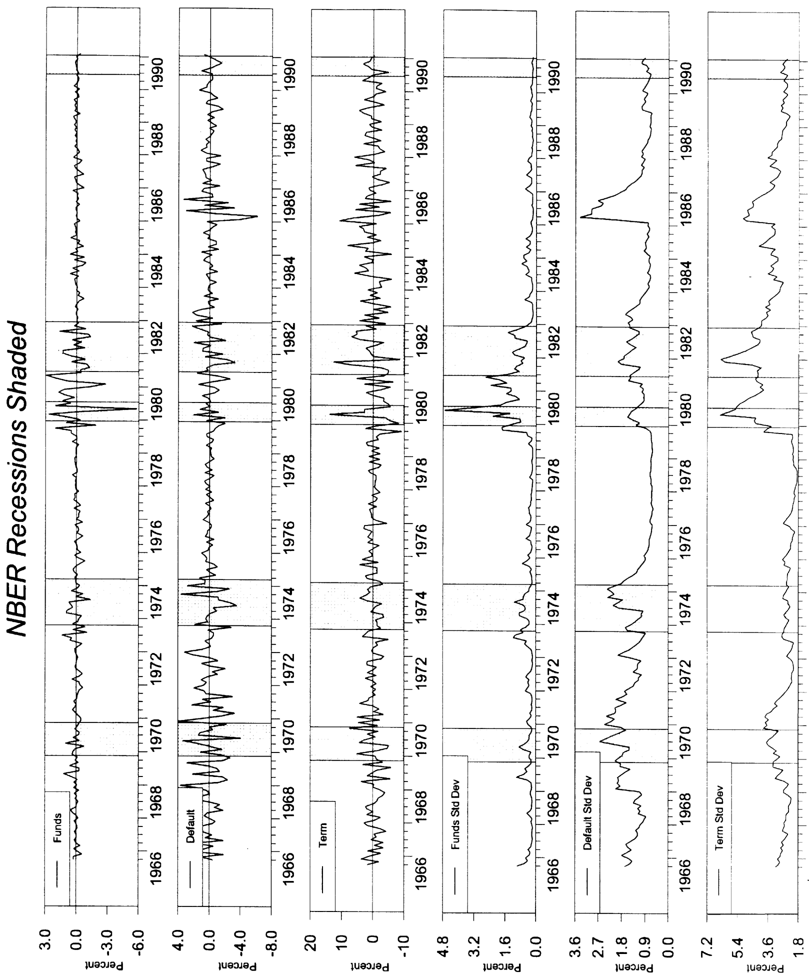


Fig. 3. Macroeconomic Factors.

and the structural features associated with this sample period are evidently important in estimating the factor innovations and factor volatilities. For example, if we reestimate the asset pricing equations over alternative sub-samples such as the post 1982:10 period, conditioning on the factor innovations and volatilities estimated utilizing the full sample information, the results are similar to those reported here. If we instead reestimate the factor innovations and volatilities over various sub-samples, their explanatory power for T-bill returns in the second stage is altered. One possible interpretation is that structural features of the full post-1966 sample are important in estimating the factor innovations and volatilities.

6. CONCLUSIONS

This paper investigates the extent to which three common macroeconomic factors can explain the time-variation in the risk premia in the short-end of the term structure. Our empirical model is motivated by a modern, dynamic asset pricing model that can be viewed as a multi-factor ICAPM with conditionally heteroskedastic factors. Our dynamic factor model explains up to 65% of the temporal variation in T-bill returns, and we find that the short end of the term structure responds significantly to innovations in the federal funds rate and shifts in the yield curve, but not to innovations in the default risk premium. We also find that the volatility of shifts in the yield curve is a priced dynamic factor at the very low end of the term structure, and that the volatility of the funds rate may have been priced over the post-1966 sample, when the federal funds rate first began to serve as a source of bank liquidity. The finding that the volatility of the funds rate is priced is somewhat sensitive to the sample period chosen. These results provide some support for the hypothesis risk premiums at the short end of the term structure tend to rise during periods when these macro factors are volatile.

ACKNOWLEDGMENTS

I thank James D. Hamilton, Tim Bollerslev and seminar participants at the University of Virginia for their comments on a very early draft of this paper. Any remaining errors are my own.

NOTES

1. Early contributions include Campbell (1987) and Harvey (1989). Time-varying risk premia in the short end of the term structure was documented by Engle, Ng and Rothschild (1990).

2. Alternatively, the observed macro factors could be filtered so that they are conditionally orthogonal, although at the risk of obscuring their economic interpretation.

3. For example, our Eq. (2-5) is very similar to Eqs (8) and (11) of King, Sentana and Wadhwani (1994).

4. Similar policy measures have been utilized in asset pricing frameworks by Thorbecke (1997).

5. Alternatively, we can impose $N(N-1)/2$ exclusion restrictions on $\mathbf{B}$ satisfying a rank condition.

6. We do not recalculate Murphy-Topel t-statistics due to the computational complexities involved for multi-factor models. The results from the single factor model, however, suggest that inference is not likely to be affected.

REFERENCES

Bernanke, B. S., & Blinder, A. S (1992). The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review*, *82*, 901–921.

Bernanke, B. S., & Gertler, M. (1995). Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *Journal of Economics Perspectives*, *9*, 27–48.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews*, *11*, 143–172.

Campbell, J. Y. (2000). Asset Pricing at the Millennium. *Journal of Finance*, *55*, 1515–1567.

Campbell, J. Y. (1987). Stock Returns and the Term Structure. *Journal of Financial Economics*, *18*, 373–399.

Chen, N-f., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *Journal of Business*, *59*, 383–403.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (1996). The Effects of Monetary Policy Shocks, Evidence from the Flow of Funds. *Review of Economics and Statistics*, *78*, 16–34.

Citibase: Citibank economic database (1978). New York: Citibank, N. A.

Cochrane, J. H. (2001). *Asset Pricing*. New Jersey: Princeton University Press.

Cochrane, J. H. (1996). A Cross-Sectional Test of an Investment-Based Asset Pricing Model. *Journal of Political Economy*, *104*, 572–621.

Elder, J. (2001). Can the Federal Funds Rate Explain the Time-Varying Risk Premium in Treasury Bill Returns? *Journal of Macroeconomics*, *23*, 73–97.

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

Engle, R. F., & Kroner, K. F. (1985). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, *11*, 122–150.

Engle, R. F., Ng, V., & Rothschild, M. (1990). Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics*, *45*, 213–237.

Fama, E. F., & French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, *33*, 3–56.

Flannery, M. J., Hameed, A. S., & Harjes, R. H. (1997). Asset Pricing, Time-Varying Risk Premia and Interest Rate Risk. *Journal of Money, Credit and Banking*, *21*, 315–335.

Harvey, C. R. (1989). Time-Varying Conditional Covariances in Tests of Asset Pricing Models. *Journal of Financial Economics*, *24*, 289–317.

Harvey, C. R. (1991). The World Price of Covariance Risk. *Journal of Finance*, *44*, 111–157.

Kan, R., & Zhou, G. (1999). A Critique of the Stochastic Discount Factor Methodology. *Journal of Finance*, *54*, 1221–1248.

King, M., Sentana, E., & Wadhwani, S. (1994). Volatility and Links Between National Stock Markets. *Econometrica*, *62*, 901–903.

Madura, J., & Schnusenberg, O. (2000). Effect of Federal Reserve Policies on Bank Equity Returns. *Journal of Financial Research*, *23*, 421–447.

McCulloch, J. H., & Kwon, H-C. (1993). U.S. Term Structure Data: 1947–1991. Ohio State University Working Paper No. 93–6.

McElroy, M., & Burmeister, E. (1988). The Arbitrage Pricing Theory as a Restricted Nonlinear Multivariate Regression Model. *Journal of Business and Economic Statistics*, *6*, 1–41.

Meulendyke, A-M. (1989). U.S. Monetary Policy and Financial Markets. New York: Federal Reserve Bank of New York.

Murphy, K. M., & Topel, R. H. (1985). Estimation and Inference in Two-Step Econometric Models. *Journal of Business and Economic Statistics*, *3*, 370–379.

Pagan, A. (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review*, *25*, 221–247.

Pagan, A., & Ullah, A. (1988). The Econometric Analysis of Models with Risk Terms. *Journal of Applied Econometrics*, *3*, 87–105.

Patelis, A. D. (1997). Stock Return Predictability and the Role of Monetary Policy. *Journal of Finance*, *52*, 1951–1972.

Scruggs, J. T. (1998). Resolving the Puzzling Intertemporal Relation between theMarket Risk Premium and Conditional Market Variance: A Two-Factor Approach. *Journal of Finance*, *53*, 575–603.

Strongin, S. (1994). The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle. *Journal of Monetary Economics*, *35*, 463–497.

Thorbecke, W. (1997). On Stock Market Returns and Monetary Policy. *Journal of Finance*, *52*, 635–654.

White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica*, *50*, 1–25.

APPENDIX

Standard errors by the method of Murphy-Topel (1985) are a combination of the estimated scores from the first-stage and second-stage regressions with a measure of the sensitivity of the second stage regression to the estimates obtained in the first stage. The method described here is based on Hamilton (1997) and Murphy-Topel (1985). Let $\hat{\boldsymbol{\theta}}_1$ be an $(a_1 \times 1)$ vector of estimated parameters from the first stage regression, which is the GARCH model of the funds rate conditioned on other macro variables, and let the log likelihood from this regression at date *t* conditional on the contemporaneous information set Ψ_t

and $\hat{\boldsymbol{\theta}}_1$ be represented by $\ln f(F_t | \Psi_t : \boldsymbol{\theta}_1)$. Also, let $\boldsymbol{\ell}_{1,t}$ be the vector of estimated scores for date t:

$$\boldsymbol{\ell}_{1,t} = \left. \frac{\partial \ln f(F_t | \boldsymbol{\Psi}_t ; \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1} \right|_{\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1}$$

Let $\hat{\boldsymbol{\theta}}_2$ be an $(a_2 \times 1)$ vector of estimated parameters from the second stage regression, which is the regression of T-bill returns on the estimated factor and its conditional variance, and let the log likelihood from this regression at date t conditional on the contemporaneous information set $\boldsymbol{\Psi}_t$, $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be represented by $\ln f(Rt | \boldsymbol{\Psi}_t : \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ and let $\boldsymbol{\ell}_{2,t}$ be the vector of estimated scores for date t:

$$\boldsymbol{\ell}_{2,t} = \left. \frac{\partial \ln f(\boldsymbol{R}_t | \boldsymbol{\Psi}_t ; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2} \right|_{\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_2, \boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2 .}$$

Also, let $\boldsymbol{\ell}_{3,t}$ be the following vector, which is a measure of the sensitivity of the second-stage regression to the parameters estimated in the first stage:

$$\boldsymbol{\ell}_{3,t} = \left. \frac{\partial \ln f(\boldsymbol{R}_t | \boldsymbol{\Psi}_t ; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_1} \right|_{\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2 .}$$

Finally, define the following

$$\hat{\boldsymbol{R}}_1 = T^{-1} \sum_{t=1}^{T} \boldsymbol{\ell}_{1,t} \boldsymbol{\ell}'_{1,t}$$

$$\hat{\boldsymbol{R}}_2 = T^{-1} \sum_{t=1}^{T} \boldsymbol{\ell}_{2,t} \boldsymbol{\ell}'_{2,t}$$

$$\hat{\boldsymbol{R}}_3 = T^{-1} \sum_{t=1}^{T} \boldsymbol{\ell}_{3,t} \boldsymbol{\ell}'_{2,t}$$

$$\hat{\boldsymbol{R}}_4 = T^{-1} \sum_{t=1}^{T} \boldsymbol{\ell}_{1,t} \boldsymbol{\ell}'_{2,t}$$

and

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{R}}'_4 \hat{\boldsymbol{R}}_1^{-1} \hat{\boldsymbol{R}}_3$$

The Murphy-Topel covariance matrix for $\hat{\boldsymbol{\theta}}_2$ is given by

$$\frac{1}{T}\hat{\Sigma} = \frac{1}{T}[\hat{\boldsymbol{R}}_2^{-1} + \hat{\boldsymbol{R}}_2^{-1}(\hat{\boldsymbol{R}}_3'\hat{\boldsymbol{R}}_1^{-1}\hat{\boldsymbol{R}}_3 - \hat{\boldsymbol{H}} - \hat{\boldsymbol{H}}')\hat{\boldsymbol{R}}_2^{-1}].$$

Note that if $\boldsymbol{\ell}_{3,t}$ is a vector of zeros, the Murphy-Topel covariance matrix becomes simply for $\hat{\boldsymbol{\theta}}_2$

$$\frac{1}{T}\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\hat{R}_2^{-1}.$$

That is, if the parameter estimates in the second stage regression are not sensitive to the parameters estimated in the first stage, then the Murphy-Topel covariance matrix for the estimated parameter vectors from the second stage $\hat{\boldsymbol{\theta}}_2$ reduces to the familiar covariance matrix based simply on the outer-product estimate of the information matrix.

STOCK SPLITS AND LIQUIDITY: EVIDENCE FROM AMERICAN DEPOSITORY RECEIPTS

Christine X. Jiang and Jang-Chul Kim

ABSTRACT

Using a sample of stock splits on NYSE listed ADRs between 1994 and 1999, we study the change in liquidity following stock splits. Our findings suggest that cost to liquidity demanders measured by percentage quoted and effective bid-ask spreads, split-factor adjusted quoted depth and trading volume increases for split-up securities. However, we observe that raw trading volume and depth both go up after splits, suggesting that liquidity may increase because market makers/brokers' higher incentives in promoting the shares for larger payments on order flows. In addition, number of small trades and number of shareholders go up 28% and 21%, respectively while institutional holdings pre- and post-splits are not significantly different, also consistent with the notion that splits provide an incentive for brokers to promote the stocks, and their efforts seem to target small investors.

1. INTRODUCTION

There has been considerable research on the return behavior of common stocks surrounding stock splits. The collective evidence shows that stock prices respond positively to the announcement of stock splits and the researchers have

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 109–131.

ISBN: 0-7623-0887-7

attributed the gain mainly to the signaling role of stock split under asymmetric information, and to the improved liquidity following the split. More recently, the focus has been on the trading behavior of the split-up securities and whether the observed trading patterns are consistent with the often quoted rationale for stock splits, that is firms split their stocks to improve liquidity and hence as a result increases shareholder wealth.

Corporate managers maintain that stock splits return stock prices closer to the "optimal range" to make them more liquid. Various models and rationales have been put forward to account for the increase in liquidity and equity value associated with stock splits. The increase in value could be due to a reduction in trading costs in the tradition of Amihud and Mendelson (1986, 1987, 1988). Along this line, it is expected that measures of trading costs such as bid-ask spread and commission costs should decrease after splits. And moreover, some key determinants of bid ask spread may also change accordingly to explain the reduction in bid-ask spread. Volume should increase so the inventory component of bid-ask spread goes down and volatility likely to decrease to reflect a reduction in the adverse selection component of the transaction cost. However, empirical evidence on stock splits suggests the contrary, that is, some measures of liquidity are lower after a split. Copeland (1979) finds that trading volume declines in the year following a split and Conroy, Harris and Benet (1990) observe that bid-ask spread increases after splits.

An increase in spreads and a decrease in volume would contradict the "market folklore" that splits improve liquidity if the price and quantity dimensions of trading are the only measures liquidity. Nevertheless, there is more to liquidity.

Merton (1987) posits that investor awareness is valuable in capital market and increasing the fraction of investors who "know about" a security reduces its risk and required rate of return. To be consistent with Merton's theory, the number of shareholders is expected to increase post split. Lamoureux and Poon (1987), Brennan and Hughes (1991), Maloney and Mulherin (1992) all find that the number of shareholders increases after splits. The limitation of this conjecture is that it fails to explain why well known firms need to rely on stock splits to broaden their shareholder base; the argument may be more appropriate in explaining the splits of smaller, less well known firms, or perhaps foreign firms listed in the U.S.

More recently, Angel (1997) maintains that firms lower share prices through splits so as to revert to optimal relative tick size. If share price goes down while the institutionally mandated tick size remains the same, this results in an increase in the relative tick size. A larger relative tick size is likely to decrease the bargaining and processing costs and providing more incentives for limit orders and market makers to provide liquidity, though bid-ask spread may likely be higher.[1] Given that an optimal tick size is a trade-off between incentives that a larger relative tick size brings to liquidity providers and the higher transaction costs that a larger tick imposes, firms may consider increase the relative tick size of their shares through stock splits to entice liquidity providers. Brennan and Hughes (1991) also posit that since brokerage commissions are likely to increase following stock splits, brokerage firms have more incentives to promote a stock. Schultz (2000) finds that number of small trades increases while trading cost becomes higher after splits. He concludes that small traders are not trading the split-up stocks because of lower costs; rather the surge in small trades is related to greater efforts of brokers in promoting and sponsoring the stocks.

Using a unique split data set compiled from American Depository Receipts (ADR), Muscarella and Vetsuypens (1996) are able to test the competing hypotheses (signaling versus liquidity) for stock splits by examining solo ADR splits. Since solo splits refers to the cases that ADRs split when the home country stocks do not, they effectively rule out the possibility that ADRs split for the purpose of signaling. They interpret their findings as supportive of the liquidity explanation of stocks split announcement effects.

Built upon the work of Muscarella and Vetsuypens (1996), this paper aims to provide further empirical evidence on the change in liquidity around an ADR stock split and makes an effort in exploring the causes of the observed change.[2] Our study differs from prior studies in several important ways. First, while Muscarella and Vetsuypens focus on documenting the positive market response to ADR stocks splits and distinguishing between signaling and liquidity explanations for stock splits, their findings on liquidity is based on a sample of 7 ADRs and a small subset of liquidity measures. We aim to offer a more comprehensive investigation of various components of liquidity on a larger sample of ADRs. For example, we look at depth in addition to liquidity premium studied in Muscarella and Vetsuypens (1996). We consider this important because the spread captures only one dimension of liquidity. As shown in Lee, Mucklow and Ready (1993), Harris (1994), Kavajecz (1999), and Goldstein and Kavajecz (2000), it is important that we consider both the price and quantity dimensions of quotes to accurately measure liquidity. We also look for evidence that ADR splits are intended to provide incentives to liquidity providers by moving its share prices closer to the "optimal tick size." Along this line, we examine various proxies such as raw trading volume, number of small trades, small buys and small sell orders, number of shareholders and institutional holdings following splits. Overall, a careful examination of the empirical patterns of various liquidity measures enables us to better detect the source of liquidity improvement. Second, prior studies on ADR have primarily focused either on the asset pricing aspect or on the diversification benefits of ADRs, the microstructure of ADRs has received little attention. Howe and Lin (1992) examine the determinants of ADR bid-ask spread and find that the transaction costs are lower on ADR transactions. Recently, using proprietary data on NYSE specialist trading, Bacidore and Sofianos (2001) investigate how the differences between non-U.S. and U.S. stocks affect specialist participation and the market quality of non-U.S. stocks. They conclude that non-U.S. stocks have wider spreads and less depth due to higher degree of information asymmetry. In this paper, the liquidity change associated with stock splits is further examined through an investigation of the intraday patterns of ADR bid-ask spread, depth and volume. We compare the patterns observed for our sample of ADRs with some stylized observations on exchange-listed stocks (e.g. a U shape for intraday spread and volume and a reverse U-shape for intraday depth). Given the substantial growth in listing and trading volume of ADRs, this should add to our understanding of the microstructure issues of ADRs.

Using transactions data on ADRs along with other pertinent information on share ownership, we find the following. First, we investigate the impact of stock splits on the cost to liquidity demanders. We find that the price dimension of liquidity, namely the bid-ask spreads, increases substantially post split. The quantity dimension of liquidity measured by volume and depth is also carefully examined. The volume and depth adjusted for split factors indicate that the dollar value of volume and depth deteriorate while the raw volume and depth increase post split. Demsetz (1977), Copeland (1979), Baker and Gallagher (1980), Lamoureux and Poon (1987) also document the negative relationship between stock split and split adjusted liquidity measures. Second, with regard to broadening shareholder base and brokers' efforts in promoting the split-up stocks, our results show that the number of shareholders and in addition the number of small trades significantly increase after ADR splits, consistent with findings of Lamoureux and Poon (1987), and Brennan and Hughes (1991). Our results also show that small buys increase significantly post splits which is similar to Schultz (2000). Percentage of institutional ownership prior and post-split is also compared and we find that institutional ownership decrease post split although the difference is not statistically significant. Third, we investigate intraday patterns of percentage bid-ask spread, depths and volume. Consistent with other studies, spreads and volume across a trading day are found to follow a U-shape. However, the depth pattern exhibits a reverse S-shape which is inconsistent with other studies and the theory of inventory control model which suggests that specialists use quotes with wider spread and low depth to avoid carrying shares overnight. In addition, we compare the absolute spreads before and post splits. Absolute spread is found to decrease significantly following splits which is comparable to other findings that support a positive relationship between share price and absolute spread.

The remainder of the paper is organized as follows. In Section 2, we discuss empirical constructs used to measure components of liquidity. Section 3 presents data and empirical results. Section 4 summarizes our findings and concludes the paper.

2. MEASURES OF LIQUIDITY

A traditional measure of market liquidity has been the quoted bid-ask spread (Huang & Stoll, 1996). Of interest to us is whether stock splits lead to lower quoted spread in those stocks. We compute the quoted percentage spread as the difference between the prevailing quoted bid and ask prices relative to the quote midpoint.

Trading on the NYSE provides investors the opportunity for price improvement. Lee and Ready (1991) present evidence that many trades take place inside the quoted bid-ask spread, thus it is suggested that the effective spread can be a better measure of transaction cost. The effective spread is defined as twice the absolute difference between the trade price and the quote midpoint existing at time of trade.

The quoted and effective spreads measure only the price dimension of the liquidity. As shown in Lee, Mucklow and Ready (1993), Harris (1994), Kavajecz (1999), and Goldstein and Kavajecz (2000), it is important that we consider both the price and quantity dimensions of dealer quotes to provide a complete characterization of market liquidity. When liquidity is defined along these two important dimensions, the change in liquidity can occur through change in one of these two measures or both.

Therefore, we also compile information on quoted depth and volume to examine the change in the quantity dimension of the liquidity following splits. We have the following three measures of depth: quoted ask depth; quoted bid depth; and quoted depth which is the sum of the depth at ask price and the depth at bid price. In addition to raw volume and raw depth measures, we also use split-factor adjusted ones. Since the volume and/or depth are likely to increase given that the number of shares outstanding increases n fold for n for 1 stock splits, liquidity are considered to have a real improvement only if they exceed the measures before split at least n times (for n for 1 stock splits). We follow the approach of Copeland (1979) to adjust the depth and volume measures by the split factors. Specifically, we divide depths and volume following splits by the ADRs split factors to make them comparable with pre-split depths and volume. The adjustment process allows us to see more clearly the whether the dollar value of shares traded and the dollar value of depth have changed following splits. We standardize the split factor adjusted measures of spread, depth, and volume so they are not unduly influenced by extreme values.

3. DATA AND EMPIRICAL RESULTS

Data Source and Sample Selection

To identify ADR splits for this research, we use the CRSP NYSE File. From the CRSP Data File, 62 ADR splits were initially identified between 1994 and 1999.[3] We include all ADR splits even though a firm has more than one split in the sample period, provided that the splits were sufficiently far apart. Since the NYSE's Trade and Quote data over 1994–1999 are used, we further eliminate splits that do not have enough transaction data pre- and post-splits (for example, splits in January 1994 and in December 1999). The final sample includes 51 splits on 44 NYSE listed ADRs originated from 22 different countries.

Once all ADR splits were identified from CRSP and filtered, quote and trade data between 1994 and 1999 were retrieved from the NYSE's Trade and Quote (TAQ) database. To minimize sampling errors and secure a well-behaved sample, we omit trades and quotes if the TAQ database indicates that they are out of time sequence or involve an error. We omit quotes if either the ask or bid price is equal to or less than zero; either the bid or ask depth is equal to or less than zero; and either the price or volume is equal to or less than zero. In addition, following Huang and Stoll (1996), we further minimize data errors by eliminating quotes with characteristics described below:

(1) quotes if the bid-ask spread is greater than \$4 or negative.
(2) quotes associated with trading halts or designated order imbalance.
(3) before-the-open and after-the-close trades and quotes.
(4) trade price, p_t, if $|(p_t - p_{t-1})/p_{t-1}| > 0.10$.
(5) ask quote, a_t, if $|(a_t - a_{t-1})/a_{t-1}| > 0.10$.
(6) bid quote, b_t, if $|(b_t - b_{t-1})/b_{t-1}| > 0.10$.

Panel A in Table 1 presents a description of the sample ADR splits by country, ticker symbols, splits dates and the prices pre- and post- splits and the split factors. For each stock, pre- and post-split prices are calculated through

Table 1. The Characteristics of ADRs.

Panel A. ADR Stock Split Sample and Average Prices Before and After Split

Our sample includes 51 stock splits on 44 ADRs originated from 22 countries. Prices before and after split are the average of the last inside bid and ask quotes prior to a split and the average of the first inside bid and ask quotes after a split. All prices are quoted in U.S. dollar.

Country	Ticker	Price before split	Price after split	Split factor	Split date
Argentina	BFR	29.6875	25.3750	23:20	12/11/96
Argentina	TAR	36.8125	18.3125	2:1	02/28/95
Argentina	TEO	57.9375	29.0000	2:1	08/28/97
Australia	BHP	57.5000	28.5000	2:1	06/25/96
Australia	NWS	46.3125	16.0000	3:1	11/21/94
Brazil	ARA	13.6250	9.3750	3:2	05/19/94
Brazil	ARA	12.0000	9.1250	4:3	04/06/95
Chile	CTC	101.0625	23.6875	5:1	01/02/97
Denmark	NVO	101.6875	25.7500	4:1	04/18/94
Denmark	TLD	54.1250	27.5938	2:1	06/08/99
Finland	MX	23.6250	12.0313	2:1	07/02/99
Germany	PV	76.0938	53.5000	3:2	07/21/98
Hong Kong	CBA	19.0625	12.6875	3:2	07/13/99
Hong Kong	CBA	25.8125	5.5938	5:1	10/04/99
Hong Kong	HKT	57.8125	19.3750	3:1	06/14/94
Indonesia	TLK	13.4063	12.2500	108:100	06/30/99
Ireland	AIB	92.0000	30.0313	3:1	05/17/99
Ireland	ELN	61.2500	30.4375	2:1	08/23/96
Ireland	ELN	50.5313	25.4375	2:1	06/07/99
Italy	LUX	92.9688	18.9063	5:1	04/17/98
Italy	NTZ	46.5000	23.2500	2:1	12/24/96
Japan	KYO	91.0938	45.0625	2:1	09/14/98
Luxembourg	ESF	29.2500	14.6875	2:1	08/03/94
Mexico	EKT	31.7500	16.1250	2:1	01/13/98
Mexico	KOF	55.6250	18.8438	3:1	01/28/98
Mexico	VTO	11.4375	9.4375	6:5	05/19/95
Netherlands	AEG	85.2500	34.1250	5:2	06/09/95
Netherlands	AEG	150.2500	77.5000	2:1	05/26/98

Table 1. Continued.

Panel A. ADR Stock Split Sample and Average Prices Before and After Split – Continued

Country	Ticker	Price before split	Price after split	Split factor	Split date
Netherlands	AHO	89.5000	29.0938	3:1	07/30/97
Netherlands	RD	213.2813	53.6875	4:1	06/30/97
Netherlands	UN	218.3438	55.4375	4:1	10/21/97
New Zealand	NZT	72.7500	36.3125	2:1	04/09/97
Norway	PGO	61.5938	30.8750	2:1	06/25/98
Philippine	PHI	65.4375	32.7500	2:1	07/14/97
South Africa	SPP	91.7500	46.1875	2:1	05/15/95
South Africa	SPP	90.5000	9.1250	10:1	10/27/99
Spain	ELE	84.3125	20.7500	4:1	08/01/97
Spain	BBV	82.2500	27.6875	3:1	07/22/97
Spain	BBV	57.5938	19.7500	3:1	07/14/98
Spain	REP	49.4688	15.8125	3:1	04/20/99
Spain	STD	51.0000	25.7500	2:1	07/06/98
Spain	TEF	149.0313	49.5313	3:1	07/27/99
Taiwan	TSM	35.2188	30.2500	123:100	08/16/99
U.K.	BAS	31.3750	15.9375	2:1	02/09/98
U.K.	AVZ	111.5000	54.8750	2:1	04/27/98
U.K.	CSG	57.2500	27.6250	2:1	06/01/99
U.K.	SBH	85.5938	43.4688	2:1	08/29/97
U.K.	SC	125.8438	42.5938	3:1	07/01/97
U.K.	UL	125.0000	31.3750	4:1	10/21/97
U.K.	VOD	87.7500	29.3125	3:1	07/25/94
U.K.	VOD	232.7500	45.9375	5:1	10/04/99

Panel B. Distribution of Share Prices Immediately Before and After Splits

	Before ($)	After ($)
Minimum	11.4375	5.5938
5%	13.4063	9.1250
25%	36.8125	16.1250
Median	61.2500	27.5938
75%	91.7500	34.1250
95%	213.2813	54.8750
Maximum	232.7500	77.5000
Mean	74.3836	28.3554

Table 1. Continued.

Panel C. Distribution of Split Factors

Split Factor	2-1	3-1	3–2	4-1	5-1	Others	Total
No. of Split Cases	21	11	3	5	3	8	51
% of Total Splits	41%	22%	6%	10%	6%	15%	100%

averaging the last inside bid and ask quotes right before split date and through taking the mean of the first inside bid and ask quotes for the first trading day post split. Observe that ADRs from the U.K. and Spain had more splits over our sample period. Interestingly, although the 22 countries in our sample have rather diverse share prices traded in their home markets and maintain different minimum tick size, they tend to split their stocks to have share prices conform to the U.S. average, which is around \$32 according to Angel (1997). ADRs originated from Brazil, Hong Kong and Mexico have relatively lower post-split prices (all are below \$20). Based on the split factors, clearly majority of the splits are two for one or three for one splits

Panel B in Table 1 lists the distribution of share prices immediately before and after splits for the entire sample. The median and mean pre-split share prices are \$61.25 and \$74.38, respectively, and the range is between \$11.44 and \$232.75. However, as expected, the median and mean post-split share prices decrease to \$27.59 and \$28.36, close to the average share price in the U.S. Angel (1997) found that for NYSE stocks, the median and mean pre-split prices are \$55.75 and \$58.95. Provided that the median U.S. stock price is about \$40, and the average NYSE share price was from \$32 to \$31 for the last half century, the post ADR split prices are rather close to the average NYSE price.

Split factors are reported in Panel C of Table 1. Of all the split factors, 2 for 1 splits account for 41% and 3 for 1 splits constitute around 22% of our sample. Clearly the distribution of ADR split factors are quite comparable to other samples focused on domestic stocks (Schultz, 2000; Gray, Smith & Whaley, 1996).

Quote and trade data such as quoted bid and ask prices, depth at bid and ask, and trading volume are retrieved for 40 trading days preceding splits and 40 trading days after splits.

To examine the potential change in investor awareness around stock splits, the data on institutional ownership and the number of shareholders were also collected from Compact Disclosure. In addition, for the number of shareholders, we supplemented our initial sample form the Compact Disclosure with information in Compustat. We identify 35 ADRs and 22 ADRs with information on institutional ownership, and the number of shareholders, respectively.[4]

Note that on June 24, 1997, NYSE changed minimum tick size from $1/8 to $1/16 for stocks traded above $1. Goldstein and Kavajecz (2000) find a decline in depths at the commencement of trading in sixteenths from eighths in 1997 while Chakraverty and Wood (2000) observe that both quoted and effective spreads are lower following decimalization. To determine changes in liquidity in our sample of ADR splits associated with the change in tick size, we form two sub-samples (before and after tick size change) to examine whether there is noticeable difference in the distribution of absolute spreads.

Intraday Pattern for Percentage Quoted and Effective Spreads, Depth, and Volume

The intraday patterns of quotes have been extensively researched for stocks listed on NYSE. Brock and Kleidon (1992), McInish and Wood (1992), Lee, Mucklow and Rudy (1993), and Chung, Van Ness and Van Ness (1999), find that, intraday patterns for spread and volume of NYSE listed stocks follow a U-shape and depths trace a reverse U-shape. Copeland and Galai (1983), and Glosten and Milgrom (1985) provide explanations based on an information asymmetry model. Assume that two types of traders present in the market, informed and liquidity traders, specialists, market makers and limit order traders are categorized as liquidity traders who possess no private information about the value of the stock. On average, they lose to informed traders. To minimize their risk and protect their profit from unfavorable trades with informed traders, liquidity traders widen the bid-ask spread in the first half-hour or early hours of trading. In general, spreads are wider during trading intervals with a high probability of information motivated trading. Amihud and Mendelson (1982) explain the wider spread before market close through an inventory model. Their model posits that specialists widen spreads to discourage trade so they do not carry unwanted inventory overnight. Lee, Mucklow and Ready (1993) document that specialists use both spread and depth as their strategic choice to limit their losses from trading against informed traders. Specifically, depth and spread have an inverse relationship in that spreads (depths) are higher (lower) in the morning, move lower (higher) during the middle of a trading day and widen (fall) just before market close.

Very little is known about the intraday patterns of ADRs, an important segment of NYSE listings.[5] We examine the patterns of percentage quoted and effective spreads, depth, and volume to see whether they conform to the extant findings for NYSE traded shares.

For a trading day divided into thirteen half-hour intervals, Figs 1, 2, 3, and 4 graph the intraday patterns for percentage quoted and effective bid-ask spread, depth, and volume before and ask split. Although split does have impact on all variables graphed, the impact is limited to shifts in levels (to be discussed in part C and D) rather than the shapes of intraday patterns. Consistent with other studies intraday patterns for spread and volume are U-shaped. Nevertheless, the distribution of depth is different from previous findings of lower depth in the first half hour of trading and near market close and higher depth during the middle of a trading day. Therefore, the inventory argument cannot explain the higher depths of ADRs in the last half-hour of trading. Howe and Lin (1992) find that ADRs are larger firms, with higher stock prices, larger number of market makers, and lower volatility than average firms listed on NYSE. Given these characteristics along with the lower bid-ask spread,

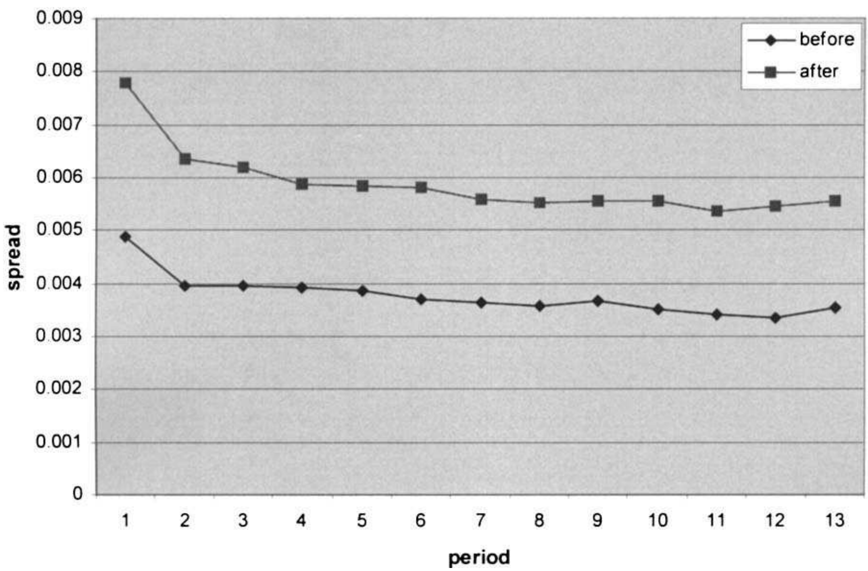


Fig. 1. Intraday Pattern for Quoted Percentage Spreads.

The NYSE is open from 9:30 A.M. to 4:00 P.M. EST and there are 13 half-hour intervals during a trading day.

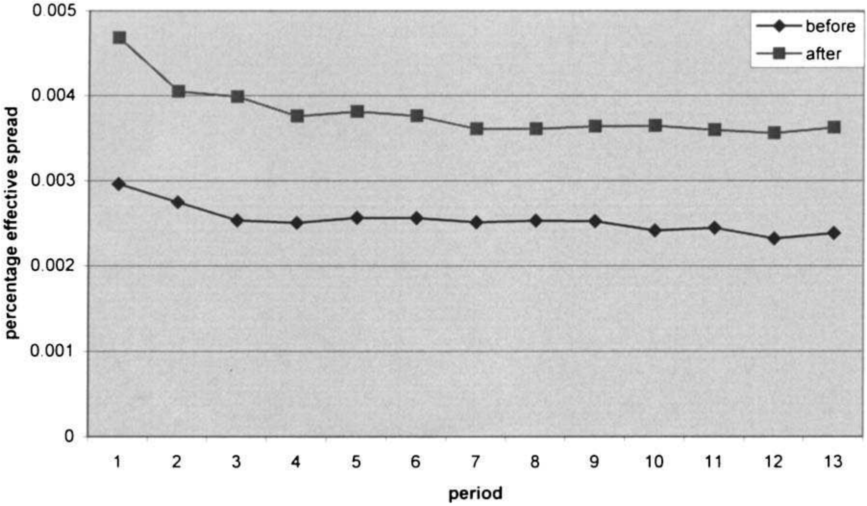


Fig. 2. Intraday Pattern for Percentage Effective Spreads.

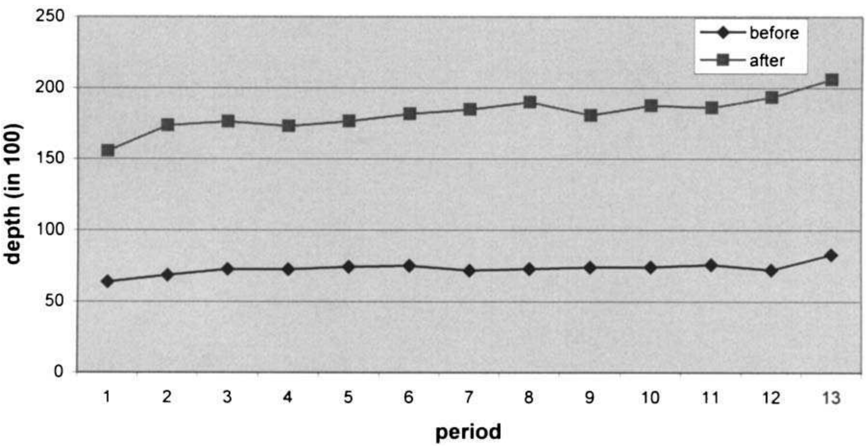


Fig. 3. Intraday Pattern for Depth.

The NYSE is open from 9:30 A.M. to 4:00 P.M. EST and there are 13 half-hour intervals during a trading day.

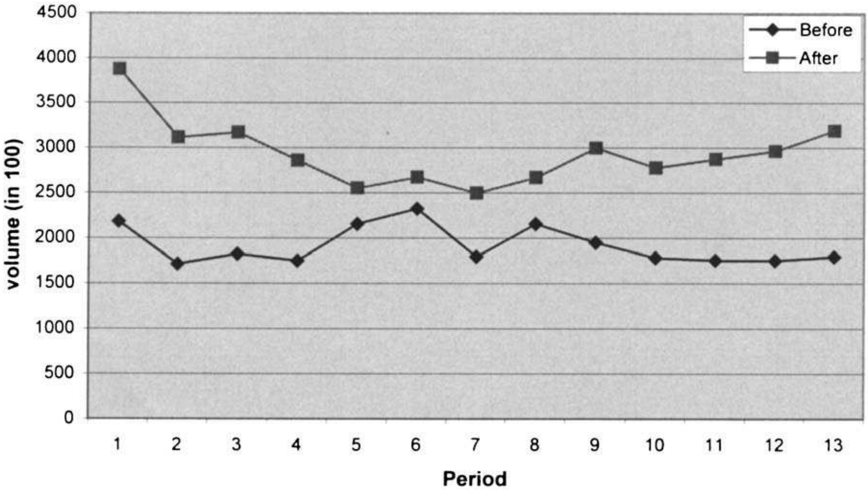


Fig. 4. Intraday Pattern for Volume.

The NYSE is open from 9:30 A.M. to 4:00 P.M. EST and there are 13 half-hour intervals during a trading day.

specialists perhaps are not as strongly averted to carrying inventory overnight as they would do for average stocks on NYSE.

Splits and Cost of Demanding Liquidity

Investors demanding liquidity from the stock market generally benefit from lower spread, higher depth and trading volume. We provide evidence on changes in these aspects of liquidity following ADR splits.

We first examine the impact of ADR splits on quoted spread. Two measures of bid-ask spread are used: the absolute dollar spread and the percentage spread. While absolute dollar spreads are intuitive, they are difficult to interpret when comparison are made across securities or for before and after splits. Percentage spreads circumvent these problems since they measure bid-ask spread per dollar. Figure 4 graphs the distribution of absolute spread before and post splits for all splits occurred before the tick size change in June 1997. Similarly, Fig. 5 presents the distribution after the tick size change.

Consistent with other studies, we find that absolute spread declines significantly following splits. Before tick size changes, 41.65% and 41.39% of the quoted absolute spread occur at 1/8 ($0.125) and 1/4 ($0.25), respectively.

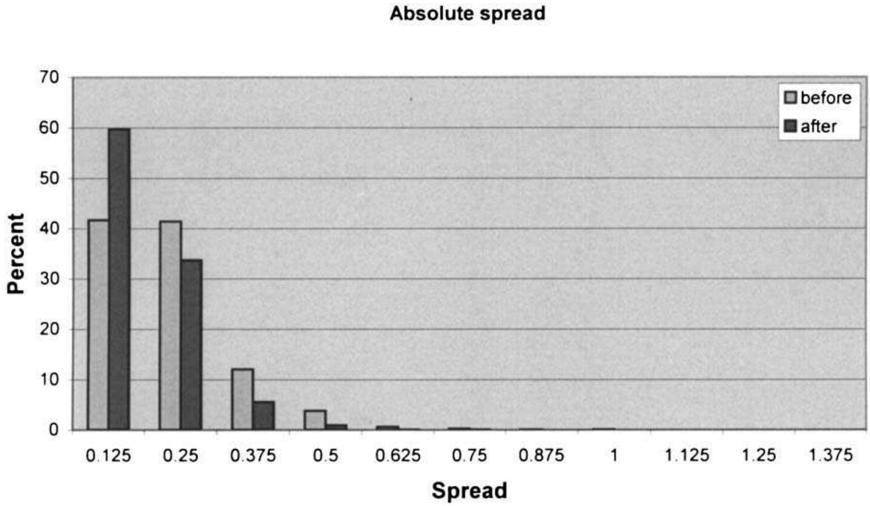


Fig. 5. Distribution of Absolute Spreads Before Tick Size Change.

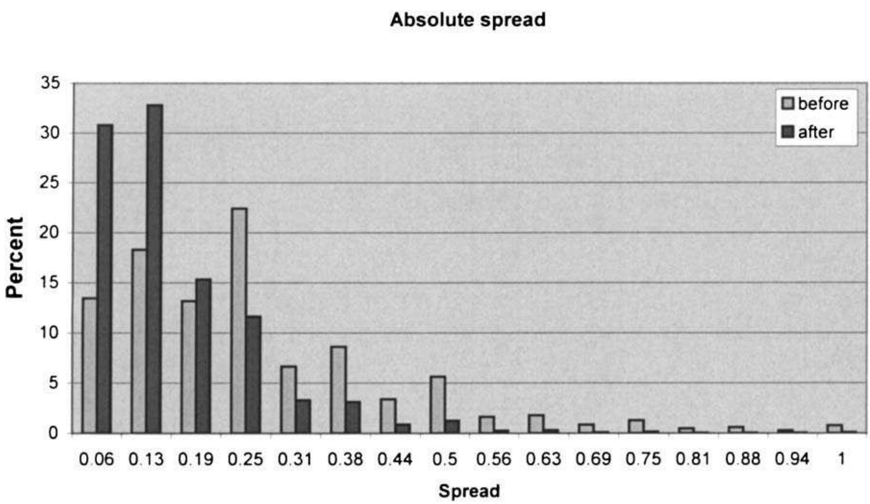


Fig. 6. Distribution of Absolute Spreads After Tick Size Change in June 1997.

Post-splits distribution of absolute spreads shows that majority of the spread (59.74%) occurs at 1/8. After the tick size change on June 24, 1997, we observe similar shift of quoted absolute spreads towards the first and second ticks. The quotes at 1/16 more than doubled (30.76% versus 13.43%) and quotes at 1/8 also increase substantially post splits. This clearly indicates that the absolute cost of transaction significantly decrease following splits, which is consistent with the positive relationship between share price and absolute spread documented in other studies.

However, in terms of percentage spread, we find that the percentage quoted and effective spreads increase as share prices become lower. As shown in Figures 1 and 2, percentage spreads following splits increase significantly for all thirteen intraday intervals.

Table 2 shows the change in the percentage spreads following ADR stock splits. The means of the standardized percentage spreads before and after splits are 0.0066 and 0.0097, respectively. Percentage spreads are higher following splits suggesting that absolute spreads do not fall proportionally with share prices. The difference in percentage spread is highly statistically significant. Similar results are observed for effective spread, which increases from the pre-

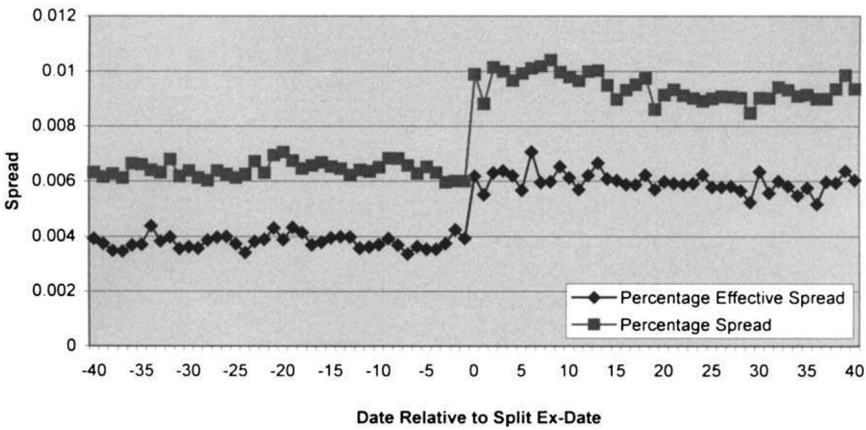


Fig. 7. Comparison of Percentage Quoted and Effective Spreads Before and After Stock Splits.

This figure presents percentage quote and effective spreads for 40 days before and after split date. Percentage quote and effective spreads are computed for each stock-day. Averages for the day are then computed as the equally weighted mean across all split stocks and date 0 indicates split date.

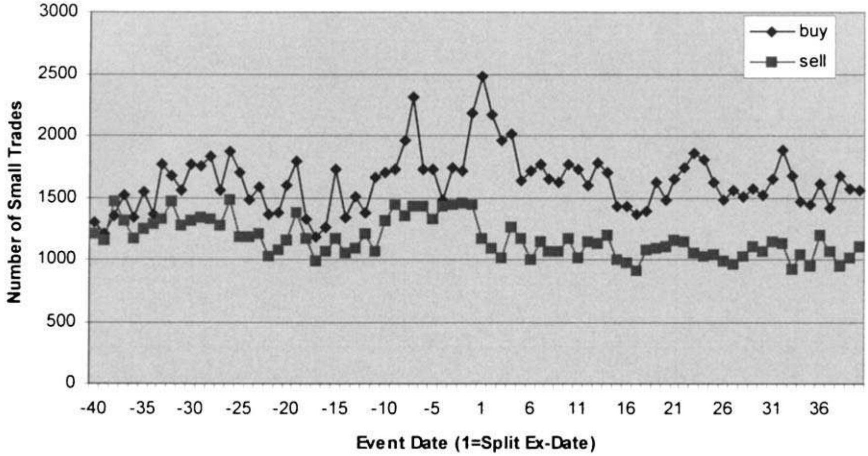


Fig. 8. The Aggregate Number of Small Buys and Small Sells Around Splits.

A trade is defined as a buy (sell) order, if the trading price is greater (less) than the average of bid and ask quote. We defined the quote as the most recent quote that were time stamped at least five seconds before trade following Lee and Ready (1991).

split level of 0.0041 to 0.0062 post split. The results clearly suggest that that liquidity costs to investors increases post splits. The association of splits with increases in percentage spreads is a direct result of the share-price effect of splits and is in agreement with previous studies such as Gray, Smith and Whaley (1996), Maloney and Mulherin (1992), and Conroy, Harris and Benet (1990).

Table 2. The Change in Percentage Quoted Spread & Percentage Effective Spread.

	Pre-Split	Post-Split	t-statistics
Percentage Quoted Spread	0.0066	0.0097	15.95**
Percentage Effective Spread	0.0041	0.0062	16.52**

** indicates significance at the 1% level.

Table 3. The Change in Depth, and Volume Following Splits.

Depth, Ask Depth, Bid Depth, and Volume are all adjusted with split-factors according to Copeland (1979). In addition, each variable is also standardized based on the following procedures: $STDEP_{k,i} = (D_{k,i} - \bar{D}_i)/\sigma(D_i)$; and $STVOL_{k,i} = (V_{k,i} - \bar{V}_i)/\sigma(V_i)STDEP_{k,i}(STVOL_{k,i})$ denotes the standardized depth (volume) of quote k for stock i; $D_{k,i}(V_{k,i})$ is the depth (volume) of quote k for stock i; $\bar{D}_i(\bar{V}_i)$ and $\sigma(D_i)(\sigma(V_i))$ are the mean and standard deviation of $D_{k,i}(V_{k,i})$.

	Pre-Split	Post-Split	t-statistics
Depth	0.0256	–0.023	–19.12**
Ask Depth	0.0141	–0.013	–10.54**
Bid Depth	0.0271	–0.025	–20.23**
Volume	0.0913	–0.083	–5.46**

** indicates significance at the 1% level.

In Table 3, we report quantity dimension of liquidity such as depth, ask depth, bid-depth, and volume. All these variables are split-factor adjusted and standardized. There is consistent evidence across all these measures that liquidity measured in dollar value deteriorates after split, suggesting that depth and volume do not increase proportionally with splits. Copeland (1979), and Lamoureux and Poon (1987) also document the negative relationship between stock split and split adjusted trading volume although depth was not investigated.

Thus far, our evidence on ADR splits seems to suggest that investors experience a deterioration of liquidity in both price and quantity dimensions following splits. Nevertheless, Angel (1997) posits that there are more to liquidity than bid-ask spreads. While the cost to liquidity demanders goes up, cost to liquidity providers may decline as a result of larger relative tick size post split. Liquidity may improve because of lower bargaining and processing costs, higher incentives for promoting the split-up stocks and for limit orders and market makers to provide liquidity. Schultz (2000) provides weak evidence in support of the notion of lower cost of market making following splits. The higher incentive in providing liquidity is due to the fact that a larger relative tick size may lead to more profitable market making, providing brokers with higher incentives to promote the split-up stocks. Shultz also suggests that since payment for order flow received by brokers is usually a fixed amount (typically 1 cent to 2 cents) per share, transactions of a given dollar amount involves larger number of shares following splits and thus more payment to the broker for the order flow.

Splits and Liquidity Improvement

To empirically examine whether cost of liquidity providing decreases as a result of stock split, we critically evaluate several liquidity measures that can be used as proxies. Since brokers benefit from payment for order flows when transaction volume goes up, we believe that increase in raw trading volume and depth is a necessary condition. To this end, we report the depth and volume without the adjustment of split factors volume in panel A of Table 4 and find that they are all much higher following splits.[6]

We report the before and after average depth and In addition, to gauge the potential impact of brokers in promoting the split-up shares, we also examine whether the number of small trades has changed. Trades up to 1000 shares are categorized as small trades similar to Muscarella and Vetsuypens (1996) and are compiled from the TAQ database for 40 days preceding the split and 40 days following the split. Reported in panel B of Table 4 are the number of trades and the number of small trades. Observe that the number of trades rises an average of 19% post splits, and at the same time the averages of small trades across all ADRs show a much larger increase of almost 44% for trades of less than 500 shares and 29% for trades of less than 1000 shares. Both t-statistic and z-statistic show that the difference in means is statistically significant.

Shultz (2000) reports that small trades increase subsequent to splits and the overwhelming majority of these orders are buys. We tabulate small buys and sells 40 days (transaction of less than 1000 shares) before and after splits for our ADR sample (see panel C of Table 4) and observe similar patterns. Small buys increase 37% while small sells go up 29%, indicating that small shareholder base is enlarged post split.

Another proxy to gauge the effort of brokers is to see whether there is an increase in the number of shareholders, provided that small investors is the primary target of brokerage promotion. We collect information on number of shareholders for the year before and after splits. Consistent with earlier studies by Demsetz (1977), Copeland (1979), Baker and Gallagher (1980), and Lamoureux and Poon (1987), we observe a significant increase in the number of shareholders (Panel D of Table 4). The average number of shareholders increases about 21% (goes up to 25,815 from 21,393 before splits). This is also consistent with the hypothesis that liquidity improves with a broader shareholder base and this improved liquidity in the form of "awareness" is expected to have a significant positive effect on the value of the firm.

Interestingly, when institutional ownership is compared before and after splits, the results reported in Panel E of table 4 indicate that the percentage of institutional ownership declines post split, although the difference is not

Table 4. Other Liquidity Measures.

Depth, Ask Depth, Bid Depth, and Volume are not adjusted with split-factors, but are standardized. Trades up to 1,000 shares are categorized as small trades and are compared for 40 days preceding and 40 days after splits. The aggregate number of small buys and small sells are computed for 40 days before and 40 days after splits. A trade is defined as a buy (sell) if the trading price is greater (less) than the average of bid and ask quotes. The quote is defined as the most recent quote that is time stamped at least five seconds before a trade following Lee and Ready (1991). The number of shareholders and institutional holdings are collected for a year prior and a year after a split.

Panel A: Depth and Volume without Adjustment of Split Factors

	Pre-Split	Post-Split	t-statistics
Depth	–0.3420	0.2811	17.06**
Ask Depth	–0.2660	0.2180	16.37**
Bid Depth	–0.2590	0.2132	14.74**
Volume	–0.2920	0.2649	11.21**

Panel B: Average Daily Number of Trades and Number of Small Trades

	Pre-Split	Post-Split	Change (%)	t-statistic	Z-statistic
Number of trades	80.2953	95.9393	19.48	2.55*	3.50**
Trade Size ≤ 500	45.6156	65.5091	43.61	2.69*	3.39**
Trade Size ≤ 1000	59.2613	76.7236	29.47	2.55*	2.83**

Panel C: Number of Small Buys and Sells

	Pre-Split	Post-Split	Change (%)	t-statistic	Z-statistic
Buy	1400.42	1916.74	36.87	2.51*	3.48**
Sell	1075.38	1389.90	29.25	2.57*	2.54*

Panel D: The Change in Number of Shareholders Before and After Splits

Pre-Split	Post-Split	Change (%)	t-statistic	Z-statistic
21,393	25,815	20.67	2.43*	3.43**

Panel E: The Change in Institutional Ownership (in%) Before and After Splits

Pre-Split	Post-Split	Change (%)	t-statistic	Z-statistic
11.38	8.70	–23.61	1.32	0.21

* and ** indicate significance at the 5% and 1% level, respectively.

statistically significant. This further collaborates our earlier findings that liquidity supply targeting small investors improves, consistent with Schultz (2000) which also provides evidence that small buy orders significantly increase and large investors sell the split-up stocks.

4. SUMMARY

Using transactions data from a sample of stock splits on NYSE listed ADRs, we investigate liquidity changes following splits.

We first examine the intraday patterns for percentage quoted and effective spreads, depths and volume before and after splits. Percentage spread and volume across a trading day follow a U-shape, consistent with other studies and the asymmetric information and inventory models of microstructure. Interestingly, the depth pattern is different from the previous findings which show lower depth in the first half hour of trading, higher depth during the middle of trading, and a reversion to lower level toward market close (a reverse U-shape). Instead, depth is higher close to the end of a trading day, contrary to the prediction of the inventory model. Since ADRs are larger firms, with higher stock prices, larger number of market makers, and lower volatility than average firms listed on NYSE (Howe & Lin, 1992), specialists perhaps are not as strongly averted to carrying inventory overnight as they would do for average stocks on NYSE.

Muscarella and Vetsuypens (1996) suggest that the positive response of share prices to ADR stock splits is likely the result of improved liquidity, since the unique construct of their sample rules out signaling as the motivation. However, their findings of liquidity change are based a sample of 7 ADRs and a small subset of liquidity measures. This paper shed light on a more comprehensive picture of liquidity changes following splits.

We find that the absolute spreads decline as share prices become lower post split. However, the reduction in absolute spreads is not proportional to change in share price, which leads to a higher percentage quoted and effective spread on split-up securities. The higher transaction cost on ADR splits is consistent with findings for NYSE listed shares and other securities (Gray, Smith & Whaley, 1996; Maloney & Mulherin, 1992; and Conroy, Harris & Benet, 1990). We also examine the quantity dimension of liquidity measures, namely, split-factor adjusted volume and depth, and find that liquidity deteriorates post splits as well.

Angel (1997) posits while the cost to liquidity demanders goes up, cost to liquidity providers may decline as a result of larger relative tick size post split.

Liquidity may improve because of higher incentives for brokers to promote the split-up stocks to benefit from payment for order flows. Thus, transactions of a given dollar amount involves larger number of shares bring more payments to the broker for the order flow.

To this end, we believe that the raw volume and depth (not adjusted for split-factors) are informative. A comparison of the depth and volume before and after split indicates that both increase significantly following splits, suggesting that this may result from brokers heightened efforts in promoting the split-up shares.

Since the effort of brokers' targets primarily smaller traders, we also examine whether there is any change in the number of shareholders and small trading following splits. Consistent with other studies, we find significant increase in liquidity as reflected in a 21% increase in number of shareholders and a 28% increase in small trades following splits. Also, small buy orders increase at a faster rate than small sells, suggesting that small shareholder base is enlarged post splits. Nevertheless, institutional holding, although lower in percentage, is not significantly different after splits. The increase in the number of shareholders is also consistent with the hypothesis that liquidity improves with a broader shareholder base.

NOTES

1. Using TORQ data for the period between November 1990 and January 1991, Angel (1997) presents evidence that there is a higher fraction of limit orders placed on lower priced stocks.
2. Since we do not have "solo splits", our study is not intended to examine whether signaling or liquidity improvement is the driving force of ADR splits. And clearly this has been well accomplished in Muscarella and Vetsupens (1996).
3. The minimum tick size on NYSE changed from $1/8 to $1/16 for stocks with a price greater than $1 on June 24, 1997.
4. We exclude those ADRs that have more than one split during the sample period since the data on ownership is collected over a longer time period than the transaction data.
5. COMPUSTAT includes 277 NYSE traded ADRs in 1999, accounts for almost 10% of all firms listed on NYSE.
6. Lamoureux and Poon (1987) posit the tax-options value of the split-up stocks rises with the increased noisiness of the security's return resulted from higher raw trading volume.

REFERENCES

Amihud, Y., & Mendelson, H. (1982). Asset pricing behavior in a dealership market. *Financial Analysts Journal*, *38*, 50–59.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*, 223–249.

Amihud, Y., & Mendelson, H. (1987). Trading mechanisms ans stock returns: An empirical investigation. *Journal of Finance*, *42*, 533–553.

Amihud, Y., & Mendelson, H. (1988). Liquidity and asset prices: Financial management implications. *Financial Management*, *17*, 5–15.

Angel, J. J. (1997). Tick size, share prices, and stock splits. *Journal of Finance*, *52*, 655–681.

Baker, K. H., & Gallagher, P. L. (1980). Management view of stock splits. *Financial Management*, *9*, 73–77.

Bacidore, J. M., & Sofianos, G. (2001). Liquidity provisions and specialist trading in NYSE-listed non-U.S. stocks. *Journal of Financial Economics*, forthcoming.

Brennan, M. J., & Hughes, P. J. (1991). Stock prices and the supply of information. *Journal of Finance*, *46*, 1665–1691.

Brock, W. A., & Kleidon, A. W. (1992). Periodic market closure and trading volume: A model of intraday bids and asks. *Journal of Economic Dynamics and Control*, *16*, 451–489.

Chakravarty, S., & Wood, R. A. (2000). The effects of decimal trading on market liquidity. Working paper. Purdue University.

Chung, K. H., Van Ness, B., & Van Ness, R. (1999). Limit orders and the bid-ask spread. *Journal of Financial Economics*, *53*, 255–287.

Conroy, R. M., Harris, R. S., & Benet, B. A. (1990). The effects of stock splits on bid-ask spreads. *Journal of Finance*, *45*, 1285–1295.

Copeland, T. E. (1979). Liquidity changes following stock splits. *Journal of Finance*, *34*, 115–141.

Copeland, T. E., & Galai, D. (1983). Information effects on the bid-ask spread. *Journal of Finance*, *38*, 1457–1469.

Demsetz, H. (1977). The cost of transacting. *Quarterly Journal of Economics*, *82*, 33–53.

Glosten, L., & Milgrom, P. R. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents. *Journal of Financial Economics*, *14*, 71–100.

Goldstein, M. A., & Kavajecz, K. A. (2000). Eighths, Sixteenths and Market Depth: Changes in Tick Size and Liquidity Provision on the NYSE. *Journal of Financial Economics*, *56*, 125–149.

Gray, S. F., Smith, T. M., & Whaley, R. E. (1996). Stock splits: Implications for models of the bid-ask spread. Working paper. Duke University.

Harris, L. (1994). Minimum price variations, discrete bid-ask spreads, and quotation sizes. *Review of Financial Studies*, *7*, 149–178.

Howe, J. S., & Lin, J-C. (1992). The bid-ask spreads of American Depository Receipts. In: S. G. Rhee & R. P. Chang (Eds), *Pacific-Basin Capital Markets Research* (Vol. III).

Huang, R. D., & Stoll, H. R. (1996). Dealer versus auction markets. *Journal of Financial Economics*, *41*, 313–357.

Kavajecz, K. A. (1999). A specialist's quoted depth and the limit order book. *Journal of Finance*, *54*, 747–771.

Lamoureux, C. G., & Poon, P. (1987). The market reaction to stock splits. *Journal of Finance*, *42*, 1347–1370.

Lee, C. M. C., & Ready, M. J. (1991). inferring trade direction from intraday data. *Journal of Finance*, *46*, 733–746.

Lee, C. M. C., Mucklow, B., & Ready, M. J. (1993). Spreads, depths, and the impact of earnings information: Intraday analysis. *Review of Financial Studies*, *6*, 345–374.

Maloney, M. T., & Mulherin, H. J. (1992). The effects of splitting on the ex: A microstructure reconciliation. *Financial Management*, *21*, 44–59.

McInish, T., & Wood, R. (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *Journal of Finance*, *47*, 753–764.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, *42*, 483–511.

Muscarella, C. J., & Vetsuypens, M. R. (1996). Stock splits: Signaling or Liquidity? The case of ADR 'solo-splits'. *Journal of Financial Economics*, *42*, 3–26.

Schultz, P. (2000). Stock splits, tick size and sponsorship. *Journal of Finance*, *55*, 429–450.

This Page Intentionally Left Blank

PORTFOLIO SELECTION WITH ROUND-LOT HOLDINGS

Clarence C. Y. Kwan and Mahmut Parlar

ABSTRACT

A round lot being the standard unit of stock trading, we consider portfolio selection with round-lot requirements in analytical settings where short sales are disallowed and allowed. In either case, by exploiting some analytical properties of the objective function in portfolio optimization, we are able to approximate the round-lot solution without the encumbrance of any algorithmic complexities that are often associated with integer programming. The efficient heuristic we use to solve the resulting nonlinear integer programming problem examines only the corner points of a "hypercube" surrounding the optimal fractional solution found without the round-lot requirements. Then, by characterizing the covariance structure of security returns with the single index model, we establish the correspondence between the round-lot solution and the solution without round-lot requirements for which security selection criteria in terms of risk-return trade-off are available. This correspondence, in turn, provides useful information regarding the sensitivity of the round-lot solution in response to changes in return expectations. Given these nice features, the analysis should enhance the practical relevance of portfolio modeling for assisting investment decisions.

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 133–163.

ISBN: 0-7623-0887-7

1. INTRODUCTION

Modern portfolio theory dates back to the pioneering work of Markowitz (1956, 1959). Although the portfolio concept is a key component of financial education, as covered extensively in finance textbooks, few academic studies devoted to normative portfolio analysis have appeared in research journals with broad readership among practitioners in investment analysis and portfolio management. The apparent lack of practical appeal of academic portfolio analysis can be attributed, in part, to the nature of the analysis itself. No matter how a portfolio model is formulated, in a mean-variance framework or otherwise, the basic idea is still to optimize an objective function, subject to some conditions or restrictions in the allocation of investment funds within a given set of securities. Analytical tractability being an overriding concern, the reliance on simplifying assumptions is inevitable. Some assumptions involved are perhaps too unrealistic for the corresponding analysis to have much practical relevance.

The textbook material on portfolio analysis can serve as an example to illustrate the above point. As in equilibrium analysis leading to asset pricing models, including the CAPM, normative portfolio analysis as covered in finance textbooks typically assumes that short sales are allowed and that the short seller of a security not only provides no margin deposit but also can immediately use the short-sale proceeds to purchase other securities. Such unpalatable assumptions have made it possible to solve portfolio selection problems by using basic algebraic tools. Predictably, the short-sale benefit being over-stated, the corresponding analysis often results in unrealistic allocations of investment funds in both long and short positions. Under these assumptions, each security considered will always be selected for the portfolio; that is, the analysis cannot filter out any securities that are suitable for neither long nor short holdings.

The relaxation of these assumptions in an attempt to capture reality adequately, however, can only make the analysis more complicated. Even in a simple no-short-sale case where no simplifying assumption is made to the covariance structure of security returns, sophisticated solution methods such as the Markowitz (1956, 1959, 1987) critical line algorithm or its variants will have to be applied. [The intuition underlying the critical line algorithm has been covered in a small number of studies containing pedagogical materials, as well as advanced investment textbooks. For example, a tangency-portfolio approach, which simplifies the algorithm, can be found in Kwan and Yuan (1993), and a geometric presentation of the critical line idea for a three-security case can be found in Haugen (1997, Ch. 5).] The major task there is to identify which of the securities considered are to be selected for the portfolio. Once such securities are identified, the allocation of investment funds among them can easily be achieved. Further conditions or restrictions to the portfolio choice – such as investment limits on individual securities for satisfying institutional requirements or for reducing the potential impact of any overly optimistic input parameters on the portfolio weights – for the sake of making the analysis more practically relevant often have just the opposite effect because of the attendant analytical complexities. For achieving computational efficiency, the corresponding solution methods are often some algorithmic exercises that are of little interest to the potential users of the analysis.

It is therefore not surprising that, although normative portfolio analysis is intended for assisting practical investment decisions, the corresponding research has not attracted much attention of portfolio managers. The availability of computer software packages that can solve numerically many constrained optimization problems without requiring the users to understand the technical details has also reduced any remaining incentives for further research in portfolio analysis. Indeed, the apparent lack of practical appeal of many earlier portfolio models, along with the academic finance profession's preference for equilibrium analysis over normative analysis, has virtually stalled the academic research in normative portfolio theory for many years. The commentary by Frankfurter (1990), in an effort to revitalize the profession's interests in normative analysis, has nicely captured such reality in the finance literature.

Following the gradual acceptance of short selling as a viable investment tool by institutional investors in recent years, however, there have been renewed academic interests in normative portfolio modeling. Recent studies, such as Alexander (1993, 1995) and Kwan (1995, 1997, 1999), have placed special emphasis on accurately portraying short-sale transactions while maintaining analytical tractability and, in some cases, computational simplicity. In view of the insight of Alexander (1993) that the short sale of a security under institutional procedures can be treated as the purchase of an artificially constructed security, these studies are able to retain many familiar analytical features of portfolio selection without short sales. These studies have responded well to the need for practical relevance in normative portfolio modeling.

To enhance the usefulness of portfolio analysis with or without short sales, we seek to provide a further refinement by considering yet another practical issue. An interesting feature of the proposed analysis is that this practical refinement can still be achieved without resorting to highly sophisticated mathematical tools. Specifically, the issue here pertains to round-lot holdings in the allocation of investment funds. A round lot, which generally represents 100 shares, is the standard unit of stock trading. As securities are normally traded in round lots, both liquidity and transaction-cost considerations would favor the selection of securities in round-lot holdings. For analytical convenience, however, available portfolio models as reported in the literature, with or without short sales, have all been formulated under the assumption that fractional shares can be held. Obviously, for portfolios with sizable investment funds, to translate each odd-lot and fractional share holding into an approximate number of round lots is an easy task. For small portfolios to be constructed for individual investors, in contrast, it is unclear as to how best odd-lot and fractional share holdings can be approximated by round-lot holdings without severely affecting each portfolio's risk-return trade-off. Indeed, the smaller the available investment capital for portfolio construction, the more pertinent is the round-lot issue.

Analytically, the search for round-lot portfolio solutions belongs to nonlinear integer programming problems. To solve an integer programming problem efficiently is typically an algorithmic exercise which may require sophisticated mathematical tools. The analysis, no matter how algorithmically elegant and computationally efficient, tends to lack practical appeal because the potential users may not be familiar with or interested in the attendant technical details. By exploiting some analytical properties of the portfolio selection problems considered, however, we are able to produce near optimal round-lot holdings by directly revising the corresponding portfolio allocations without round-lot requirements. The proposed approach is an easily implemented "hypercube" method. This method, as explained below, allows the analysis to bypass any advanced mathematical tools that are normally associated with integer programming, thus making the analysis more accessible to investors.

Unlike the more traditional formulations where the input parameters for the analysis are defined in terms of rates of return, their counterparts in the present analysis are defined in terms of prices. This allows the decision variables – the individual security holdings in a portfolio – to be stated in terms of numbers of lots rather than fractions of the available investment funds. As will be explained in more detail later in this study, if the optimal holding of a security i in a portfolio without imposing the round-lot constraints is N_i lots, the corresponding optimal round-lot holding must be one of the two consecutive integers bracketing the real number N_i. For example, if $N_i = 2.67$, the optimal round-lot holding will be in the neighborhood of 2 or 3. Therefore, for a portfolio selection problem based on n securities labeled as $1, 2, \ldots, n$, where in the absence of round-lot requirements the holdings $N_1, N_2, \ldots, N_n$ are known, a near-optimal round-lot holdings must correspond to one of the 2^n corners of the hypercube with unit-length edges enclosing the point $(N_1, N_2, \ldots, N_n)$ in an n-dimensional space. More precisely, only integer values pertain to the coordinates of each corner of the hypercube.

Given the above analytical property, the search for the approximately optimal round-lot holdings can be performed by evaluating the corresponding objective function and the feasibility of the resulting portfolios for the 2^n potential cases. Although the computations as required by the hypercube method increase exponentially with the number of securities considered, the analysis should not be burdensome given its intended purpose. Since round-lot solutions pertain primarily to cases of small investment capital, the corresponding portfolio construction is most likely based on a small number of securities. In the case of $n = 15$, for example, the total number of corners of the hypercube to be evaluated, $2^{15} = 32768$, should correspond to an easily manageable amount of computations for any Pentium-based microcomputers.

The analytical equivalence between the short sale of a security and the purchase of its artificially constructed counterpart is an interesting feature which allows the hypercube method to be applied equally well to both short-sale and no-short-sale cases. Technical issues pertaining to sophisticated integer programming algorithms not being of any concern, we are able to pursue the underlying economic rationale of the portfolio decision, as well as the analytical correspondence between the round-lot solution and the solution without imposing the round-lot constraints. Specifically, by characterizing the covariance structure of security returns by a single index model as in Elton, Gruber and Padberg (1976, 1978), a ranking approach which reveals the portfolio-selection criteria can likewise be established in the present setting. Considering that the analysis is intuitively appealing, easy to implement, and able to capture the institutional environment for portfolio investment, it should be of interest to investors to assist their practical investment decisions.

The remainder of the paper is organized as follows. Section 2 first formulates a portfolio selection problem without short sales under a full-information covariance structure of security returns, where the decision variables are security holdings in numbers of lots. Then, in view of the analytical equivalence between the short sale of a security and the purchase of an artificially constructed security, we extend this formulation to the short-sale case in the same section. The extension allows us to formulate and solve portfolio selection problems with round-lot restrictions, where institutional procedures for short selling are followed. In Section 3 we introduce and justify the hypercube method with geometric illustrations. Starting with a two-security case and then a three-security case, where the analytical aspect of the method can be visualized and explained geometrically, we formalize the intuitive idea that each round-lot solution tends to be in the vicinity of its corresponding solution without round-lot restrictions.

Section 4 solves the no-short-sale case under a single index characterization of the covariance structure. In so doing, we are able to establish an analytical correspondence between the round-lot solution and the solution without round-lot requirements, for which the economic rationale of the portfolio decision in terms of risk-return trade-off is available. A corresponding analysis for the short-sale case is presented in Section 5. In each case a numerical example is provided to illustrate the analytical features involved. As explained analytically and illustrated numerically for each case, the comparison between the round-lot solution and the corresponding solution without round-lot requirements provides useful information regarding the sensitivity of the portfolio optimization results in response to changes in price expectations of individual securities. Section 6 summarizes and concludes the present study.

2. PORTFOLIO SELECTION WITH AND WITHOUT SHORT SALES: MODEL FORMULATION

In a single-period setting without transaction costs where there are n risky securities for portfolio consideration, we define each decision variable $N_i(\geq 0)$ as the number of units of security i to be held, for $i = 1, \ldots, n$, with a unit representing a round lot of the corresponding shares. Among the input parameters for the analysis, P_{0i} is the initial (beginning-of-period) unit price of security i and W_0 is the investor's initial wealth (total amount of funds for the portfolio investment). We also define P_i as the random terminal (end-of-period) unit price including any dividend component and W as the random terminal wealth. Then, the total investment among the n securities amounts to $\sum_{i=1}^{n} N_i P_{0i}$. The remainder, $W_0 - \sum_{i=1}^{n} N_i P_{0i} \geq 0$, is invested risk-free for a given rate of return r over the period. The random terminal wealth, therefore, is

$$W = \sum_{i=1}^{n} N_i P_i + (1 + r)\left(W_0 - \sum_{i=1}^{n} N_i P_{0i} \right).$$

Accordingly, the expected terminal wealth is

$$E(W) = \sum_{i=1}^{n} N_i \bar{P}_i + (1 + r)\left(W_0 - \sum_{i=1}^{n} N_i P_{0i} \right),$$

where $\bar{P}_i \equiv E(P_i)$ is the expected terminal unit price of security i, and the variance of the terminal wealth is

$$\text{Var}(W) = \sum_{i=1}^{n} \sum_{j=1}^{n} N_i N_j \,\text{Cov}(P_i, P_j),$$

where $\text{Cov}(P_i, P_j)$ is the covariance of the terminal unit prices P_i and P_j.

If the entire initial wealth is invested risk-free instead, the terminal wealth will be $(1 + r)W_0$. The portfolio selection problem without short sales can be formulated as a constrained optimization of the portfolio's expected performance

$$\theta = \frac{E(W) - (1 + r)W_0}{\sqrt{\text{Var}(W)}}. \tag{1}$$

The objective function θ, a dimensionless quantity, captures the incremental expected wealth from the above investment beyond a purely risk-free investment per unit of risk exposure. The optimization problem, with $N_1, N_2, \ldots, N_n$ being the decision variables, can be stated as

$$\max \theta = \frac{\sum_{i=1}^{n} N_i[\bar{P}_i - (1 + r)P_{0i}]}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} N_i N_j \,\text{Cov}(P_i, P_j)}} \tag{2}$$

subject to the constraints

$$\sum_{i=1}^{n} N_i P_{0i} \leq W_0,$$

$$N_i \geq 0 \text{ and integer valued, for } i = 1, 2, \ldots, n. \tag{3}$$

Notice that condition (3) has been stated as an inequality to accommodate a risk-free investment and the integer requirements on the n decision variables. This inequality allows the available investment funds W_0 to be fully utilized, considering that, for the given prices $P_{01}, P_{02}, \ldots, P_{0n}$, the integer requirements on $N_1, N_2, \ldots, N_n$ may render the equality between W_0 and $\sum_{i=1}^{n} N_i P_{0i}$ unattainable.

This inequality, however, can lead to multiple solutions. The reason is that, the objective function θ in (2) being homogeneous of degree zero, its value is unaffected when the decision variables $N_1, N_2, \ldots, N_n$ are scaled up or down by a common positive multiplicative constant. Analytically, suppose that the integers $N_1, N_2, \ldots, N_n$ for which θ is maximized without violating condition (3) have their greatest common divisor $d > 1$. Suppose further that the integer d has k divisors (including 1 and d) labeled as $d_1, d_2, \ldots, d_k$. Then, the scaling of the integers $N_1, N_2, \ldots, N_n$ can be achieved feasibly by any of $1/d_h$, for $h = 1,2, \ldots, k$. That is, each of the k sets of round-lot holdings of the n securities, $N_1/d_h, N_2/d_h, \ldots, N_n/d_h$, for $h = 1,2, \ldots, k$, corresponds to the same optimal θ for which condition (3) is satisfied.

For example, for a two-security case where a round-lot solution is $(N_1, N_2) = (5, 15)$, since the greatest common divisor of 5 and 15 is $d = 5$ which has the divisors $d_1 = 1$ and $d_2 = 5$, both $(N_1, N_2) = (5/d_1, 15/d_1) = (5, 15)$ and $(N_1, N_2) = (5/d_2, 15/d_2) = (1, 3)$ are acceptable solutions. Similarly, for a three-security case where a round-lot solution is $(N_1, N_2, N_3) = (12, 6, 18)$, since $d = 6$ which has 4 divisors (e.g. 1, 2, 3, and 6), all 4 round-lot solutions, (12, 6, 18), (6, 3, 9), (4, 2, 6), and (2, 1, 3), correspond to the same θ. As obvious in these numerical illustrations, the relative distributions of investment funds among risky securities, represented by the ratio N_i/N_j for all i and j, are unaffected by the scaling. In order to ensure a unique round-lot solution, therefore, we simply assume that the investor intends to allocate as much as possible his/her available investment funds among risky securities.

Before introducing the hypercube method for an integer solution, we extend the above formulation to the short sale case. As in Alexander (1993, 1995) and Kwan (1997), we treat the short sale of a security as the purchase of an artificially constructed security. Therefore, the corresponding optimization problem becomes the case of portfolio selection among $2n$ risky securities, where the non-negative decision variables are $N_1, N_2, \ldots, N_{2n}$, representing the numbers of lots held. For $i = 1,2, \ldots, n$, each security $n + i$ is the artificial counterpart of security i. As will be clear later, under this formulation many analytical features of the original no-short-sale case can be retained.

Letting λ be the margin deposit as a fraction of the share value that is sold short, the total investment in risky securities becomes $\sum_{i=1}^{n} (N_i P_{0i} + \lambda N_{n+i} P_{0i})$. The initial wealth W_0 net of this amount is invested risk-free. Following institutional procedures for short selling, the margin deposit earns a risk-free return r and the short seller may receive a partial interest rebate on the short-sale proceeds that the brokerage firm earns. With $0 \le \nu < 1$ being the fraction of interest rebated, the random terminal value is

$$W = \sum_{i=1}^{n} \{N_i P_i + N_{n+i}[-P_i + (1+\nu r)P_{0i} + (1+r)\lambda P_{0i}]\}$$

$$+ (1+r)\left[W_0 - \sum_{i=1}^{n} (N_i P_{0i} + \lambda N_{n+i} P_{0i})\right].$$

Then, the expected incremental terminal wealth beyond a purely risk-free investment is

$$E(W) - (1+r)W_0 = \sum_{i=1}^{n} \{N_i[\bar{P}_i - (1+r)P_{0i}] + N_{n+i}[-\bar{P}_i + (1+\nu r)P_{0i}]\}. \quad (4)$$

The corresponding variance of terminal wealth is

$$\mathrm{Var}(W) = \sum_{i=1}^{n} \sum_{j=1}^{n} (N_i N_j - N_i N_{n+j} - N_{n+i} N_j + N_{n+i} N_{n+j})\mathrm{Cov}(P_i, P_j). \quad (5)$$

The expected portfolio performance θ, as given by Eq. (1), can be established accordingly.

Geometrically, the solution to either of the above portfolio selection problems without round-lot requirements, represented by $(N_1, N_2, \ldots, N_n)$ or $(N_1, N_2, \ldots, N_{2n})$ as the case may be, is a point inside the corresponding hypercube as described in the introductory section. As shown in the next section, given the functional form of the objective function in each case, the round-lot solution is likely to be in the vicinity of the hypercube, and the search for an approximate solution can easily be performed by evaluating all corners of the hypercube. Given the analytical similarities between the no-short-sale and short-sale cases, the search method works equally well for both. For ease of exposition, we use the no-short-sale case to justify the method.

3. ROUND-LOT HOLDINGS AND THE HYPERCUBE METHOD

In general, when the n non-negative decision variables $N_1, N_2, \ldots, N_n$ of a constrained optimization problem are required to be integer valued, to solve the problem is a non-trivial matter. To illustrate the idea underlying the hypercube method for reaching round-lot portfolio holdings, we present in the following a simple numerical example with $n = 2$. For the two-security problem, assume that the initial prices are $(P_{01}, P_{02}) = (2000, 1000)$, the expected terminal prices are $(\bar{P}_1, \bar{P}_2) = (2200, 1140)$, the variances of terminal prices are $(\sigma_1^2, \sigma_2^2) = (100^2, 80^2)$, the covariance of terminal prices is $\mathrm{Cov}(P_1, P_2) = 800$, and the risk-free interest rate is 5%. Assume for now that there are no round-lot requirements and the initial wealth $W_0 = 20000$ is to be invested entirely in the two risky securities. Then, with the expected wealth $E(W) = 100N_1 + 90N_2 + 21000$ and the variance of wealth $\mathrm{Var}(W) = 10000N_1^2 + 6400N_2^2 + 1600N_1N_2$, the portfolio's expected performance is

$$\theta = \frac{1}{20}\,\frac{100N_1 + 90N_2}{\sqrt{25N_1^2 + 16N_2^2 + 4N_1N_2}}.$$

The objective here is to maximize the function θ subject to the budget constraint $2000N_1 + 1000N_2 = 20000$ and the non-negativity constraints $N_1, N_2 \geq 0$. Given $N_2 = 20–2N_1$, θ can be written as a function of N_1. It follows from $d\theta/dN_1 = 0$ that $N_1 = 5.8077$ and $N_2 = 8.3844$, with the corresponding $\theta = 1.4357$. Given that the non-negativity constraints are completely satisfied, the problem is solved as long as there are no round-lot requirements. Of course, if any non-negativity constraint is violated by the achieved values of N_1 and N_2 (for a different set of input data), a corner solution is implied; that is, the whole investment will be confined to a single risky security.

Before proceeding to solve the same problem with round-lot requirements, it is worth noting that, since round-lot investments in the two risky securities do not always exhaust the initial wealth, the corresponding budget constraint is $2000N_1 + 1000N_2 \leq 20000$ instead. Geometrically, the optimal round-lot solution must be among the lattice points in a triangle on the (N_1, N_2) plane as defined by the N_1 and N_2 axes and the line $2N_1 + N_2 = 20$. Here, a lattice point is a point with integer coordinates. Although in the case of $n = 2$ the round-lot solution can still be reached by exhaustive enumerations of the individual lattice points in the search of the optimal θ, such a method will become hopelessly cumbersome when more securities are involved. (When $n = 2$ a simple count indicates that there are 121 such lattice points including the origin.) In order to achieve algorithmic efficiency without resorting to sophisticated integer-programming tools, we exploit some functional properties of θ, as illustrated below with the same numerical example.

A property of the function θ, as indicated earlier, is that it is homogeneous of degree zero. That is, in the case of $n = 2$, multiplying the decision variables N_1 and N_2 by an arbitrary common positive constant has no impact on the corresponding value of θ. Therefore, in the absence of any budget constraint and round-lot requirements, an iso-performance locus – the set of points that includes all potential combinations of non-negative real numbers N_1 and N_2

corresponding to a common value of θ – as plotted on the (N_1, N_2) plane must consist of straight lines emanating from the origin. Strictly speaking, since θ is undefined at $N_1 = N_2 = 0$, these lines do not reach the origin. As shown in Fig. 1, each iso-performance locus is a V-shaped figure with two line segments. The family of all iso-performance loci is a nest of V-shaped figures; as θ increases, the angle between the two line segments narrows. Once the maximum of θ is reached at θ_{max}, the corresponding V-shaped figure reduces to a single line from the origin. The intersection of this line and the budget line $2N_1 + N_2 = 20$, which occurs at $(N_1, N_2) = (5.8077, 8.3844)$, is the solution without round-lot requirements. It is easy to visualize that the graph showing the values of θ versus the corresponding locations along the budget line has a single maximum at this particular combination of N_1 and N_2. Analytically, the uniqueness of the solution given the budget line is due to the concavity of the function θ (the concavity is not confined to the case of $n = 2$).

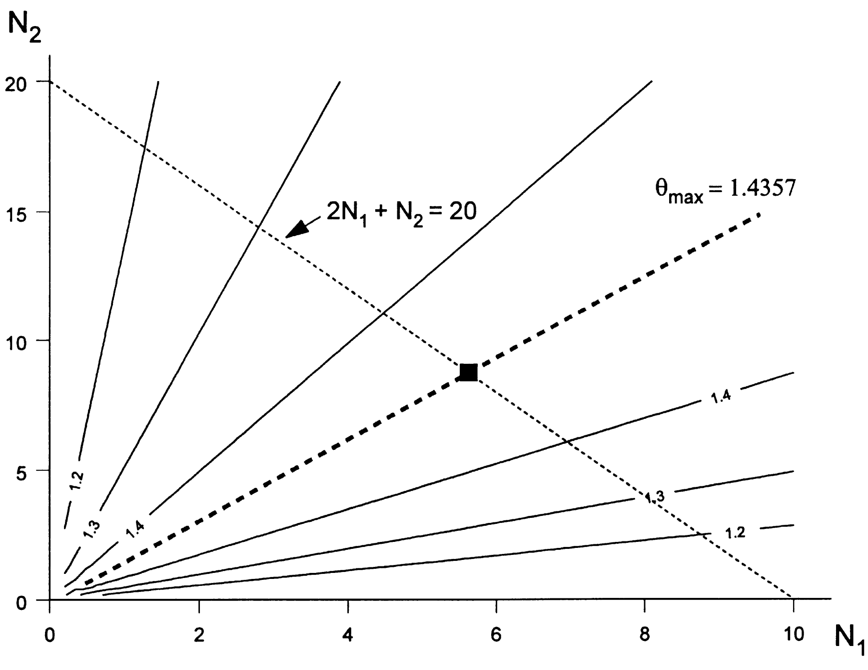


Fig. 1. Family of iso-performance loci for $n = 2$. The budget line $2N_1 + N_2 = 20$ and the line θ_{max} intersect at $(N_1, N_2) = (5.8077, 8.3844)$ indicated by the symbol ■. The line $\mathcal{L}$ emanating from the origin and passing through the point (5.8077, 8.3844) is defined by $N_2 = 1.444N_1$.

Now consider all lattice points in the triangle as defined by the N_1 and N_2 axes and the line $2N_1 + N_2 = 20$. The optimal round-lot solution satisfying the budget constraint of $2N_1 + N_2 \leq 20$ must be among these lattice points. In view of the geometric pattern of the iso-performance loci, however, the number of candidates can be reduced drastically.

Recall that θ is a concave function reaching its constrained maximum at $\theta_{\max} = 1.4357$ on the boundary $2N_1 + N_2 = 20$. The iso-performance line $\mathcal{L}$ corresponding to $\theta_{\max}$ emanates from the origin and passes through the point (5.8077, 8.3844); thus $\mathcal{L}$ has a slope of $8.3844 \div 5.8077 = 1.444$ and it can be written as $N_2 = 1.444N_1$. In order to find a round-lot solution that maximizes the concave θ function while satisfying the budget constraint $2N_1 + N_2 \leq 20$, it is sufficient to examine the feasible integer points in the neighborhood of the line $N_2 = 1.444N_1$. For the present two-security case this is a relatively easy matter as we now demonstrate.

The following table gives the feasible (N_1, N_2) combinations on the line $\mathcal{L}$ defined by $N_2 = 1.444N_1$ for integer values of N_1 ranging from 1 to 6. Since the optimal round-lot solution must be at an integer point in the neighborhood of the line $\mathcal{L}$, this table also provides a list of all candidate integer (N_1, N_2) values. For example, when $N_1 = 2$, we obtain $N_2 = 1.444 \times 2 = 2.888$. Thus, when $N_1 = 2$ the feasible candidate values for N_2 are the "floor" and "ceiling" of 2.888, respectively; i.e., $\lfloor 2.888 \rfloor = 2$ *and* $\lceil 2.888 \rceil = 3$.

N_1:	1	2	3	4	5	6
N_2:	1.444	2.888	4.332	5.776	7.220	8.664
$\lfloor N_2 \rfloor$:	1	2	4	5	7	8
$\lceil N_2 \rceil$:	2	3	5	6	8	–

The next table examines each feasible integer point in the neighborhood of $\mathcal{L}$ and shows that the optimal round-lot solution is (5, 7) with the maximum performance value of $\theta_{RL} = 1.4355$. Note that this solution gives rise to a performance measure that is only 0.013 below the optimal $\theta_{\max} = 1.4357$ obtained without the round-lot requirements.

N_1	1	1	2	2	3	3
N_2	1	2	2	3	4	5
θ	1.4167	1.4214	1.4161	1.4354	1.4347	1.4328

N_1	4	4	**5**	5	6
N_2	5	6	**7**	8	8
θ	1.4326	1.4354	**1.4355**	1.4342	1.4347

This approach for generating the optimal round-lot solution is reasonably easy to implement for simple problems with only a few securities. However, as the number of securities increases it can quickly lead to a very time consuming search effort due to the increase in the number of feasible integer points neighboring the iso-performance line $\mathcal{L}$. We now describe an alternative heuristic method that locates a near-optimal solution without the attendant difficulties of the previous approach. Instead of examining *all* points in the neighborhood of $\mathcal{L}$, we only evaluate the integer points *surrounding* the optimal solution found without the round-lot requirements. Since the maximum value θ_{max} of the portfolio measure is an upper bound to any round-lot solution, one can easily measure the quality of a round-lot solution found by the heuristic by comparing its θ value to the upper bound θ_{max}.

For the case of two-securities presented above, and under the originally stated assumption that the investor is to allocate *as much as possible* his/her initial wealth to the two risky securities, the approximately optimal round-lot solution must be among the lattice points that are in the vicinity of $(N_1, N_2) = (5.8077, 8.3844)$. The lattice points in question, therefore, are the four corners of the unit square, at (5, 8), (5, 9), (6, 8), and (6, 9), that enclose the point (5.8077, 8.3844), as shown in Fig. 2. Although (6, 9) obviously violates the budget constraint and the feasibility of the lattice points (5, 9) and (6, 8) has yet to be verified, searching for the round-lot solution among these four candidates is by far more efficient than among all feasible lattice points in the above triangle.

The search results for the highest θ among these specific lattice points are summarized below. Although the lattice point (6, 9) violates the budget constraint, it is included here to reveal the achievable θ if more investment funds are available.

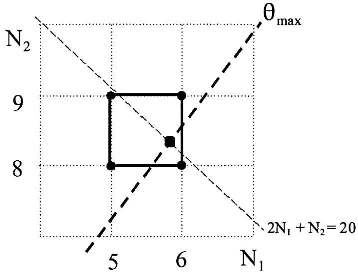


Fig. 2. The lattice points (5, 8), (5, 9), (6, 8) and (6, 9) enclose the point $(N_1, N_2) = (5.8077, 8.3844)$. The approximately optimal round-lot solution is at (6, 8).

N_1	N_2	θ	Amount Invested	Feasible?	Optimal?
5	8	1.4342	18000	Yes	No
5	9	1.4289	19000	Yes	No
6	**8**	**1.4347**	**20000**	**Yes**	**Yes**
6	9	1.4354	21000	No	–

The lattice point (6, 8), for which $\theta = 1.4347$, corresponds to the round-lot solution generated by the heuristic. Notice that the lattice point (3, 4), which is on the line joining the origin and (6, 8) and thus is on the iso-performance locus for $\theta = 1.4347$, is also a round-lot solution unless it is assumed (as the case here) that the investor intends to allocate as much as possible his/her initial wealth in risky securities. In the case of (3, 4), only half of the initial wealth is allocated to the two risky securities. Finally, it is worth noting that the performance value of 1.4347 obtained with this method is only 0.069% below the upper bound $\theta_{max} = 1.4357$.

The same approach can easily be extended to $n > 2$. Geometrically, in the case of $n = 3$ without round-lot requirements, the family of iso-performance loci in the (N_1, N_2, N_3) space can be described as a nest of cone-shaped surfaces with a common vertex at the origin. The plane of budget constraint (or, simply, the budget plane) shows a cross section of these cones. Each iso-performance contour on the budget plane is enclosed by other iso-performance contours with successively lower values of θ. As optimality is reached, the corresponding iso-performance contour on the plane reduces to a point. Each point on the budget plane being a feasible combination of N_1, N_2, and N_3, it can be visualized that a three-dimensional graph of the values of θ versus the corresponding locations on the budget plane has a single maximum of θ as the solution without round-lot requirements. Also, since the function θ is homogeneous of degree zero, the line from the origin to the optimal point on the budget plane in the (N_1, N_2, N_3) space is an iso-performance locus for the highest achievable value of θ. Any plane parallel to the budget plane has a proportionally scaled display of the same pattern of iso-performance contours.

All feasible allocations of the initial wealth with round-lot holdings of risky securities can be captured in the (N_1, N_2, N_3) space by the lattice points in a tetrahedron as defined by the three Cartesian planes and the budget plane. Given the above geometric pattern of the iso-performance loci, if the investor intends to allocate as much as possible his/her initial wealth in risky securities, the near-optimal round-lot solution must be located at one of the $2^3 = 8$ corners of the unit cube as defined by the lattice points enclosing the point corresponding to the solution without round-lot requirements. The search for the round-lot solution, therefore, can be performed efficiently by evaluating each of the corner of this cube. Of course, the corners that are most distant from the origin are most likely to be infeasible because of their violation of the budget constraint. The round-lot solution is captured by the corner corresponding to the highest achievable θ without violating the budget constraint.

When there are $n \geq 4$ securities, the geometry of the algorithm becomes difficult to visualize as it is impossible to graph objects in four or higher dimensions. The analytical justification for the approach, however, remains the same. The 2^n corners of an n-dimensional unit-sized "hypercube" bracketing the optimal solution without round-lot requirements can still be examined algebraically. Very simply, if the solution without round-lot requirements has a holding of N_i units in security i, the only integer values to be attempted for this security are the pair of consecutive integers bracketing N_i. Then, by comparing the 2^n sets of integer values of $N_1, N_2, \ldots, N_n$, the feasible case with the highest achievable value of θ can easily be identified. As stated formally in the following Proposition, this simple computational method allows us to solve the portfolio selection problem with round-lot requirements without resorting to sophisticated integer-programming tools.

Proposition 1: As the objective function θ of the portfolio selection problem based on n risky securities is concave, a near-optimal round-lot solution can be found by examining the 2^n corners of the n-dimensional unit-sized hypercube, with integer coordinates, bracketing the solution without round-lot requirements.

Having established a simple computational method, we now turn our attention to a covariance structure that allows us to explore the economic rationale underlying the portfolio selection. Specifically, in the next two sections we implement the hypercube method for the single-index model. We first consider the case where short-sales are disallowed. Next, by treating the short sale of a security under institutional procedures as an investment in an artificially constructed security, we extend the analysis to allow short sales. In order to facilitate the attainment of the round-lot solution, the problem formulation here differs from the more conventional cases in that the input parameters here are in prices rather than in returns. Nevertheless, as will be shown in the next two sections, the underlying economic rationale of the portfolio decision in Elton, Gruber and Padberg (1976, 1978) is retained.

4. THE NO-SHORT-SALE CASE UNDER THE SINGLE INDEX MODEL

In a single-index setting the random return R_i from investment in security i is generated by a linear model

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i, \quad \text{for } i = 1,2, \ldots, n, \tag{6}$$

where R_m is the random return on the market index (labeled as m), α_i and β_i are parameters, and ε_i is random noise. The usual assumptions are $E(\varepsilon_i) = 0$, $\text{Cov}(R_m, \varepsilon_i) = 0$, and $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, for $i, j = 1, 2, \ldots, n$, and $i \neq j$. Letting $\sigma_i^2 = \sigma_{ii} = \text{Var}(R_i)$, $\sigma_{ij} = \text{Cov}(R_i, R_j)$, $\sigma_m^2 = \text{Var}(R_m)$, and $\sigma_{\varepsilon_i}^2 = \text{Var}(\varepsilon_\text{i})$, the covariance structure of returns under these assumptions can be characterized as:

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{\varepsilon_i}^2$$

and

$$\sigma_{ij} = \beta_i \beta_j \sigma_m^2, \quad \text{for } i, j = 1, 2, \ldots, n \text{ and } i \neq j.$$

In terms of the random terminal and known initial prices for security i and the market m, Eq. (6) becomes

$$\frac{P_i - P_{0i}}{P_{0i}} = \alpha_i + \beta_i \left(\frac{P_m - P_{0m}}{P_{0m}} \right) + \varepsilon_i,$$

or, equivalently,

$$P_i = a_i + b_i P_m + e_i, \quad \{\text{for } i = 1, 2, \ldots, n,$$

where $a_i = (1 + \alpha_i - \beta_i)P_{0i}$, $b_i = \beta_i P_{0i}/P_{0m}$, and $e_i = P_{0i}\varepsilon_i$, with the properties that $E(e_i) = 0$, $\text{Cov}(P_m, e_i) = 0$, and $\text{Cov}(e_i, e_j) = 0$, for $i, j = 1,2, \ldots, n$, and $i \neq j$. Letting $s_{ii} = s_i^2 = \text{Var}(P_i)$, $s_{ij} = \text{Cov}(P_i, P_j)$, $s_m^2 = \text{Var}(P_m)$, and $s_{e_i}^2 = \text{Var}(e_i)$, the corresponding covariance structure can be characterized as:

$$s_i^2 = b_i^2 s_m^2 + s_{e_i}^2$$

and

$$s_{ij} = b_i b_j s_m^2, \quad \text{for } i, j = 1,2, \ldots, n, \text{ and } i \neq j.$$

Given the algebraic equivalence of the above two versions of the covariance structure, the optimality conditions in terms of security returns, as derived in Elton, Gruber and Padberg (1976, 1978), can easily be rewritten in terms of security prices. The optimality conditions in the current setting without round-lot requirements, therefore, are as follows:

$$Z_i = \frac{b_i}{s_{e_i}^2}\left[\frac{\bar{P}_i - (1+r)P_{0i}}{b_i} - C^*\right] + \frac{\delta_i}{s_{e_i}^2}, \tag{7}$$

$$Z_i \geq 0,\ \delta_i \geq 0, \text{ and } Z_i\delta_i = 0, \quad \text{for } i = 1,2,\ldots,n, \tag{8}$$

where each δ_i is a slack variable, and C^* is the cutoff rate of security performance. Each variable Z_i is related to the portfolio holdings $N_1, N_2, \ldots, N_n$ in the following manner:

$$Z_i = N_i\left\{\frac{\sum_{j=1}^{n} N_j[\bar{P}_i - (1+r)P_{0i}]}{\sum_{j=1}^{n}\sum_{k=1}^{n} N_j N_k s_{jk}}\right\}, \quad \text{for } i = 1,2,\ldots,n. \tag{9}$$

Letting L represent the set of selected securities (i.e. the set of securities with positive holdings in the portfolio) and defining

$$A_i \equiv \frac{s_m^2 b_i^2}{s_{e_i}^2} > 0 \tag{10}$$

and

$$t_i \equiv \frac{\bar{P}_i - (1+r)P_{0i}}{b_i}, \quad \text{for } i = 1, 2, \ldots, n, \tag{11}$$

for notational simplicity, an explicit expression of the cutoff rate of security performance is

$$C^* = s_m^2 \sum_{j\in L} b_j Z_j = \frac{\sum_{j\in L} t_j A_j}{1 + \sum_{j\in L} A_j}. \tag{12}$$

Combining $\sum_{i=1}^{n} N_i P_{0i} = W_0$ and the definition of each Z_i in Eq. (9), it follows that

$$N_i = \frac{Z_i W_0}{\sum_{j\in L} Z_j P_{0j}}, \quad \text{for } i = 1, 2, \ldots, n.$$

Following the same approach as in Elton, Gruber and Padberg (1976, 1978), as long as $b_i > 0$, for $i = 1, 2, \ldots, n$, the determination of set L can be achieved by first ranking and labeling the n securities such that $t_1 \geq t_2 \geq \cdots \geq t_n$. For each portfolio consisting of the h highest ranking securities (i.e. securities $1, 2, \ldots, h$), for $h = 1, 2, \ldots, n$, let

$$C_h = \frac{\sum_{j=1}^{h} t_j A_j}{1 + \sum_{j=1}^{h} A_j}.$$

A comparison between t_h and C_h, for $h = 1, 2, \ldots, n$, will reveal the lowest ranking security h in the optimal portfolio without round-lot requirements. That is, if $t_h > C_h$ (implying that $t_1 \geq t_2 \geq \cdots \geq t_h > C_h$) and $t_{h+1} \leq C_{h+1}$, set L consists of securities $1, 2, \ldots, h$. This C_h is the same as the cutoff rate of security performance C^* in Eq. (12), satisfying the conditions $t_1 \geq t_2 \geq \cdots \geq t_h > C_h \geq t_{h+1} \geq \cdots \geq t_n$. Then, according to Eq. (7), we have $Z_i = (t_i - C_h) b_i / s_{e_i}^2 > 0$, for $i = 1, 2, \ldots, h$, and $Z_i = 0$, for $i = h + 1, h + 2, \ldots, n$.

While each security i, for $i = 1, 2, \ldots, n$, is ranked according to its excess-return-to-beta ratio $(\bar{R}_i - r)/\beta_i$ – a measure of security performance – in Elton, Gruber and Padberg (1976, 1978), Eq. (7) provides an equivalent ranking hierarchy according to t_i, the ratio of expected excess price $[\bar{P}_i - (1 + r)P_{0i}]$ to systematic risk b_i, which is the same as $P_{0m}(\bar{R}_i - r)/\beta_i$. Also, C^* is the same as P_{0m} times the cutoff rate of security performance in Elton, Gruber and Padberg (1976, 1978). With P_{0m} being a multiplicative constant, the same ranking hierarchy of securities is maintained regardless of which of the two measures of security performance is used. With Z_i and N_i, for $i = 1, 2, \ldots, n$, successively known, the round-lot solution can be reached via the hypercube method. An illustrative example is provided below.

In view of the correspondence between the model formulation in Elton, Gruber and Padberg (1976, 1978) and that in the present setting, we translate the return data in Table 9.1 of Elton and Gruber (1995, Ch. 9) to corresponding data in terms of prices. Specifically, for an investment capital of $W_0 = \$100000$, consider a set of $n = 10$ securities with $r = 5\%$, $s_m^2 = P_{0m}^2 \sigma_m^2 = 64000$ (in $\$^2$), and the following data:

i	$\bar{P}_i$ (in \$)	P_{0i} (in \$)	b_i	$s^2_{e_i}$ (in $\$^2$)
1	2875	2500	0.3125	31250
2	3510	3000	0.5625	36000
3	2240	2000	0.2500	8000
4	1755	1500	0.3750	2250
5	4440	4000	0.5000	64000
6	3330	3000	0.5625	27000
7	2220	2000	0.5000	16000
8	2675	2500	0.2500	10000
9	3210	3000	0.3750	18000
10	1056	1000	0.0750	600

Here, the 10 securities have been ranked and labeled in such a way that $t_1 \geq t_2 \geq \cdots \geq t_{10}$. With $\bar{R}_i = (\bar{P}_i - P_{0i})/P_{0i}$, $\beta_i = b_i P_{0m}/P_{0i}$, and $\sigma^2_{\varepsilon_i} = s^2_{e_i}/P^2_{0i}$, for $i = 1, 2, \ldots, 10$, along with $P_{0m} = \$8000$ and $\sigma_m = \sqrt{10}\%$, the same input data as in Elton and Gruber (1995, Ch. 9) can be reproduced. [Notice that, in Elton and Gruber (1995, Ch. 9), all return data are provided implicitly in percentage terms. For example, the value of $\bar{R}_1 = 15$ with % being its unit implies that $\bar{R}_1 = 15\% = 0.15$ and the value of $\sigma^2_m = 10$ with $(\%)^2$ being its unit implies that $\sigma^2_m = 10 \times (\%)^2 = \dfrac{10}{(100)(100)} = \dfrac{1}{1{,}000}$ or, equivalently, $\sigma_m = \sqrt{10}\% = 0.03162$. Therefore, to translate properly the return data to the corresponding price data, the return data must first be stated in their proper decimal forms.]

In the present setting, the computed values of t_h and C_h, for $h = 1, 2, \ldots, 10$, are as follows:

h	t_h (in \$)	C_h (in \$)
1	800	133.333
2	640	295.035
3	560	353.591
4	480	434.333
5	480	436.085
6	320	424.097
7	240	401.817
8	200	392.497
9	160	379.811
10	80	361.384

Since $t_1 > t_2 > \cdots > t_5 > C_5$ and $t_6 < C_6$, set L consists of securities 1, 2, ..., 5, with the corresponding cutoff rate of security performance $C^* = C_5$. Then, the optimal portfolio holdings without round-lot requirements, corresponding to $\theta_{max} = 1.9997$, are as follows:

$$\begin{array}{cccccccccc} N_1 & N_2 & N_3 & N_4 & N_5 & N_6 & N_7 & N_8 & N_9 & N_{10} \\ \hline 9.391 & 8.222 & 9.993 & 18.887 & 0.885 & 0 & 0 & 0 & 0 & 0 \end{array} \tag{13}$$

The round-lot solution generated by the heuristic is the one, among the following $2^{10} = 1024$ corners of a hypercube bracketing the above point in a 10-dimensional space, that corresponds to the highest achievable value of θ:

$$\begin{array}{cccccccccc} N_1 & N_2 & N_3 & N_4 & N_5 & N_6 & N_7 & N_8 & N_9 & N_{10} \\ \hline 9 \text{ or } 10 & 8 \text{ or } 9 & 9 \text{ or } 10 & 18 \text{ or } 19 & 0 \text{ or } 1 & 0 \text{ or } 1 & 0 \text{ or } 1 & 0 \text{ or } 1 & 0 \text{ or } 1 & 0 \text{ or } 1 \end{array} \tag{14}$$

To implement the hypercube method efficiently, we first check the feasibility of the individual corners. If a corner is feasible (i.e. if it satisfies the budget constraint), then we evaluate the corresponding value of θ for that corner. If the corresponding value of θ is inferior to the best one found so far, we eliminate the corner from any further consideration. If this θ is better than the highest value found so far, we keep it and proceed to the next corner. Once the examination of all the corners is completed, the heuristic identifies the integer round-lot solution.

To facilitate further the computational efficiency in the search of the round-lot solution, we express the objective function in Eq. (2) in matrix notation as $\theta = \mathbf{N}'_L \boldsymbol{\mu}_L (\mathbf{N}'_L \mathbf{v}_L \mathbf{N}_L)^{-1/2}$, where for an n-security case

$$\boldsymbol{\mu}_L = \begin{bmatrix} \bar{P}_1 - (1+r)P_{01} \\ \vdots \\ \bar{P}_n - (1+r)P_{0n} \end{bmatrix} \quad \text{and} \quad \mathbf{N}_L = \begin{bmatrix} N_1 \\ \vdots \\ N_n \end{bmatrix}$$

are n-dimensional column vectors (with $\mathbf{N}'_L$ denoting the transpose of $\mathbf{N}_L$), and

$$\mathbf{v}_\mathrm{L} = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix} \tag{15}$$

is an $n \times n$ matrix. In the present example of $n = 10$, since $\boldsymbol{\mu}_L$ and $\mathbf{v}_L$ can be computed directly from the input data, the search for the optimal $\mathbf{N}_L$ is among the 2^{10} possible vectors as indicated in (14). This can easily be achieved by nesting 10 loops in any standard computer programming language. (We wrote a Maple program for this illustrative example. The total CPU time as required on a Pentium 200 personal computer was about 24 seconds. The program is available from the authors upon request.) The round-lot holdings in the present example, corresponding to $\theta = 1.9996$ and $W_0 = 97500$, are:

$$\begin{array}{cccccccccc} N_1 & N_2 & N_3 & N_4 & N_5 & N_6 & N_7 & N_8 & N_9 & N_{10} \\ \hline 9 & 8 & 10 & 18 & 1 & 0 & 0 & 0 & 0 & 0 \end{array} \tag{16}$$

Note that the $\theta = 1.9996$ value found by the hypercube heuristic is only 0.005% lower than the upper bound of $\theta_{\max} = 1.9997$.

We now show that there is an analytical equivalence between the round-lot solution and the solution for the same portfolio selection problem without round-lot requirements that is based on a slightly different set of input data. With this analytical equivalence, the economic rationale for security selection that is concealed by the hypercube algorithm for the original portfolio selection problem will become more noticeable. In order to establish this analytical equivalence, notice that in the above example, if there are any changes in the input data in the portfolio selection problem without round-lot requirements, the numerical solution in (13) will be affected. Suppose that, for instance, the only change in the input data is that $\bar{P}_1$ becomes slightly less than \$2875, the original figure. Intuitively, given that security 1 is now less attractive, the optimal holding of this security will be less than what is shown in (13). Further, the accompanying redistribution of the unused portion of the investment funds originally intended for this security will cause some increases in the individual holdings of the remaining securities in the portfolio.

To generalize this intuitive idea, suppose instead that in the same example there are minor changes – including positive and negative changes – in all the expected prices, labeled as $\Delta\bar{P}_i$, for $i = 1, 2, \ldots, 10$. The sign of change in the holding of each security due to the redistribution of investment funds will be less obvious. However, for $W_0 = \$97500$, there must be some $\Delta\bar{P}_1, \Delta\bar{P}_2, \ldots, \Delta\bar{P}_{10}$ that allow the solution without round-lot requirements to match exactly the round-lot solution in (16). As shown below, the required values of $\Delta\bar{P}_1, \Delta\bar{P}_2, \ldots, \Delta\bar{P}_5$ in this example can easily be determined by first writing Eq. (7) as

$$Z_i = \frac{b_i}{s_{e_i}^2}\left[\frac{\bar{P}_i + \Delta\bar{P}_i - (1+r)P_{0i}}{b_i} - C^*\right], \quad \text{for } i \in L, \tag{17}$$

$$= 0, \quad \text{for } i \notin L,$$

where C^* is also given by Eq. (12), except that each t_j there is now defined as $[\bar{P}_j + \Delta\bar{P}_j - (1+r)P_{0j}]/b_j$ instead. Since

$$Z_i = \frac{N_i\theta}{\sqrt{\sum_{j \in L}\sum_{k \in L} N_j N_k s_{jk}}}, \quad \text{for } i = 1, 2, \ldots, n,$$

the result in (16), along with the corresponding $\theta = 1.9996$, will allow each Z_i, for $i = 1, 2, \ldots, 10$, to be determined. Then, with $C^*\left[= s_m^2 \sum_{j \in L} b_j Z_j\right.$ according to Eq. (12)$\left.\right]$ subsequently known, the $\Delta\bar{P}_i$ term in Eq. (17) for each $i \in L$ (i.e. $i = 1,2, \ldots, 5$) can easily be solved. For $i \notin L$ (i.e. $i = 6, 7, \ldots, 10$), however, the $\Delta\bar{P}_i$ term is not unique; as long as $[\bar{P}_i + \Delta\bar{P}_i - (1+r)P_{0i}]/b_i \leq C^*$, it follows that $Z_i = 0$. That is, to violate the optimality conditions in (7) and (8), the $\Delta\bar{P}_i$ term for any $i \notin L$ must be high enough for that security to be deemed attractive for inclusion in the portfolio. Therefore, for the portfolio selection problem without round-lot requirements to have the same security holdings as in (16), the required adjustments to the expected prices are as follows:

$\Delta\bar{P}_1$	$\Delta\bar{P}_2$	$\Delta\bar{P}_3$	$\Delta\bar{P}_4$	$\Delta\bar{P}_5$	$\Delta\bar{P}_6$	$\Delta\bar{P}_7$	$\Delta\bar{P}_8$	$\Delta\bar{P}_9$	$\Delta\bar{P}_{10}$
–1.35	0.39	1.00	–0.25	3.65	65.35 (Max)	98.09 (Max)	59.05 (Max)	103.57 (Max)	26.71 (Max)

(18)

For the five securities that are excluded from the portfolio, the highest possible adjustments without affecting the results in (16) are shown.

The above idea of adjusting the expected prices draws from the same analytical treatment of portfolio selection with investment limits as in Kwan and Yip (1987). For instance, while $N_1 = 9.391$ in the same example is the desired holding of security 1, it has been reduced to 9 lots to satisfy a specific round-lot requirement. This can be viewed as the outcome of specifically imposing a 9-lot limit on security 1. As shown analytically in Kwan and Yip (1987) the optimality conditions for the portfolio selection problem with upper bounds on individual securities differ from the corresponding case without such limits in that there is an additive adjustment term to the expected return for each security. When formulated in terms of prices rather than returns as in the present setting, the optimality conditions satisfying various upper and lower investment limits can still be written as those in (7) and (8), but with each $\bar{P}_i$ replaced by $\bar{P}_i + \Delta\bar{P}_i$, for $i = 1, 2, \ldots, n$, as indicated intuitively above. (Of course, the round-lot holdings of $N_1 = 9$, $N_2 = 8$, etc., in the example are initially unknown. Therefore, it is impossible to set up correctly a portfolio selection problem for some predetermined upper or lower bounds on security holdings without first obtaining the round-lot solution. Rather, the purpose here is to establish an analytical equivalence between the corresponding solutions with and without round-lot requirements.)

The presence of the $\Delta\bar{P}_i$ terms in Eq. (17) allows the round-lot solution in (16) to accommodate different price expectations for each security in the input data. Since reliable forecasts of security prices (or returns) are much more difficult to achieve than reliable forecasts of risk parameters, it is useful to examine whether the portfolio solution is robust given this difficulty. The signs and the magnitudes of the individual $\Delta\bar{P}_i$ terms can provide useful information regarding the sensitivity of the portfolio results in response to changes in individual price expectations. The most striking numbers in (18) pertain to the five excluded securities; some very large positive adjustments in price expectations are required for any of these securities to reach the threshold performance level. For example, in the absence of any other changes to the input data, $\bar{P}_6$ has to increase from \$3330 to beyond \$3395.35 to make security 6 attractive enough for inclusion in the portfolio.

5. THE SHORT-SALE CASE UNDER THE SINGLE INDEX MODEL

We now extend the above analysis to allow short-sale transactions under institutional procedures. As described in Section 2, the short sale of security i, for $i = 1, 2, \ldots, n$, is analytically equivalent to the purchase of an artificially constructed security $n + i$. Therefore, under the covariance structure for securities $1, 2, \ldots, n$ as characterized by the single index model, the optimality conditions without round-lot requirements are analogous to those of the no-short-sale case. Following Kwan (1995), but in a formulation where the input data for the analysis are in prices rather than returns, we obtain

$$Z_i = \frac{b_i}{s_{e_i}^2}\left[\frac{\bar{P}_i - (1+r)P_{0i}}{b_i} - C^*\right] + \frac{\delta_i}{s_{e_i}^2}, \tag{19}$$

$$Z_{n+i} = -\frac{b_i}{s_{e_i}^2}\left[\frac{\bar{P}_i - (1+\nu r)P_{0i}}{b_i} - C^*\right] + \frac{\delta_{n+i}}{s_{e_i}^2}, \tag{20}$$

$$Z_i \geq 0,\ \delta_i \geq 0,\ Z_{n+i} \geq 0,\ \delta_{n+i} \geq 0,$$

$$Z_i\delta_i = 0,\ Z_{n+i}\delta_{n+i} = 0, \text{ and } Z_iZ_{n+i} = 0, \quad \text{for } i = 1, 2, \ldots, n. \tag{21}$$

Here, for $i = 1, 2, \ldots, 2n$, each δ_i is a slack variable and each Z_i is related to the portfolio holdings in the following manner:

$$Z_i = N_i\left[\frac{E(W) - (1+r)W_0}{\mathrm{Var}(W)}\right], \tag{22}$$

where the bracketed terms are given by Eqs (4) and (5). We let L and S represent, respectively, the sets of securities among $i = 1, 2, \ldots, n$ and $i = n+1, n+2, \ldots, 2n$ that are selected for the portfolio. Defining

$$u_i \equiv \frac{\bar{P}_i - (1+\nu r)P_{0i}}{b_i}, \quad \text{for } n+i \in S, \tag{23}$$

for notational simplicity, we can express the cutoff rate of security performance as

$$C^* = s_m^2\left(\sum_{j\in L} b_jZ_j + \sum_{(n+j)\in S} b_jZ_j\right) = \frac{\sum_{j\in L} t_jA_j + \sum_{(n+j)\in S} u_jA_j}{1 + \sum_{j\in L} A_j + \sum_{(n+j)\in S} A_j},$$

where A_j and t_j are as defined in Eqs (10) and (11). Notice that, with u_i expressed equivalently as the ratio of $-\bar{P}_i + (1+\nu r)P_{0i}$ and $-b_i$, the economic rationale of using this ratio to rank securities in short positions becomes more obvious. Since the purchase of security $n+i$ requires a beginning-of-period investment of νP_{0i}, the expected end-of-period payoff in excess of what can be achieved in a risk-free investment of νP_{0i} is $(-\bar{P}_i + P_{0i}) - (1+r)\nu P_{0i} =$

$-\bar{P}_i + (1 + \nu r)P_{0i}$. Thus, keeping risk exposure constant, the lower the amount $\bar{P}_i - (1 + \nu r)P_{0i}$, the more attractive is security $n + i$ for holding. In the risk dimension, the investment in security $n + i$ is the same as investing in a negative-beta security; the corresponding systematic risk in the present setting is $-b_i$. As shown in Ben-Horim and Levy (1980), an investment in a negative beta security has a stabilizing, risk reducing effect on the portfolio. Thus, keeping the expected payoff from investing in security $n + i$ constant, the higher the positive value of b_i, the more attractive is security $n + i$ in a portfolio context. The ratio u_i nicely captures the expected payoff in excess of a risk-free outcome, adjusted for the systematic risk exposure; the lower the ratio, the more attractive is security $n + i$.

As long as L and S are known, the portfolio selection problem without round-lot requirements can easily be solved. The present task, therefore, is to determine the two sets. To this end, we follow the approach in Kwan (1995) to establish two alternative ranking hierarchies for securities $1, 2, \ldots, n$. We first rank and label the n securities in such a way that $t_1 \geq t_2 \geq \cdots \geq t_n$. This ranking hierarchy facilitates the selection of securities for purchases. For the selection of securities for short selling, however, we rank and label the same n securities as $[1], [2], \ldots, [n]$ in accordance with $u_{[1]} \leq u_{[2]} \leq \cdots \leq u_{[n]}$. The idea behind these two ranking hierarchies is that, if security h is selected for the portfolio in a long position, so are securities $1, 2, \ldots, h - 1$, and if security $[k]$ is held short, so are securities $[1], [2], \ldots, [k–1]$. The optimal portfolio without round-lot requirements can be constructed by successively adding securities to an initially empty portfolio following the two ranking hierarchies. With the parameter

$$C_{h,[k]} = \frac{\sum_{j=1}^{h} t_j A_j + \sum_{[j]=[1]}^{[k]} u_{[j]} A_{[j]}}{1 + \sum_{j=1}^{h} A_i + \sum_{[j]=[1]}^{[k]} A_{[j]}}, \text{ for } h = 1, 2, \ldots, n \text{ and } [k] = [1], [2], \ldots, [n],$$

calculated successively for each portfolio consisting of securities $1, 2, \ldots, h$ and $[1], [2], \ldots, [k]$, optimality is reached when $t_h > C_{h,[k]} \geq t_{h+1}$ and $u_{[k]} < C_{h,[k]} \leq u_{[k+1]}$. This specific $C_{h,[k]}$ corresponds to the cutoff rate of security performance, C^*, in Eqs (19) and (20), and the optimal L and S consist of securities $1, 2, \ldots, h$ and $[1], [2], \ldots, [k]$, respectively.

Having described how the optimal L and S are formed, we now return to the original labeling of securities (i.e. $1, 2, \ldots, 2n$) so that each security and its artificial counterpart can be matched more readily. For each security i $(= 1, 2, \ldots, 2n)$ that is selected to the portfolio, we have $Z_i > 0$ and $\delta_i = 0$

according to the optimality conditions in (19), (20), and (21); for each excluded security i, we have $Z_i = 0$ and $\delta_i \geq 0$ instead. In view of Eq. (22), the optimal portfolio holdings can be computed from

$$N_i = \frac{Z_i W_0}{\sum_{j=1}^{n} Z_j P_{0j} + \lambda \sum_{j=n+1}^{2n} Z_j P_{0j}}, \quad \text{for } i = 1, 2, \ldots, 2n.$$

These portfolio holdings, represented by a point in a $2n$-dimensional space, will allow the hypercube method to be implemented in the search of the round-lot solution. Notice that, since each security and its artificial counterpart cannot both be selected for the portfolio, not all 2^{2n} corners of the hypercube need to be attempted. Therefore, the computational burden is not as onerous as the number 2^{2n} may suggest.

We now provide an illustration of the short-sale case using the same example in Section 4. In addition to the input data as provided earlier, let $\nu = 0.5$ be the fraction of interest rebate and $\lambda = 0.2$ be the margin deposit. To solve the portfolio selection problem with short sales not subject to round-lot requirements, we can start with the no-short-sale case where securities $1, 2, \ldots, 5$ are held long. The artificially constructed securities, $11, 12, \ldots, 20$, are ranked according to Eq. (23) and relabeled as follows:

$n+i$	$[h]$	$u_{[h]}$ (in \$)
17	[1]	340
19	[2]	360
20	[3]	413
18	[4]	450
16	[5]	453.333
14	[6]	580
15	[7]	680
13	[8]	760
12	[9]	773.333
11	[10]	1000

With $C_5 = 436.085$ from the no-short-sale case, securities [1], [2], and [3] can successively be added to the portfolio, resulting in $C_{5,[1]} = 423.295$, $C_{5,[2]} = 419.345$, and $C_{5,[3]} = 418.926$, respectively, in accordance with Eq. (24). This $C_{5,[3]}$ corresponds to cutoff rate C^* because the security selection criteria $t_1 \geq t_2 \geq \cdots \geq t_5 \geq C_{5,[3]} \geq t_6 \geq \cdots \geq t_{10}$ and $u_{[1]} \leq u_{[2]} \leq u_{[3]} \leq C_{5,[3]} \leq u_{[4]} \leq \cdots \leq u_{[10]}$ are completely satisfied.

The resulting portfolio holdings, for an achievable $\theta_{max} = 2.0380$, are

N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
7.981	7.234	9.233	21.319	0.999	0	0	0	0	0
N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{19}	N_{20}
0	0	0	0	0	0	5.166	0	2.571	1.464

where $N_{17} = N_{[1]}$, $N_{19} = N_{[2]}$ and $N_{20} = N_{[3]}$. The positive holdings of securities 17, 19, and 20 are analytically equivalent to short selling securities 7, 9, and 10, respectively, in the numbers of lots as indicated. Given the above results, the search for the optimal θ among the individual corners of a 20-dimensional hypercube is similar to the case where short sales are not allowed. Since each security and its artificial counterpart cannot both be selected for the same portfolio, the corners to be evaluated are confined to the following coordinate positions in a 20-dimensional space:

N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
7 or 8	7 or 8	9 or 10	21 or 22	0 or 1	0 or 1	0	0 or 1	0	0
N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{19}	N_{20}
0	0	0	0	0 or 1	0 or 1	5 or 6	0 or 1	2 or 3	1 or 2

Further, there is no need to evaluate any case where $(N_i, N_{10+i}) = (1, 1)$, for $i = 5$, 6, and 8, for the same reason.

For computational efficiency, we first re-write Eqs (1), (4), and (5) in matrix notation. Specifically, we form two n-element column vectors of excess prices,

$$\boldsymbol{\mu}_L = \begin{bmatrix} \bar{P}_1 - (1+r)P_{01} \\ \vdots \\ \bar{P}_n - (1+r)P_{0n} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu}_S = \begin{bmatrix} \bar{P}_1 - (1+\nu r)P_{01} \\ \vdots \\ \bar{P}_n - (1+\nu r)P_{0n} \end{bmatrix},$$

and stack these vectors to form a new $2n$-element vector,

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_L \\ \boldsymbol{\mu}_S \end{bmatrix}.$$

With the $n \times n$ covariance matrix $\mathbf{v}_L$ for securities $1, 2, \ldots, n$ defined as in (15), we construct the augmented matrix $\mathbf{v}$ for the $2n$ securities as

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_L & \mathbf{v}_L \\ \mathbf{v}_L & \mathbf{v}_L \end{bmatrix}.$$

For the decision variables $N_1, N_2, \ldots, N_{2n}$, we define two n-element vectors,

$$\mathbf{N}_L = \begin{bmatrix} N_1 \\ \vdots \\ N_n \end{bmatrix} \quad \text{and} \quad \mathbf{N}_S = \begin{bmatrix} -N_{n+1} \\ \vdots \\ -N_{2n} \end{bmatrix},$$

and form the $2n$-element vector

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_L \\ \mathbf{N}_S \end{bmatrix}.$$

Then, the objective function can conveniently be computed as $\theta = \mathbf{N}'\boldsymbol{\mu}(\mathbf{N}'\mathbf{v}\mathbf{N})^{-1/2}$ (with $\mathbf{N}'$ denoting the transpose of $\mathbf{N}$) for the individual corners of the $2n$-dimensional hypercube.

In the above example, the round-lot long and short holdings, corresponding to $\theta = 2.0377$ (a performance value that is only 0.014% lower than the upper bound) and $W_0 = \$98500$, are as follows:

N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
8	7	9	21	1	0	0	0	0	0

N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{19}	N_{20}
0	0	0	0	0	0	5	0	3	1

These results indicate that, among the 10 securities for portfolio consideration, securities $1, 2, \ldots, 5$ are held long and securities 7, 9, and 10 are held short, in round-lot holdings as indicated. The analytical equivalence between the above round-lot solution with short sales and the corresponding solution without round-lot requirements (for a portfolio selection problem based on the same input data but with some revised price expectations) can also be established in the same manner as the no-short-sale case in Section 4. Once each $\bar{P}_i$ is substituted by $\bar{P}_i + \Delta\bar{P}_i$, for $i = 1, 2, \ldots, 10$ (and, in general, $i = 1, 2, \ldots, n$), the optimality conditions in (19), (20), and (21) for the portfolio selection problem without round-lot requirements are also applicable to the round-lot case. Therefore, the security selection criteria that are concealed by the hypercube algorithm now becomes more readily noticeable.

For the short-sale solution without round-lot requirements to have exactly the same security holdings as in (25), the required adjustments in expected prices, in the case of $W_0 = \$98500$, are

$\Delta\bar{P}_1$	$\Delta\bar{P}_2$	$\Delta\bar{P}_3$	$\Delta\bar{P}_4$	$\Delta\bar{P}_5$	$\Delta\bar{P}_6$	$\Delta\bar{P}_7$	$\Delta\bar{P}_8$	$\Delta\bar{P}_9$	$\Delta\bar{P}_{10}$
2.25	–2.51	–0.62	–0.56	–0.26	–20.41 to 54.59	–0.49	–8.24 to 54.26	–4.94	–0.01

As in the no-short-sale case, the sign and the magnitude of each $\Delta\bar{P}_i$, for $i = 1, 2, \ldots, 10$, can provide useful information regarding the sensitivity of the round-lot solution in response to changes in price expectations. The decision to exclude securities 6 and 8 from the portfolio is not likely to be reversed unless there are major changes to the input data. For example, assuming that there are no other changes in the input data, to make security 6 attractive enough for purchasing, its price expectation has to increase from \$3330 to beyond \$3384.59 (= \$3330 + \$54.59); alternatively, to make the same security attractive enough for short selling, a decrease of its price expectation from \$3330 to below \$3309.59 (= \$3330 − \$20.41) will be required.

6. SUMMARY AND CONCLUSION

The present study has extended the portfolio literature by considering a practical issue in portfolio selection. In particular, a round lot being the standard unit for stock trading, we have solved portfolio selection problems with round-lot requirements in settings where short sales are disallowed and allowed. Drawing on recent developments in the portfolio literature, we have followed the institutional procedures for short selling as required for portfolio analysis so that the short-sale benefit can be captured accurately. The translation of portfolio solutions which allow holdings of fractional shares into round-lot solutions without severely sacrificing the portfolio performance is a non-trivial issue for investors with relatively small investment capitals. By exploiting some analytical properties of the objective function in portfolio optimization, we have presented a hypercube method to search for the round-lot solutions.

With security holdings in the solution of each n-security portfolio selection problem without short sales and without round-lot requirements represented geometrically as a point in an n-dimensional space, the corresponding hypercube bracketing this point has unit lengths for all its edges and has integer coordinates for all its 2^n corners. We have shown that a near-optimal round-lot solution can be reached by evaluating these 2^n corners for portfolio performance, as well as budgetary feasibility. Since the short sale of a security can be viewed as the purchase of an artificially constructed security, a portfolio selection problem with short sales is analytically equivalent to a $2n$-security case without short sales. Therefore, the round-lot solution with short sales can be found among the corners of the corresponding $2n$-dimensional hypercube.

In order to provide the economic rationale for portfolio selection, we have characterized the covariance structure of security returns using the single index model. This characterization of the covariance structure has allowed us to establish the correspondence between the round-lot solution and the solution without round-lot requirements for which security selection criteria are available. It has also allowed us to examine the sensitivity of the round-lot solution in response to changes in price expectations. This is a useful feature considering that reliable forecasts of security prices (or returns) are much more difficult to achieve than reliable forecasts of risk parameters.

Since the usefulness of a normative portfolio model often hinges on its ability to capture reality in the investment world, as clearly indicated in recent developments in the portfolio literature, the present study has contributed in improving further the practical relevance of academic portfolio modeling. On the technical side, we have illustrated that using the hypercube method, high-quality round-lot solutions can be achieved without the encumbrance of any algorithmic complexities that are often associated with integer programming. This technical simplification will make it easier for the analysis to be implemented in practical settings. On the economic side, we have provided the rationale, based on risk-return considerations, as to why a security should be included in or excluded from the portfolio. Given these nice features, the analysis as presented in this study should enhance the usefulness of portfolio modeling for assisting practical investment decisions.

ACKNOWLEDGMENTS

Financial support for this research was provided by the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

Alexander, G. J. (1993). Short selling and efficient sets. *Journal of Finance*, *48*, 1497–1506.

Alexander, G. J. (1995). Efficient sets, short-selling, and estimation risk. *Journal of Portfolio Management*, (Winter), 64–73.

Ben-Horim, M., & Levy, H. (1980). Total risk, diversifiable risk and non-diversifiable risk: A pedagogic note. *Journal of Financial and Quantitative Analysis*, *15*, 289–297.

Elton, E. J., & Gruber, M. J. (1995). *Modern Portfolio Theory and Investment Analysis* (5th ed.). Wiley, New York.

Elton, E. J., Gruber, M. J., & Padberg, M. W. (1976). Simple criteria for optimal portfolio selection. *Journal of Finance*, *31*, 1341–1357.

Elton, E. J., Gruber, M. J., & Padberg, M. W. (1978). Optimal portfolios from simple ranking devices. *Journal of Portfolio Management*, *4* (Spring), 15–19.

Frankfurter, G. M. (1990). Is normative portfolio theory dead? *Journal of Economics and Business*, *42*, 95–98.

Haugen, R. A. (1997). Modern Investment Theory (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Kwan, C. C. Y. (1995). Optimal portfolio selection under institutional procedures for short selling. *Journal of Banking and Finance*, *19*, 871–889.

Kwan, C. C. Y. (1997). Portfolio selection under institutional procedures for short selling: Normative and market-equilibrium considerations. *Journal of Banking and Finance*, *21*, 369–391.

Kwan, C. C. Y. (1999). A note on market-neutral portfolio selection. *Journal of Banking and Finance*, *23*, 773–799.

Kwan, C. C. Y., & Yip, P. C. Y. (1987). Optimal portfolio selection with upper bounds for individual securities. *Decision Sciences*, *18*, 505–523.

Kwan, C. C. Y., & Yuan, Y. (1993). Optimal portfolio selection without short sales under the full-information covariance structure. *Journal of Economics and Business*, *45*, 91–98.

Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, *3*, 111–133.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New Haven, CT: Yale University Press.

Markowitz, H. M. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. New York: Basil Blackwell.

This Page Intentionally Left Blank

DEFINING A SECURITY MARKET LINE FOR DEBT EXPLICITLY CONSIDERING THE RISK OF DEFAULT

Jean L. Heck, Michael M. Holland and David R. Shaffer

ABSTRACT

While a major consequence of the use of debt by a business is generally assumed to be a change to the risk of default, theoretical work relating this risk to the lender's required rate of return is notably sparse. This paper defines an equilibrium model to value debt given a non-zero probability of default by extending previous research and then formulates the corresponding appropriate security market line. Also, a model to value debt is synthesized that compensates a lender for both capital market risk and default risk.

INTRODUCTION

As demonstrated in the popular press, excessive use of debt by businesses would seem to be a major contributor to the eventual failure of many of those businesses. Accounting for the risk of financial loss associated with the use of debt would seem to be integral in such areas as analysis of optimal capital structure, loan portfolio management, and capital market equilibrium. However, there has been a dearth of work examining the theoretical linkages between the probability of default (POD), the risk of financial loss associated

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 165–179.

ISBN: 0-7623-0887-7

with default, and a mechanism to determine an appropriate interest rate to compensate an investor for taking these risks.

Early seminal works, such as Modigliani and Miller (1963), Hamada (1969) and Rubenstein (1973), and most derivative works, took an expedient path by assuming risk-free borrowing. Others, notably Haugen and Senbet (1979, 1988) and Bierman and Oldfield (1979) have presented compelling arguments in support of this assumption by maintaining that, in an efficient market, agency costs associated with bankruptcy will be virtually eliminated and that the risk of loss associated with default and bankruptcy will be essentially zero. Work by Altman (1984) and other empirical evidence has suggested that the occurrence of bankruptcy may result in many unavoidable costs to the business that must ultimately be borne by the investors; i.e. as long as there are lawyers and accountants, bankruptcy will not be costless. Adequate evidence exists to support the notion that risk of default may be a significant investor risk and that this risk has not been effectively accounted for by earlier research.

Two early papers related to quantifying the risk of default (Bierman & Hausman, 1970; Dirickx & Wakeman, 1976) first proffered an approach to linking default, collateral and interest rates. However, the focus in these works was more on the extension of trade credit by a business. The most definitive work on the relationship between risk premiums and risk of default, and work that parallels the methodologies of this paper is that of Bierman and Hass (1975), Yawitz (1977) and Fons (1987) who examine default premiums in the bond market.

Besides the obvious relationship between the risk of default and investor compensation, an equally important concern is the quantification of capital market risk associated with debt and the setting of an appropriate rate of return to take that risk. In this vein, Callahan and Mohr (1989) synthesized the determinants of systematic risk and the capital asset pricing model. Other significant studies include those of Bierman and Oldfield and an extension by Conine (1980) attempting to define debt risk and to then link that risk to security valuation within the context of the CAPM.

In the next section an alternative methodology is specified to extend the earlier work of Fons (1987) regarding the quantification of a borrower-specific risk of default, where the risk of default refers to the possibility of suffering financial loss given default, while the probability of default refers to the specific probability of default occurring. In the third section, Fons' work is extended by formulating a model that compensates a lender for both default risk and capital market risk – which will lead to the development of a security market line for debt. Section four applies the model to a market traded bond and examines the performance of the model as key parameters are varied. The final section offers a summary of the paper's findings.

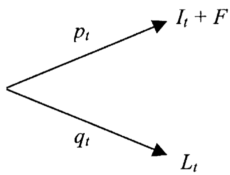
VALUATION OF DEBT INVESTMENTS GIVEN RISK OF DEFAULT

A debt investment (loan) is defined as returning a predefined amount of principal and interest to the investor through periodic payments that terminate after some specified number of payments. It is assumed that default on a payment would terminate the payment stream and lead inexorably to the liquidation of any assets of the business (collateral), whether or not formal bankruptcy procedures are invoked. The key to understanding the solution methodology is to recognize that the repayment schedule is an *n*-payment stream of income that may be truncated at some random time offering a single, alternative payment. From this, evaluation of the borrower-specific risk of a loan subject to a POD can be viewed as being composed of three fundamental steps: (1) an estimation of the POD; (2) a forecast of the proceeds that would be realized in the event collateral had to be liquidated; and (3) defining the correct repayment schedule.

Much work has been done on both techniques suitable to forecast a POD and the accuracy of such forecasts (Altman, 1983, 1982, 1968; Robertson, 1985; Mensah, 1984; Lincoln, 1984; Srinivasan, 1984; Cowen & Page, 1982; Ohlson, 1980; Vinso, 1979). It is assumed that current appraisal techniques allow a lender to forecast the liquidation value of the collateral of the business with reasonable accuracy. What has remained relatively ill-defined in the literature, save for Fons (1987), is a quantitative methodology linking these two elements to the setting of the amount of interest to charge to compensate for the risk of default.

One-Period Loan

The amount of the original loan, the term of the loan, and the timing of the payments would be generally defined by the needs of the borrower. The liquidation value of the collateral and the POD would then be estimated by the lender. Thus, all of the elements of the loan are known except the specific amounts of the interim and final payments; i.e. the repayment schedule. In fact, this development begins by making no assumptions regarding the relationship between the interim and final payments. At its simplest, the process is represented by the decision tree below:



where,

t = the period in which various events occur for a one-period loan, $t = 1$.
I_t = the interim payment made at the end of time period t.
F = the final, or balloon payment made at the end of time period n.
L_t = the funds obtained from liquidation of the collateral at the end of time period t, as the result of a default during this time period, net of associated bankruptcy costs.
p_t = the probability of receiving the contracted payments at the end of period t.
q_t = the POD during the interval, t, and thus the probability of not receiving the contracted payment at the end of the interval; $q_t = 1 - p_t$.

The next step is to reconcile the time disparity in the flow of funds. If lenders chose to make a loan free of default risk, they could invest their capital in some otherwise comparable alternative offering a rate of return of k percent over the interval. Finally, dropping all the subscripts, the expected value, EV, of the risky one-period loan – stated in present value terms – would appear as:

$$EV = \frac{qL + (1 - q)(I + F)}{1 + k} \tag{1}$$

By defining a present value term to be $U = (1 + k)^{-1}$, Eq. (1) can be rewritten in a more generalized form as:

$$EV = qLU + (1 - q)IU + (1 - q)FU \tag{2}$$

Clearly, there is some combination of the interim and final payments that would make the lender indifferent between the risky loan and the default-free alternative. At its simplest, this would occur when the expected present value of the risky loan was equal to the amount lent.

Multi-Period Loans

Extending this logic next to a two-period loan, the expected value equation becomes:

$$D = (1 - q_1)I_1U + q_1L_1U + (1 - q_1)q_2L_2U^2 + (1 - q_1)(1 - q_2)I_2U^2 + (1 - q_1)(1 - q_2)FU^2 \tag{3}$$

Under this rationale, D would be the market value of a 2-payment stream of cash flows subject to risk of default.

It can be seen by reflecting on Eqs (2) and (3) that the general form is composed of three basic terms; a collateral term (CT), an interim payment term (IPT) and the final payment term (FPT); i.e. $D = CT + IPT + FPT$. For the two-period loan, the terms would be:

$$CT = q_1L_1U + (1 - q_1)q_2L_2U^2$$
$$IPT = (1 - q1)I_1U + (1 - q_1)(1 - q_2)I_2U^2$$
$$FPT = (1 - q_1)(1 - q_2)FU^2$$

Finally, for an n-payment loan, the terms can be shown to have the following general form:

$$D = CT + IPT + FPT \tag{4}$$

where,

$$CT = \sum_{t=1}^{n} q_tL_tU^t \prod_{j=0}^{t-1} (1 - q_j) \tag{4a}$$

$$IPT = \sum_{t=1}^{n} q_tI_tU^t \prod_{j=1}^{t} (1 - q_j) \tag{4b}$$

$$FPT = FU^n \prod_{j=1}^{t} (1 - q_j) \tag{4c}$$

Here, $(1 - q_0) = 1$, $U^t = (1 + k)^{-t}$, and k is the per-period rate of return available on an otherwise similar n-payment investment but offering no risk of default.

Generalizing the Results

By making some simplifying assumptions it is easy to demonstrate the generality and utility of this development. For example, if we simply assume a constant POD $= q$, then the number of periods that a firm will survive – based on $(1-q)$ – and consequently the number of possible interim payments available to a creditor, is a random variable that is geometrically distributed. Taking the expected value of a geometrically distributed random variable yields the expected number of interim payments available to a creditor: $\mu = (1-q)/q$. If we further assume that the interim payments, I, and the liquidation value of the collateral, L, are all constant, Eqs (4a), (4b) and (4c) can be simplified to the following:

$$CT = qLU[1 + (1-q)U + (1-q)^2U^2 + (1-q)^3U^3 + \cdots + (1-q)^{n-1}U^{n-1}] \quad (5a)$$

$$IPT = (1-q)IU[1 + (1-q)U + (1-q)^2U^2 + (1-q)^3U^3 + \cdots + (1-q)^{n-1}U^{n-1}] \quad (5b)$$

$$FPT = F(1-q)^nU^n \quad (5c)$$

Recognizing that the CT and IPT terms are finite geometric series, we can simplify D to a more manageable form shown in Eq. (6)[1]:

$$D = \frac{[(qL + (1-q)I)(1 - (1-q)^nU^n)]}{(1-q)} + F(1-q)^nU^n \quad (6)$$

All of the discussions that follow are based on this simplified equation, essentially the same equation developed by Fons (1987).

Next, we let D_0 equal the par value of the loan ($D_0 = F$) and define the liquidation value of the collateral, L, as proportion of the amount lent – i.e. $\delta_D = L/D_0$. We can now rewrite Eq. (6) in terms of a conventional interest-only loan as:

$$D = \frac{[(q\delta_D D_0 + (1-q)I)(1 - (1-q)^nU^n)]}{(1-q)} + D_0(1-q)^nU^n \quad (7)$$

Finally, at the time a loan is initiated we have $D = D_0$. Therefore, substituting D_0 for D in Eq. (7) and solving for D_0, we get the equilibrium relationship between the amount to be lent and the parameters specifying the risky, future repayment schedule:

$$D_0 = \frac{(1-q)I}{[1 - q(1-\delta_D)]} \quad (8)$$

It is interesting to note that for conventional interest-only lending, the correction for risk of default *at the time of issue* is independent of the term of the loan, a result suggested by the simulation studies of Bierman and Hass (1975). Yawitz's (1977) work suggested this result would also hold for bonds sold at premium or discount but Eq. (7) does not support that conclusion. The simplification afforded by Eq. (8) applies specifically to the analysis done by the lender before actually issuing the bond. At the time of the issue, the bond would likely be sold at a discount or premium due to changes in the market place. Once issued, Eq. (6) would apply.

In summary, Eqs (4a), (4b) and (4c) provide the general mathematical model to allow the valuation of a debt investment given the potential for default and truncation of the future income stream. Unfortunately, these are unwieldy and of little pedagogical value. However, by making some conventional assumptions, these equations can be simplified to the more manageable forms of Eqs (7) and (8).

As noted earlier, Fons claimed that his development was based on the assumption of risk neutrality of the lender, thus leaving undefined the relationship between debt valuation, risk of default, and capital market risk. However, that assumption has a shortcoming in that Fons ignored the linkage to capital market risk in choosing a correct discount rate. In fact, an appropriate calculation of the discount rate, *k*, provides the linkage to capital market risk.

DEVELOPMENT OF THE SECURITY MARKET LINE

For the development, the following form of the CAPM is employed:

$$r_i = r_f + \lambda\sigma_{si} \tag{9}$$

where, r_i is the required rate of return on investment *i*.

Within this frame of reference, capital market risk is related to variability in the actual rate of return and quantified by the term, $\lambda\sigma_{si}$. Finally, the SML could be rewritten as:

$$r_i + f_f + \lambda\rho_{i,m}\sigma_i \tag{10}$$

where,

$\rho_{i,m}$ = the coefficient of correlation between the returns on investment *i* and the returns on the market portfolio.

σ_i = the standard deviation in the returns on investment *i*.

The correlation definition for systematic risk has been long recognized in the literature (Francis & Archer, 1979), but has mostly given way to the use of beta. However, to put the returns on a debt investment into the frame of reference of capital market theory defined by Eq. (10), we need to specify a standard deviation in the per period rate of return and an appropriate coefficient of correlation. It is important to note that the assumption of static equilibrium is critical to this step of the analysis.

Francis and Archer were specific about static equilibrium in that at the time of analysis the prices of all securities are correct and the debt and equity markets are in equilibrium. Thus, in Eq. (10), r_i, r_f, $\rho_{i,m}$ and σ_i are assumed constant. That further specifies that the market line is stable – i.e. that λ is a constant – and that the term structure of interest rates is constant. Capital market theory, then, is related to actual rates of return that are random within this otherwise static environment.

A common mistake made at this point is to define a beta for debt without basing it on the correct frame of reference, as Fons suggested, and without relating it to a correct model for the periodic rate of return. For example, in their attempt to deal with the risk of default, Bierman and Oldfield assumed that the actual value of the interim payments on a loan would be a random variable, made no provision for the termination of these payments as a result of default, and assumed that bankruptcy would be costless. However, by allowing for the partial payment of interest, and no recourse for non-payment, they avoided the inherent differences between debt and equity investments and assumed away the risk of default.

Default will occur when the profit after taxes in a given year, Y, is less than or equal to Y_L where Y_L is the amount of profit so low (i.e. negative) as to precipitate default and liquidation. Some of the current assets of the business would certainly have been liquidated in an attempt to remain solvent. Thus, profits below Y_L would mean that unpaid claims that would normally have been paid as expenses still existed; e.g. wages. Such claims could certainly impinge on the liquidation value of the business to both debt and equity investors to such an extent as to render the value of equity equal to zero and to reduce the value of debt to less than its book value. In summary, the interest payment to a debt investor is not a random variable; indeed, it is defined by contract in accordance with the equilibrium relationships specified by Eqs (4) and (6). Rather, the life of the business and the liquidation value of the collateral given bankruptcy is the source of default risk to all investors.

From this reference point, the model for calculating the per-period return on a debt investment subject to default can be show to be:

$$r_D = \frac{q\delta_D D_0 + (1-q)(I + D_{t+1})}{D_t} - 1 \tag{11}$$

where D_t is the market value of a debt investment at any time t given that n payments remain as given by Eq. (6); $D_t = D_0$ at the time of issue.[2]

Related to capital market risk, we are concerned with the potential correlation between the two potential random variables – i.e. q and δ_D – and the return on the market portfolio. Given static equilibrium, the assumption is made that the conditions that could precipitate default – i.e. the value of Y_L and the probability distribution of Y – are constant. As was previously demonstrated, the actual liquidation value of the assets is a function of the actual value of Y; and Y is clearly correlated to returns on the market portfolio. Thus, the POD will be considered to be constant while δ_D will be assumed to be the random variable. Then the equation for the random returns on a debt instrument can be rewritten as:

$$\tilde{r}_D = \frac{q\tilde{\delta}_D D_0 + (1-q)(I + D_{t+1})}{D_t} - 1 \tag{11a}$$

We can now define the standard deviation in returns on the debt investment, the coefficient of correlation between those returns and those of the market portfolio, and thus the market risk of a debt investment.

From Eq. (11a), variability in the rate of return is clearly related to the first term. Unfortunately, this term is not constant over time but rather, given a maturing loan, is always changing as the security rides down the yield curve. Indeed, given a normal yield curve and a monotonically increasing term structure of interest rates, $D_t \geq D_0$. Therefore, assuming either a relatively flat yield curve or specifying the correction for the worst case as $D_t = D_0$, it can be seen that in Eq. (10) $\sigma_i = q\sigma_t$ and $\rho_{i,m} = \rho_{t,m}$ where,[3]

σ_i = the standard deviation in the liquidation value of the business.
σ_t = the standard deviation in the returns on investment t.
$\rho_{t,m}$ = the coefficient of correlation between the liquidation value of the business and the returns on the market portfolio.

From Eq. (10) and the definitions above, the resulting form of the SML, given a non-zero POD, can then be defined as:

$$k = r_f + \lambda\rho_{t,m} q\sigma_i \tag{12}$$

Several points are worth noting. First, it is the objective of the security market line – as conventionally defined in Eq. (10) – to provide the rate of return an investor would require to accept an investment whose returns were random, but correlated with the returns on the market portfolio. By correctly calculating the periodic returns on an investment subject to premature termination due to default, Eq. (12) provides the compensation for capital market risk explicitly considering a non-zero POD; i.e. it corrects for variability in the future cash flows associated with default but not for the risk of loss due to default. In other words, Eq. (8) provides the value of an investment correcting for the risk of loss associated with default given an appropriate rate of return while Eq. (12) provides that rate of return.

Substituting the correction for capital market risk from Eq. (12) into Eq. (8), the equilibrium relationship between the amount to be lent and the parameters defining the risky, future repayment schedule can be shown to be:

$$D_0 = \frac{(1-q)I}{r_f + q(1 + \lambda\rho_{t,m}\sigma_t - \delta_D)} \tag{13}$$

We can now specify the interest payment, I, to be:

$$I = \frac{D_0[r_f + q(1 + \lambda\rho_{t,m}\sigma_t - \delta_D)]}{(1-q)} \tag{14}$$

and the coupon rate on a loan, d, to be:

$$d = I/D_0 = \frac{[r_f + q(1 + \lambda\rho_{t,m}\sigma_t - \delta_D)]}{(1-q)}\text{[4]} \tag{15}$$

By partitioning the risks to indicate the specific compensation for interest rate risk, capital market risk, and default risk (but ignoring other risks such as marketability or call risk), Eq. (16) represents the formulation for an integrated security market line for debt at issue.

$$d = \frac{r_f + q\lambda\rho_{t,m}\sigma_t}{(1-q)} + \frac{q(1 + r_f - \delta_D)}{(1-q)} \tag{16}$$

Conventional thinking suggests that debt securities should be on the capital market line just as equity securities. And, indeed, Eq. (16) is consistent with this logic, but in a quite different form than might be expected. Furthermore, rewriting Eq. (16) to be more in line with conventional specifications of a security market line would give:

$$d = \frac{[r_f + q(1 - \delta_D)]}{(1-q)} + \frac{q\lambda\rho_{t,m}\sigma_t}{(1-q)} \tag{17}$$

The required rate of return on all loans thus has two major components – in addition to the risk-free rate – both linked to the risk of default. In particular, the intercept specified by the first term is itself loan-specific and, thus, this logic basically denies the existence of some generalized capital market line for debt. It is interesting to note that a security market line for traded debt does not exist in the conventional sense. It is not possible to resolve Eq. (7) for an explicit rate of return given a current market value, D, different from the issuing price, D_0. Rather, because k appears in the term U, the solution for the equilibrium value of k must be done by recursive methods.

Another important point is that the risk-free rate in the SML for debt must be of the same maturity as the loan itself. This can be seen in Eq. (15) by assuming either a zero POD or costless bankruptcy; i.e. $\delta_D = 1 + r_f$ and $\sigma_t = 0$. Given that, borrowing would be at the risk-free rate with no regard for capital market effects. But, for that rate to be consistent with conventional bond valuation equations, that rate must be from the term structure of interest rates at the appropriate maturity. In short, the risk-free rate in the security market line related to a debt investment is not the short-term treasury security rates as often assumed.

APPLICATION OF THE DEFAULT PRICING MODEL

The application of the model is straightforward as demonstrated by an example in this section. Most of the inputs are either observable or can be easily estimated or obtained from third party sources. We consider a bond issue from Ford Motor Company. The bond has a 6.375% coupon rate and matures on November 5, 2008. We estimate Eq. (15) (or equivalently Eqs (16) or (17)) using weekly return data from January 2, 1998 to June 13, 2001 to obtain the required return. The risk-free rate on a U.S. Treasury note maturing in 2008 is 5.06%. The market portfolio employed is an equally weighted average between the S&P 500 stock index and the Dow Jones Corporate Bond Index.

We assume that the collateral backing the bonds has a liquidation value equal to the value of the loan – i.e. $\delta_D = 1.00$. Further, we assume that the lender has private information regarding Ford's ability to generate cash flows, and estimates the POD, q, on the bonds to be 0.10.[5] Applying the data to Eq. (15) yields a required rate of return of 5.66%. This return seems reasonable given the bond's 5.96% yield-to-maturity as of June 2008. Notice that the required rate of return is less than the yield-to-maturity. This is to be expected since the model considers only default risk and abstracts from all other sources of risk such as liquidity risk and call risk.[6] Consequently, the required return obtained from the model will be systematically less than current market rates.

Figure 1 illustrates the effect of the POD on the required rate of return. The POD varies from $q = 0$ to $q = 0.90$, while all other factors are held constant. At $q = 0$, the required rate of return is the risk-free rate. This is reasonable, since absent the risk of default, the riskiness of a corporate debt security is identical to that of a Treasury security.[7] However, as q increases, the required rate of return increases at an increasing rate.[8]

Figure 2 shows the bond's required rate of return as a function of its beta, holding all other parameters in the model constant. This is the Security Market Line for debt. With a beta of zero, the Ford bonds have a required rate of return of 5.62% which is greater than the risk-free rate of 5.06%. This result is intuitive since an absence of correlation with a market index does not protect

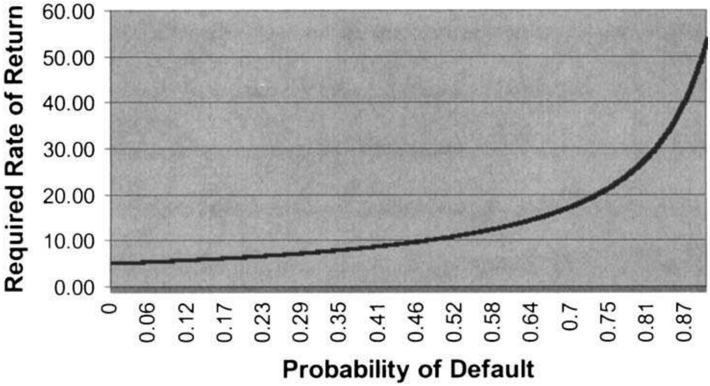


Fig. 1. Required Rate of Return as Function of Probability of Default.

or "diversify" the investor against default. However, as the bond's beta increases, the required rate of return on the bond increases linearly as nondiversifyable risk increases.

Last we consider the effect of δ_D, the liquidation value of the collateral as a percentage of the amount of the loan. A priori, we expect an increase in δ_D to reduce a lender's return demanded on the bonds. Figure 3 confirms this intuition. The liquidation value of the collateral as a percentage of loan value, δ_D, is examined over a range from 0% (essentially equivalent to a debenture) to 52%. At $\delta_D = 0.00$, the required rate of return is 5.77%. However, the required

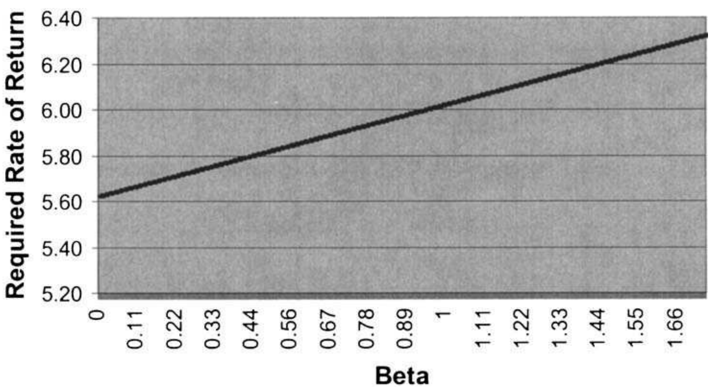


Fig. 2. Required Rate of Return as Function of Beta.

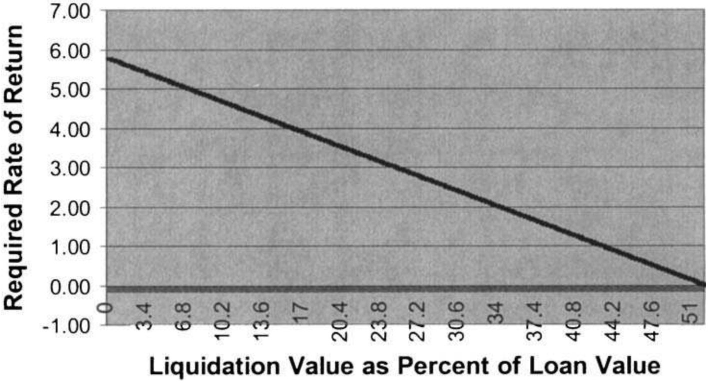


Fig. 3. Required Rate of Return as Function of Liquidation Value.

rate of return declines linearly to 0.00% at $\delta_D = 52\%$. The rationale behind this decline is that as the liquidation value of the collateral increases, the bonds become more valuable to the lender if the borrower defaults. That is, the firm is more valuable to creditors "dead than alive."

SUMMARY

The common specification for "risky" debt (e.g., see Callahan & Mohr, 1989) neither includes any correction for risk of default, nor is it based on a correct formulation for capital market risk given static equilibrium. As in Fons, this paper has clearly specified the correct frame of reference within which valuation of a security with a non-zero POD must occur. Extending the logic to account for capital market risk led to the specification of a security market line for debt (Eq. 15), differing in form from conventional models. Consistent with the traditional theory of capital structure, Eq. (15) specifies an interest rate that rises – and Eq. (7) a market value that falls – as POD increases above zero and/or the value of the collateral falls.

Many authors, in developing models related to an analysis of optimal capital structure, loan portfolio management, or capital market equilibrium have based their conclusions upon an implicit assumption that there is no risk of default. Building on the methodology first delineated by Fons, an explicit correction for the risk of loss given default was developed leading to the correct model for valuation of debt investment. This model could be simplified to current models popular in the literature by abstracting from the possibility that a firm would be unable to generate sufficient cash flows to make the promised payments on its debt, thus casting the firm into default and triggering a subsequent liquidation of assets. However, this study overcomes more simplified models by providing a calculation methodology for a security market line appropriate to a debt investor explicitly considering both default risk and capital market risk

NOTES

1. Here, $CT = qL[(1-(1-q)^n U^n)]/(1-q)$ and $IPT = I(1-(1-q)^n U^n)$. Factoring out common terms and adding the *FPT* term yields Eq. (6).
2. The numerator in Eq. (11) is the expected value of the bond at time t including the expected liquidation value of the collateral in the event of default. The denominator is the current market value of the bond. Dividing the former by the latter and subtracting unity, gives us r_D.
3. As previously indicated, the liquidation value of the firm's assets, L (a random variable), is a function of Y, a random variable of the firm's after tax profits. It is clear that Y is correlated with the market portfolio. When a realization from Y drops below Y_L, the firm is in default. The probability of a firm entering into default is q. Consequently, $\sigma_i = q\sigma_t$ and $\rho_{i,m} = \rho_{t,m}$.
4. Note that $\lambda\rho_{t,m}\sigma_t = \rho_{t,m}\sigma_t/\sigma_m(r_m - r_f)$. We can rewrite this as $\beta_t(r_m - r_f)$, where β_t is the beta coefficient estimated from a regression of bond returns on the returns of a market index. Thus, β_t has an interpretation similar to that of an equity beta.
5. A lender without access to private information could obtain the POD from publicly available default rate statistics published by debt rating agencies such as Moody's or Standard & Poors.
6. The model does, however, take into account interest rate risk since the model requires the risk-free rate to be obtained from a risk-free bond with the same maturity as the bond under consideration.
7. Again, this ignores other risk factors such as liquidity risk.
8. However, if the liquidation value of the collateral substantially exceeds the value of the loan, then the required rate of return on the firm's debt *decreases* at an increasing rate as q increases. This occurs because as δ_D increases, the value of the assets in the event of liquidation exceeds the value of the cash flows from the loan itself. Hence, a lender would be better off in the event of default and therefore demands a lower rate of return as q increases.

REFERENCES

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy. *Journal of Finance*, *23*, 589–609.

Altman, E. I. (1980). Commercial Bank Lending: Process, Credit Scoring and Costs of Errors in Lending. *Journal of Financial and Quantitative Analysis*, *15* (November), 813–932.

Altman, E. I. (1982). Accounting Implications of Failure Prediction Models. *Journal of Accounting, Auditing and Finance*, *6* (Fall), 4–19.

Altman, E. I., & Spivack, J. (1983). Prediction Bankruptcy: The Value Line Relative Financial Strength Systems vs. The Zeta Bankruptcy Classification Approach. *Financial Analysis Journal*, *39* (November/December), 60–67.

Altman, E. I. (1984). A Further Investigation of the Bankruptcy Cost Question. *Journal of Finance*, *39* (September), 1067–1089.

Bierman, H., & Haas, J. E. (1975). An Analytic Model of Bond Risk Differentials. *Journal of Financial and Quantitative Analysis*, *10* (December), 757–773.

Bierman, H., & Hausman, W. (1970). The Credit Granting Decisions. *Management Science*, *16*, B519-B532.

Bierman, H., & Oldfield, G. S. Jr. (1979). Corporate Debt and Corporate Taxes. *Journal of Finance*, *34* (September), 951–956.

Callahan, C. M., & Mohr, R. M. (1989). The Determinants of Systematic Risk: A Synthesis. *The Financial Review*, *24* (May), 157–181.

Conine, T. E. (1980). Corporate Debt and Corporate Taxes: An Extension. *Journal of Finance*, *35* (September), 1033–1037.

Cowen, S. S., & Page, A. L. (1982). A Note on the Use of Selected Nonfinancial Ratio Variables to Predict Small-Business Loan Performance. *Decision Sciences*, *13* (January), 82–87.

Dirickx, Y., & Wakeman, L. (1976). Extension of the Bierman-Hausman Model for Credit Granting. *Management Science*, *22* (July), 1229–1237.

Francis, J. C., & Archer, S. H. (1979). *Portfolio Analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Fons, J. S. (1987). The Default Premium and Corporate Bond Experience. *Journal of Finance*, *42* (March), 81–97.

Hamada, R. S. (1969). Portfolio Analysis, Market Equilibrium and Corporate Finance. *Journal of Finance*, *24* (May), 13–31.

Haugen, R. A., & Senbet, L. W. (1988). Bankruptcy and Agency Costs: Their Significance to the Theory of Optimal Capital Structure. *Journal of Financial and Quantitative Analysis*, *23*, 27–38.

Lincoln, M. (1984). An Empirical Study of the Usefulness of Accounting Ratios to Describe Levels of Insolvency Risk. *Journal of Banking and Finance*, *9* (June), 321–34.

Mensah, Y. M. (1984). An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A Methodological Study. *Journal of Accounting Research*, *22* (Spring), 380–395.

Modigliani, F., & Miller, M. H. (1963). Corporate Income Taxes and the Cost of Capital: A Correction. *American Economic Review*, *53* (June), 433–443.

Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18* (Spring), 109–131.

Rubinstein, M. E. (1973). A Mean-Variance Synthesis of Corporate Financial Theory. *Journal of Finance*, *28* (March), 167–181.

Srinivasan, V. (1984). More on Break-even Analysis. *Journal of Commercial Bank Lending*, *66* (July), 53–58.

Vinso, J. (1979). A Determination of the Risk of Ruin. *Journal of Financial and Quantitative Analysis*, *14* (March), 77–100.

Yawitz, J. B. (1977). An Analytical Model of Interest Rate Differentials and Default Recoveries. *Journal of Financial and Quantitative Analysis*, *12* (September), 481–490.

This Page Intentionally Left Blank

SHAREHOLDER HETEROGENEITY: FURTHER EVIDENCE

Yi-Tsung Lee and Gwohorng Liaw

ABSTRACT

Some studies, such as Bagwell (1992) and Bernardo and Cornell (1997), provided evidences that the shareholders' valuations differ dramatically. They argued that the valuations differ substantially, implying a significantly small supply or demand elasticity. However, Kandel et al. (1999) indicated quite an elastic demand for stocks of Israeli IPOs that were conducted as nondiscriminatory auctions. To resolve these controversial findings, this paper discusses the procedure of measuring price elasticity and provides some measures of elasticity. In addition to indicating that Bagwell's measure tends to underestimate the actual elasticity, this study supplements previous work by testing under another auction mechanism, discriminatory pricing rule, and our results are consistent with Kandel et al.'s findings.

1. INTRODUCTION

Theoretical and empirical studies provide different views on the nature of price elasticity of common stock. Whether price elasticity of common stock is perfectly elastic or not has significant implications on the cause of price fluctuation when the number of outstanding shares changes substantially. According to finance theories, when there is a perfect market or when individual stocks have perfect substitutes, stock prices would not change even if there are large amounts of new shares offering or stock repurchase.

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 181–202.

ISBN: 0-7623-0887-7

Researchers attribute large price movements, if any, to the information effect, which causes a shift in the supply and/or demand curve. However, this prediction is not supported by empirical findings such as Harris and Gruel (1986), Shleifer (1986), and Loderer et al. (1991).

In the case of IPOs (stock repurchase), supply (demand) in the stock market is given and, hence, the supply curve is a vertical line. On the demand (supply) side, investors' valuations for the same stock will be different, if shareholders are heterogeneous. Then, investors will submit their buy (sell) orders on various sell prices. If this were the case, the demand (supply) curve drawn from investors' order placement is a downward (upward)-sloping curve and is not perfectly elastic. The less heterogeneous the shareholders are, the more flat and elastic the demand (supply) curve will be. Harris and Gruel (1986), Shleifer (1986), and Loderer et al. (1991) found that the demand curve for individual stocks is downward sloped rather than horizontal. When the demand curve is not perfectly elastic, a huge increase in supply leads to a movement along the demand curve.[1] As a result, stock price will drop even if information effect does not exist.[2]

Whether stock prices move due to a change in the number of shares outstanding appears to be determined jointly by information effect and finite price elasticity. However, it is empirically difficult to isolate the finite price elasticity effect from the information effect. Bagwell (1991b, 1992), Bernardo and Cornell (1977) and Kandel et al. (1999) have attempted to deal with this issue by using different auction datasets. An auction dataset allows the shareholders to directly evaluate the same security; if shareholders' evaluations of an asset are homogenous, bid prices offered by different shareholders would be similar. This result would be consistent with the prediction of price movement according to the perfect elasticity hypothesis. On the contrary, shareholder heterogeneity implies that the demand curve of common stock is not perfectly elastic.

Using data on Dutch Auction repurchases of 31 U.S. firms, Bagwell (1992) reported that the average arc (point) elasticity of the supply schedule of common stock is 1.65 (0.68). This indicates that firms face upward-sloping curves and that shareholders' valuations are very different when they repurchase shares in Dutch auction (nondiscriminatory pricing). Bernardo and Cornell (1997) also found that the average price range (the dispersion of bids) for a sample of mortgage securities during a selling auction is 63%, further supporting the claim that heterogeneous investor valuations imply a significantly small demand elasticity. This implies that the assumption of traditional finance theories (such as CAPM, APT, and MM propositions) that all investors are homogeneous may not hold, since in the above-mentioned empirical evidence the demand or supply schedules in the auctions appear to be downward- or upward-sloping, further implying that investors are heterogeneous. However, Kandel et al. (1999) provided evidence that price is relatively quite elastic, using Israeli IPO auctions that were conducted as non-discriminatory (uniform pricing) auctions. They found that the average (median) demand elasticity is 37.1 (21.0) around the auction clearing price. Given the contradictory findings of Bagwell (1992) and Kandel et al. (1999), it is important to provide further evidence using discriminatory pricing auctions sample.[3]

The first purpose of this paper is to re-examine Bagwell's finding using the original data he released, and to explore if the findings also apply to the Taiwan market using 50 Taiwanese IPOs auction data. We show that Bagwell's approach tends to underestimate the price elasticity of common stock. Underestimation occurs when the absolute rather than the relative concept of elasticity used. In addition, Bagwell's estimation ignores the fact that some firms may face a perfectly elastic supply curve. Consequently, Bagwell's averaging procedure would also underestimate average price elasticity substantially.

To avoid the problem of under-estimation, we measured arc-elasticity using Bagwell's samples. Our results indicate that the average values of the arc-elasticity are 27.7, 24.45, 18.45, and 22.44, respectively, compared to 1.10, 1.96, 2.51 and 1.65 in Bagwell's results for 1st, 2nd, 3rd and full intervals. This phenomenon reveals that after adjusting measurements of elasticity, the elasticity is larger than it was in Bagwell's study.

The second purpose of this paper is to test whether the price is elastic or not using different auction samples. Our dataset consists of 50 Taiwanese IPOs that were conducted as discriminatory auction (i.e. you pay what you bid). The results indicate that the average point elasticity of demand is 14.2, which is consistent with Kandel et al. (1999). In the following section, the debate on the nature of price elasticity of common stock is delineated. Section 3 analytically and empirically compares the elasticity measure used by Bagwell and a measure based on a relative concept of arc elasticity. Section 4 estimates price elasticity of demand for a sample of Taiwanese IPOs. Conclusions are stated in the final section.

2. DEBATE ON PRICE ELASTICITY OF COMMON STOCK

Stock repurchase has been widely used as a tool to resist hostile takeover. The effectiveness of stock repurchase as a defensive tool is largely dependent on the price elasticity of the supply curve. Using a game theory model, Bagwell (1991a) proved that if the supply schedule of common stock is positively sloped, stock repurchases would impose a high cost on bidders. This is because shareholders selling the stock back to the firm value the firm‘s stock less than those retaining their shares. Therefore, bidders need to offer a higher price to acquire shares of the target firm. Using Dutch auction data, Bagwell (1991b, 1992) found that supply curves are positively sloped with an average slope of 1.46 and an average arc (point) price elasticity of 1.65 (0.68). This indicates that firms face upward-sloping curves and shareholders' valuations are very heterogeneous when they repurchase shares in (nondiscriminatory pricing) Dutch auction.

On the demand side, many empirical works show that stock prices drop significantly when the supply of stocks is increased dramatically. Loderer et al. (1991) found that price decreases upon the announcement of SEOs are due to finite elasticity of demand curve rather than information effect. An average inverse elasticity of –11.12 for a sample of 409 firms is reported. The claim of finite elasticity of demand for stock is further supported by Bernardo and Cornell (1997). The slope (or price elasticity) of the demand or supply curve is largely related to how different shareholders evaluate the same security. The heterogeneous valuation hypothesis proposes that investors place different values on the same security and, therefore, the demand (or supply) schedule possesses small elasticity. In contrast, the homogenous valuation hypothesis suggests that different investors evaluate the same security identically; in this case, demand (or supply) is perfectly elastic. In other words, shareholder heterogeneity implies that the individual demand (or supply) curve is downward- (upward-) sloping, while shareholder homogeneity implies that the curve has infinite price elasticity.

The nature of price elasticity of common stock has important implications regarding the cause of price movement. For example, if stock repurchase is regarded as a positive signal, (i.e. the stock is undervalued), an increase in stock price after repurchase announcement can be attributed to two effects: the information effect and the substitution (or price elasticity) effect. The information effect leads to an upward shift of the supply curve while the price elasticity effect causes an upward movement along the supply curve. Figure 1 illustrates the possible price movement after a repurchase announcement. In Fig. 1, S_o and P_o denote the supply curve and the equilibrium price, respectively, before repurchase announcement. When the firm announces a stock repurchase, demand curve shifts from P to D_1, and the equilibrium price shifts from point a to point b, *ceteris paribus*. The movement from point a to point b indicates the substitution effect. If the repurchase announcement conveys a

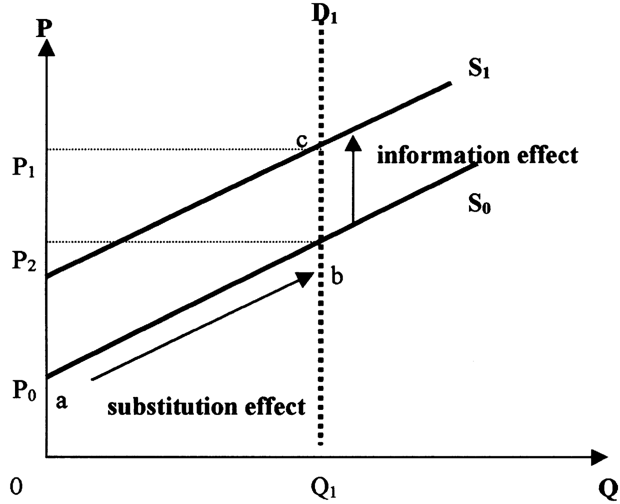


Fig. 1. The Impact of Information and Substitution Effects.

positive signal, the supply curve would shift from S_0 to S_1, and the final equilibrium point would be at c. As pointed out by Bagwell (1992), the information effect and the substitution effect are not mutually exclusive. Since there is a lack of a definitive model of the sources of shareholder heterogeneity, Bagwell refers to such hypotheses, relying on movement along an upward-sloping supply curve, collectively as the heterogeneous valuation hypothesis. This is clearly illustrated in Fig. 1.

In examining Dutch auction repurchase, Bagwell (1992) reported a price increase of approximately 7.7% upon the repurchase announcement. From his table II, 9.1% price premium obtains 15% change of the outstanding shares. It implies an average elasticity measure of 1.65. Bagwell (1992) suggested that if we assume that the repurchase caused only movement along the supply curve, then the pre-announcement price change of 7.7% is lower than the 9.1% premium based on the estimated price elasticity (1.65). This implies that the price increase is merely due to substitution effect, and that the information effect plays no role in explaining the price movement due to repurchase announcement. Referring to Fig. 1, Bagwell's result indicates that price moves from a to b, and the shift from b to c never takes place.

According to Bagwell's heterogeneous valuation hypothesis, the explanation of price change can be accompanied with price elasticity measurement: the smaller the estimated price elasticity, the more likely it is that the price change is caused by shareholder heterogeneity. In contrast, if price elasticity is large, price change is most likely caused by information effect. That is, when price elasticity is large enough (i.e. the supply curve is horizontal in Fig. 1) to cause trivial price change, we can claim that securities have perfect substitutes, as assumed under the conventional theory.

3. MEASURING PRICE ELASTICITY OF COMMON STOCK

In this section, we used the general formula for arc elasticity to demonstrate that Bagwell's measure tends to underestimate price elasticity. In addition, other problems inherent in Bagwell's average elasticity measure are discussed. Let P_i and Q_i stand for the price and relatively accumulative quantity offered by investors in the auction for firm i, $i = 1, 2, \ldots, N$; let P_i^0 and S_i^0 be the original price and original number of outstanding shares; and let p_i and q_i stand for the percentage price and related standardized accumulative quantity, where $p_i = P_i / P_i^0$ and $q_i = Q_i / S_i^0$. Bagwell's measure of average elasticity can be expressed as:[4]

$$\eta = \Delta q \Bigg/ \left[\sum_{i=1}^{N} \Delta p_i / N \right] \tag{1}$$

Bagwell's measure of price elasticity is equal to the ratio of all firms' average standardized quantity change (Δq) to the average of percentage price changes (Δp_i). Bagwell assumed Δq_i for all firms is 5%, as a result, Δq is also equal to 5%. Since the average repurchase rate for the sample of Dutch auction repurchase is 15.28%, Bagwell also estimated average price elasticity for a range of Δq to be from 1% to 15%. We reproduce Bagwell's results (p. 78) in Table 1. In order to show the reason why his elasticity is underestimated, let ε_i be the arc elasticity over an interval in a supply curve for firm i, ε_i can be written as

$$\varepsilon_i = \frac{\Delta q_i}{\Delta p_i} \cdot \frac{p_i^m}{q_i^m} \tag{2}$$

For a given interval of the supply curve, p_i^m (q_i^m) is the average of the beginning and ending prices (accumulative quantity) for firm i. When the original price and quantity instead of the standardized value are used, Eq. (2) becomes:

$$\varepsilon_i = \frac{\Delta Q_i / S_i^0}{\Delta P_i / P_i^0} \cdot \frac{P_i^m / P_i^0}{Q_i^m / S_i^0} = \frac{\Delta Q_i}{\Delta P_i} \cdot \frac{P_i^m}{Q_i^m} \tag{3}$$

A comparison of Eqs (2) and (3) indicates that standardization has no impact on the value of arc elasticity. Therefore, arc elasticity is invariant to original price level and measurement unit. However, Bagwell's price elasticity measure for an individual firm using non-standardized data is equal to:

$$\eta_i = \frac{\Delta q_i}{\Delta p_i} = \frac{\Delta Q_i}{\Delta P_i} \cdot \frac{P_i^0}{S_i^0} \tag{4}$$

It is clear that Eq. (4) tries to measure the elasticity of a supply curve at the point (P_i^0, S_i^0). However, P_i^0/S_i^0 is not the average position before and after price changes as in Eq. (3). Strictly speaking, Eq. (4) only indicates the slope relationship between standardized price and standardized quantity; it ignores the impact of relative position in a supply curve on price elasticity.[5] The following equation demonstrates why η_i underestimates the widely used elasticity measure ε_i

$$\eta_i = \varepsilon_i \left[\frac{Q_i^m}{P_i^m} \cdot \frac{P_i^0}{S_i^0} \right] \tag{5}$$

In Eq. (5), the difference between P_i^m and P_i^0 is relatively small, but S_i^0 is much larger than the averaged quantity (Q_i^m) in general. Therefore, compared with the common arc elasticity measure, Bagwell's measure tends to substantially underestimate price elasticity.

A simple numerical example can illustrate the above proposition. Let us assume that the original stock price (P^0) is \$25, and the number of shares outstanding (S^0) is 100,000. Consider two points on a supply curve: (P1,Q1) = (\$20,1000) and (P2,Q2) = (\$25,1500), the standardized price and quantity mix according to Bagwell's measures are (p1,q1) = (80,1) and (p2,q2) = (100,1.5). The resulting price elasticity (η) is 0.025 ($\Delta q/\Delta p = 0.5/20$). On the other hand, the arc elasticity (ε) for the interval is 1.8 ($(\Delta q/\Delta p)*(p^m/q^m) = 0.025*(90/1.25)$). Suppose that the number of shares outstanding is 10,000 rather than 100,000, then after normalized by 10,000, (p1,q1) and (p2,q2) become (80,10) and (100,15) respectively, and η becomes 0.25 (i.e. 5/20) while ε is still equal to 1.8 (0.25(90/12.5)). Obviously, the assumption of the original number of shares outstanding will affect Bagwell's measure of price elasticity. In this example, if the original number of shares outstanding shrinks ten times from 100,000 to 10,000, Bagwell's measure of price elasticity increases tenfold. Actually, even if the original number of shares outstanding of the firm (the normalization unit) is assumed differently, the supply schedule will remain unchanged. Therefore, the elasticity measure shouldn't change. This signifies an underestimation of Bagwell's elasticity measure – the measure of price elasticity is negatively related to the number of outstanding shares. On the other hand, the common arc elasticity measure does not have this problem.

Some of the other characteristics of Bagwell's average elasticity measure deserve to be mentioned. Let η_i be the price elasticity for firm i, $i = 1, \ldots, N$, the average price elasticity for all firms in the sample should be

$$\eta = \left(\sum_{i=1}^{n} \eta_i \right) \Big/ N, \qquad \text{if } \eta_i \neq \infty.$$

Substituting Eq. (4) into the equation above, we get

$$\eta = \left(\sum_{i=1}^{N} \Delta q_i / \Delta p_i \right) \Big/ N, \qquad \text{if } \Delta p_i \neq 0. \tag{6}$$

When the relative changes in accumulative quantity for all the firms are the same, meaning $\Delta q_i = \Delta q$, Eq. (6) can be rewritten as

$$\eta = \Delta q \left(\sum_{i=1}^{N} \frac{1}{\Delta p_i} \right) \Big/ N, \qquad \text{if } \Delta p_i \neq 0. \tag{7}$$

Comparing Eq. (1) with Eq. (7), we get the following result:

$$\Delta q \left(\sum_{i=1}^{N} \frac{1}{\Delta p_i} \right) \Big/ N \neq \Delta q \Big/ \left(\sum_{i=1}^{n} \Delta p_i / N \right)$$

The above inequality implies that Eq. (1) may not be an appropriate measure for average elasticity. Equation (7) is different from Eq. (1) in that Eq. (7) sums up elasticity of individual firms first and then calculates the average. Conversely, in Eq. (1), relative price changes for all firms are summed first, after which the sum is divided by the total number of firms. Finally, the percentage change in quantity is divided by the average to obtain the average price elasticity.

Equation (7) is a meaningful measure if and only if $\Delta p_i \neq 0$. In a stock repurchase auction, firms usually set an upper and a lower price limit between which shareholder bids will be accepted. Shareholders bid the price within the boundary. As illustrated in Fig. 2, this means that there is a horizontal portion in the supply curve at the lower and upper price limit.

In a Dutch auction stock repurchase, priority is first given to the shareholder bidding the lowest price, and then to those with the next lowest price until the targeted amount is repurchased (point D). The firm repurchases the targeted

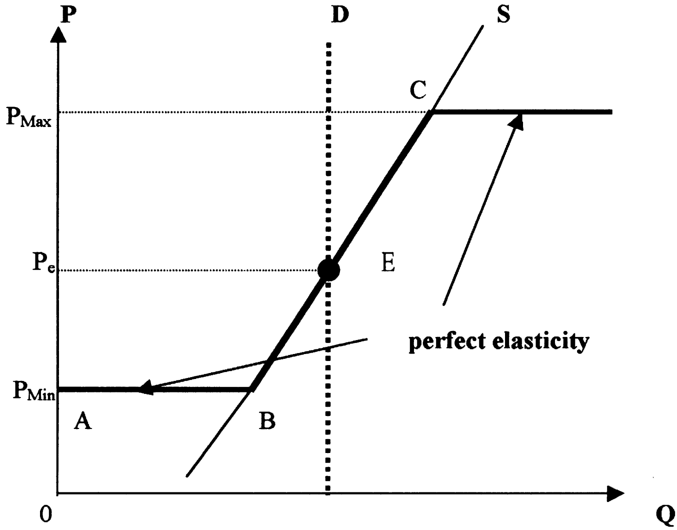


Fig. 2. Supply Curves of Stock Repurchase Dutch Auction.

amount at a single price level, P_e, which is the lowest price to acquire the number of shares sought. The proportion that is repurchased is indicated by the curve ABE. Unless the minimum price bid by shareholders is extremely low, a horizontal section of supply curve such as AB in Fig. 2 exists. In other words, it is highly possible for $\Delta p_i = 0$. This causes some difficulty in calculating average price elasticity using Eq. (7). As discussed below, an adjustment will be needed when Eq. (7) is used to measure average price elasticity.

Does Bagwell's measure of average price elasticity take into account this potential problem? Again, we used a simple example to address this question. Let $\Delta p_1 = 0\%$ and $\Delta p_2 = 5\%$ be the relative price changes for two firms; and $\Delta q_1 = 5\%$ and $\Delta q_2 = 5\%$ be the corresponding percentage changes in quantity. According to Eq. (7), elasticity for the two firms, η_1 and η_2, are ∞ and 1, respectively. Mathematically, it is impossible to find the average of η_1 and η_2. However, the average price elasticity based on Bagwell's equation (η) is equal to 2.5 [i.e. $5\%/((0\% + 5\%)/2)$]. Apparently, a value of 2.5 cannot reasonably represent the average elasticity of the firms. The underlying problem is that Eq. (1) ignores the possibility of perfect elasticity of common stock for individual firms. Since the purpose of this research is to investigate whether price elasticity is finite, we believe that only finite elasticity should be included in calculating average price elasticity.

The above discussion pinpoints two problems of Bagwell's measure of average price elasticity: (1) it does not consider relative position in the demand (or supply) curve; (2) it is inappropriate to sum up relative price changes of individual firms before calculating the average price elasticity. These two problems lead to serious underestimation of price elasticity.

To test the validity of our argument, we used Bagwell's Dutch auction repurchase (1992) data to re-estimate price elasticity using Eq. (7). Since Bagwell only listed the data for 17 firms out of a sample of 31 firms, Eq. (1) was first used to estimate price elasticity for the 17 firms. The results are compared with those obtained by Bagwell for the full sample.

Table 1 shows that the estimated price elasticity for the 17 firms are similar to those calculated with the full sample. The average elasticity for the 17 firms is 1.58 compared with 1.65 for the full sample. Therefore, when a firm repurchases 15% of its outstanding shares, a movement along the supply curve causes a 9.08% (n = 31) or 9.48% (n = 17) price increase. This indicates that the omission of 14 firms from the full sample seems to have limited impact on the evaluation of average price elasticity.

To compare the results of alternative elasticity measures specified in Eqs (1) and (2), we used Eq. (2) to re-estimate the price elasticity for the 17 firms. As shown in Table 2, the supply curve for S. Brands, and NL Ind. are perfectly elastic within the considered range. For the firms with available data, ε_i is ranged from 6.97 to infinite for each 5% change interval in quantity. For all

Table 1. Comparison of Bagwell's Elasticity Measures for 31 Firms and for 17 Firms.

The elasticity for 17 firms provided from the appendix in Bagwell (1992) is calculated by using $\eta = \Delta q \bigg/ \left[\sum_{i=1}^{N} \Delta p_i / N \right]$, where Δq is the average standardized quantity change of all firms, Δp_i is the percentage price change for firm i, and N is the number of sample firms.

Quantity Change Intervals (%)	Mean % Price Change (n = 31)	Mean Elasticity η (n = 31)	Mean % Price Change (n = 17)	Mean Elasticity η (n = 17)
1–6	4.54	1.10	4.24	1.18
6–11	2.55	1.96	3.15	1.59
11–16	1.99	2.51	2.09	2.38
Sum 1–16	9.08	1.65	9.48	1.58

Source: Table 2 and appendix B in Bagwell (1992).

Table 2. The Common Arc Elasticity Measures of the Dutch Auction Stock Repurchase Supply Curves for 17 Firms.

The common arc elasticity measure is calculated by $\varepsilon_i = \frac{\Delta q_i}{\Delta p_i} \cdot \frac{p_i^m}{q_i^m}$, where Δp_i and Δq_i are the percentage price change and the related standardized accumulative quantity change for firm i, respectively, and p_i^m (q_i^m) is the average of the beginning and ending prices (accumulative quantity) for firm i. n.a. means that the value is not available since the accumulative quantity for the individual firm is less than the minimum of the range.

Quantity Change Interval (%)	Axia	FMC	Jostens	G. Signal	Penwalt	R. Purina	S. Brands.	C. Karche
1–6	15.83	infinite	50.3	infinite	infinite	14.72	infinite	20.7
6–11	16.15	30.2	30.7	6.97	infinite	n.a.	infinite	26.81
11–16	n.a.	13	n.a.	9.65	infinite	n.a.	infinite	32.56
Sum 1–16	13.17	36.78	33.85	14.39	infinite	14.72	infinite	17.34

Quantity Change Interval (%)	B. Group	J. R. Stever	Whittaker	T. Shipyar	Knogo	NL Ind.	RJR Nabi.
1–6	51.21	22.31	22.14	37.1	14.99	infinite	infinite
6–11	21.68	42.89	75.38	n.a.	13.23	infinite	10.48
11–16	n.a.	27.37	24.14	n.a.	8.69	infinite	13.74
Sum 1–16	31.1	19.34	20.14	37.1	10.14	infinite	21.26

Source: Appendix B in Bagwell (1992).

ranges of changes (1%–16%), the value of ε_i lies between 10.14 and infinite. In calculating the mean of elasticity, we excluded firms indicating infinite price elasticity. Table 3 indicates that average elasticity for our samples ranged from 18.45 to 27.70 for each 5% change interval in quantity, with an average value of 22.44 for the entire range.

Our estimate of average price elasticity (22.44) is about 14 times higher than that calculated by Bagwell's measure (1.58) for the same 17 firms and 13 times higher than that reported by Bagwell (1.65). This difference has an important implication on the estimation of price changes. Based on Bagwell's result, a 15% change in quantity supply would lead to approximately 9% of price changes. However, according to our elasticity measure, the resulting price change will only be 0.7%. This result implies that even if the supply

Table 3. The Mean of Common Arc Elasticity Measures of the Dutch Auction Stock Repurchase Supply Curves for All 17 Firms.

The mean elasticity is calculated by $\varepsilon = \left(\sum_{i=1}^{n} \varepsilon_i\right) \Big/ N$, where ε_i represents the common arc elasticity measure for firm i and N is the number of sample firms.

Quantity Change Interval (%)	No. of Firms	Mean Elasticity ε
1–6	9	27.70
6–11	10	24.45
11–16	7	18.45
Sum 1–16	12	22.44

Source: Appendix B in Bagwell (1992).

curve for individual firms is not perfectly elastic, its elasticity is high enough to nullify the effect of quantity change on stock price. This finding is consistent with the prediction of traditional theories assuming perfect price elasticity.

Furthermore, Bagwell's measure may not be suitable for analyzing the price range effect of an announcement of Dutch auction repurchase. As shown in Fig. 2, the lower section of the supply curve (dotted line portion) is replaced by a horizontal portion, AB. Due to the presence of a minimum price, P_{min}, the upward sloping section of the supply curve below point b does not exist in a Dutch auction. The price elasticity for section AB definitely overestimates the real supply curve indicated by the dotted line. To avoid this problem, we calculated average price elasticity only for the section of the supply curve denoted by BC. Section BC shows that the actual shareholders' evaluations and the resultant elasticity would not be infinite, after it alleviates the problem of summing a positive number with an infinite large number. Table 4 presents the estimate of the slope of supply curve (b), constant price elasticity (E) from logarithm regression, and the point elasticity from linear regression at the weighted winning bid price (Ew) for the portion of the supply schedule that lies between the upper and lower price limit.

The supply schedules of firms repurchasing shares are estimated by regressing bid price on the accumulated quantity. Columns (1) and (2) of Table 4 show the slope coefficients and the corresponding t-statistics. The average slope reported in Table 4 of 1.9 is higher than that of 1.46 documented by Bagwell (1992). The slight discrepancy may be due to the difference in sample

Table 4. The Supply Curves and Alternative Price Elasticity Measures of the Dutch Auction Stock Repurchases.

The supply schedules of firms repurchasing shares are estimated by regressing bid price on the accumulated quantity: $P = a + bQ$. Columns (1) and (2) show the slope coefficients b and the corresponding t-statistics. To estimate fixed price elasticity (E), we regressed the natural logarithm of accumulated quantity on the logarithm of price: $\ln Q = A + E \ln P$. The point elasticity at the weighted winning bid price traded (Ew) is equal to $(1/b)(P_w/Q_w)$, where b is the estimated slope of the supply curve; P_w and Q_w is the price-quantity mix at the weighted winning bid price traded. * (**) indicates significance at the 5% (1%) level.

Name	(1) B	T statistic	(3) E	T statistic	(5) Ew
Axia	2.43	7.33**	17.05	13.84**	6.31
B. Group	0.96	6.39**	30	9.58**	10.32
C. Karche	0.67	11.91**	17.2	17.25**	10.64
Far West	7.36	1.94	9.63	2.84*	6.72
FMC	0.35	7.47**	20.64	13.18**	17.47
G. Signal	0.6	11.33**	8.55	17.75**	6.18
J. P. Stever	0.4	8.33**	20.72	13.54**	21.15
Jostens	0.67	11.69**	34.21	12.79**	11.79
Knogo	1.57	8.81**	8.05	12.07**	4.91
NL Ind.	0.78	8.39**	8.82	8.09**	8.21
Penwalt	2.16	29.96**	0.59	29.56**	1.08
R. Purina	1.44	7.22**	12.66	9.14**	10.3
RJR Nabi	0.94	29.67**	8.12	22.38**	14.42
S. Energy	4.95	3.33**	23.35	6.27**	5.17
S. Brands	4.79	5.08**	0.16	5.08**	0.456
T. Shipyar	1.86	4.88**	33.13	6.93**	8.83
Whitaker	0.37	9.91**	21.42	12.38**	9.32
Mean	1.9	10.21	15.66	12.51	9.01
Median	0.96	8.33	17.05	12.38	8.83
Standard deviation	1.98	7.88	10.89	6.61	5.31
Minimum	0.35	1.94	0.16	2.94	0.456
Maximum	7.36	29.96	34.21	29.56	21.15

Source: Appendix B in Bagwell (1992).

size (17 vs. 31). Except for Far West, slope coefficients for all sample firms are statistically significant. Therefore, supply curves for individual firms are upward sloping. To estimate constant price elasticity (E), we regressed the logarithm of accumulated quantity on the logarithm of price. Alternatively, the point elasticity at the weighted winning bid price traded (Ew) is measured by

$(1/b)(P_w/Q_w)$, where b is the estimated slope of the supply curve; P_w and Q_w is the price-quantity mix at the weighted winning bid price. As illustrated in Columns (3) and (4) of Table 4, the average value of constant elasticity is 15.66; also, all the elasticity are significantly different from zero. Finally, the average price elasticity at the weighted winning bid price is 9.01. The values of E and Ew in this study are much higher than those reported by Bagwell (1992). Using the lowest estimates ($E = 9.01$) to predict price change, a 15% quantity change only leads to a 1.66% price change. While the supply curves are upward sloping, the average price elasticity of supply is not small. Our results contradict those reported in previous studies.

IV. THE PRICE ELASTICITY OF DEMAND FOR IPOs

After providing evidence of the heterogeneity of shareholders from supply side by using part of Bagwell's data, it is of interest if the demand curve also is not perfectly elastic. In examining market response to announcement of the inclusion into S&P 500 index and of the seasoned equity offering, Shleifer (1986) and Loderer et al. (1991) documented negatively sloped demand curves. Their findings do not support the notion of perfect price elasticity. In this section, based on Bagwell's measure, we compare the price elasticity of demand for a sample of Taiwanese IPOs with that on Eq. (2). Data on 50 Taiwanese IPOs, wherein price and allocation were determined by public auction, was analyzed. Data for the period from Dec. 1, 1995 to Oct. 31, 1998 was obtained from the Taipei Securities Dealers Association.

In Taiwan, when IPOs are underwritten via public auction, priority is given to the investors offering the highest bid. The computer will randomly select the successful bidder when there is excessive demand. This system is equivalent to the first price sealed bid auction. In an auction for multiple units, the system is known as multiple-price (discriminatory) sealed bid auction, or English Auction (you pay as you bid)[6]. Like IPOs auction discussed in Kandel et al. (1992), the investor bids in the IPOs shares in Taiwan were not capped by a maximum price, while investor bids in Bagwell's repurchases were bound by the maximum price each firm announced it was willing to pay for the stock.

Table 5 reports the price elasticity of the 50 IPOs using Bagwell's measure. On the average, the IPOs account for 60% of the firms' outstanding shares with a maximum of 10% and a minimum of 2.67%. Therefore, price elasticity is estimated for a range of 3% change in quantity (i.e.1%–4%, 4%–7%, 7%–10%). As shown in Table 5, Bagwell's measure of average price elasticity ranges from 0.34 to 0.52, and the average elasticity for all ranges is 0.48.

Table 5. The Mean of Price Elasticity Measures of the English Auction IPOs Demand Curves in Taiwan Stock Market Based on Bagwell's Approach.

The mean elasticity is calculated by using $\eta = \Delta q \Big/ \left[\sum_{i=1}^{N} \Delta p_i / N \right]$, where Δq is the average standardized quantity change of all firms, Δp_i is the percentage price change for firm i, and N is the number of sample firms. The price elasticity is estimated for a range of 3% change in quantity (i.e. 1%–4%, 4%–7%, 7%–10%). There is one firm in which the total shares offered is less than 4% of its outstanding shares and one firm in which the total shares offered is less than 7% but greater than 4% of its outstanding shares. Hence, the total number of firms is 48 for the entire range considered.

Quantity Change Interval (%)	Mean % of Price Change	Mean Elasticity η	No. of Firms
1–4	8.80	0.34	50
4–7	6.04	0.49	49
7–10	5.73	0.52	46
Sum 1–10	18.7	0.48	46

Source: The data was obtained from the Taipei Securities Dealers Association, the 50 IPOs underwritten via public auction in Taiwan stock market took place from December 20, 1995 to October 22, 1998.

Compared with the estimate by Bagwell (1.65), the price elasticity of demand for our sample IPOs is even smaller. In other words, this result seems to further support the hypothesis of finite price elasticity.

As mentioned in Section III, Bagwell's measure is sensitive to the number of shares outstanding. Therefore, comparing results using data on Taiwan's IPOs with those based on U.S. stock repurchase may not be meaningful. We reestimated price elasticity for the sample of Taiwan's IPOs using Eq. (2), and report the results in Table 6. Since the demand curve is negatively sloped, absolute values of estimated price elasticity are reported.

In Table 6, the average price elasticity according to Eq. (2) for all ranges of quantity change is 21.10, and the median is 20.60. These figures are comparable with the average price elasticity for U.S. firms repurchasing shares as reported in the previous section (22.44). For an average 9% shares offering in Taiwan, the stock price is expected to fall about 0.43% based on a price elasticity of 21.10. Therefore, the price elasticity effect is too small to be influential. For both sets of data (U.S. stock repurchase and Taiwan IPOs), we

Table 6. The Common Arc Elasticity Measures of the English Auction IPOs Demand Curves.

The price elasticity is estimated for a range of 3% change in quantity (i.e. 1%–4%, 4%–7%, 7%–10%). The common arc elasticity measure is calculated by $\varepsilon_i = \left| \frac{\Delta q_i}{\Delta p_i} \cdot \frac{p_i^m}{q_i^m} \right|$, where Δp_I and Δq_I are the percentage price change and the related standardized accumulative quantity change for firm I, respectively, and p_I^m (q_I^m) is the average of the beginning and ending prices (accumulative quantity) for firm i. n.a. means that the value is not available since the accumulative quantity for the individual firm is less than the minimum of the range.

	Quantity Change Interval (%)			
Name of Firms	1–4	4–7	7–10	Sum 1–10
Dahin	14.48	14.59	4.41	8.18
Toung Loong	23.42	68.82	18.77	21.00
Central Insurance	40.82	27.06	28.84	26.49
Inventec	9.20	47.71	n.a.	n.a.
Kee Tai Property	61.60	20.54	20.82	25.99
Chin.a. United Trust	17.39	20.09	14.65	13.62
Tonlin Department Store	64.26	14.17	4.32	11.80
Mustek Systems	19.81	34.65	13.15	15.87
Sheng Yu Steel	24.75	32.55	24.83	20.61
Les Enphans	14.88	17.04	20.14	12.58
Hotal Motor	28.62	41.57	46.29	26.12
Ultima Electronic	50.90	91.18	35.17	41.33
Fortune Electric	36.40	32.27	188.06	31.63
Tsann Kuen	64.05	19.80	12.43	21.93
Universal Tannery	19.14	13.31	14.99	12.88
Chia-Ta Textile	11.00	9.63	7.39	7.68
Ta Yih	32.61	18.79	47.87	22.37
Kye System	77.74	43.01	156.12	53.87
Shan-Loong Trans.	17.70	92.26	59.97	20.56
Askey Elec.	33.67	149.82	19.24	28.41
United	20.97	7.14	53.12	11.69
Zyxel Communition	3.13	n.a.	n.a.	n.a.
Fu-Ta Textile	8.04	4.29	31.88	5.73
Sakura Development	624.85	275.13	22.55	83.70
Silicon Integrated	46.69	165.87	7.82	22.09
Twinhead	30.63	40.49	87.75	28.88
President Chain Store	25.88	6.90	4.81	8.25
Wisher	29.48	7.04	11.49	20.88
Hin Nan Construction	11.49	4.68	7.23	6.09
Fu Sheng	35.36	63.38	16.73	25.72
Tyc Brother	48.60	14.27	13.41	18.35

Table 6. Continued.

Name of Firms	Quantity Change Interval (%) 1–4	4–7	7–10	Sum 1–10
Gold Circuit	30.36	17.09	407.41	22.63
Formosa International Hotel	23.04	2.66	infinite	6.38
Shin Hai Gas	49.85	11.48	45.82	20.65
Lingsen	26.46	30.21	10.68	16.97
Sumagh	35.40	15.54	infinite	23.72
Yung Shin Construction	132.76	5.89	1.58	5.05
Hwa Hsia Leasing	32.90	10.95	5.41	10.82
Broad Electronics	31.09	18.53	19.48	19.01
Kai Yih Ind.	27.79	46.76	30.06	24.57
Teapo Electronics	603.80	52.23	1.99	8.65
Pan Overseas	18.55	49.90	60.08	20.08
Giga-byte	59.16	66.19	24.53	38.14
Universal Microelectronics	25.62	27.27	3.62	9.67
Micro Star	33.63	49.95	12.75	20.01
Shinkao Gas	6.60	1.78	n.a.	n.a.
Tyn Tek Co.	44.08	54.47	5.76	16.62
Hitron Technology	73.23	54.82	10.04	26.61
Davision Inc.	111.04	5.73	1.08	3.82
Arima Computer	41.42	179.46	48.64	41.67
Mean	59.08	44.30	37.41	21.10
Median	30.86	30.63	18.77	20.60
Standard deviation	117.15	52.23	67.26	14.05
Minimum	3.13	2.66	1.08	3.82
Maximum	624.85	275.13	407.41	83.70

Source: The data was obtained from the Taipei Securities Dealers Association, the 50 IPOs underwritten via public auction in Taiwan stock market took place from December 20, 1995 to October 22, 1998.

document very large price elasticity even when the demand (supply) curve is not horizontal (vertical).

Table 7 shows that at the time of IPOs, Taiwanese firms faced very flat demand curves for their stocks. The absolute average slope is 2.78, and the median is 2.22. Intuitively, we always take a non-horizontal (non-vertical) demand (supply) curve as an indication of finite price elasticity. In fact, we should consider the impact of relative position in the curve on price elasticity. When arc elasticity measure is employed, we find that the average price elasticity is greater than 20 for both the U.S. repurchase sample and the Taiwan IPOs sample.

Table 7. The Schedule and Alternative Price Elasticity Measures of the English Auction IPOs Demand Curves.

The demand schedules of firms' IPOs shares are estimated by regressing bid price on the accumulated quantity: $P = a - bQ$, where b is the absolute slope of demand curve. Column (1) shows the slope coefficients b and the corresponding t-statistics. To estimate fixed price elasticity (E), we regress the log of accumulated quantity on the log of price: $\ln Q = \alpha - E \ln P$. The point elasticity shown in column (3) at the weighted winning bid price traded (Ew) is equal to the absolute value of $(1/b)(P_w/Q_w)$, where b is the estimated slope of the demand curve; P_w and Q_w is the price-quantity mix at the weighted average winning bid price traded. * (**) indicates significance at the 5% (1%) level.

Name of Firms	(1) b	T statistic	(3) E	T statistic	(5) Ew
Dahin	1.89**	37.57	10.78**	12.32	19.95
Toung Loong	1.14**	41.40	42.04**	10.19	17.6
Central Insurance	0.44**	55.47	33.63**	33.18	30.50
Inventec	4.13**	22.58	31.66**	19.45	8.64
Kee Tai Property	1.29**	54.62	31.87**	7.66	16.98
China United Trust	1.62**	45.42	54.15**	13.82	21.11
Tonlin Department Store	1.33**	60.16	5.62**	17.98	44.61
Mustek Systems	0.85**	54.71	36.66**	11.67	25.84
Sheng Yu Steel	1.33**	11.33	18.55**	24.49	8.82
Les Enphans	1.14**	25.03	21.0**	59.81	13.16
Hotal Motor	1.11**	49.69	35.11**	40.70	16.23
Ultima Electronic	1.13**	51.26	13.87**	19.71	18.69
Fortune Electric	0.66**	42.31	32.57**	36.38	13.32
Tsann Kuen	1.63**	81.36	47.31**	45.65	25.68
Universal Tannery	4.65**	71.99	28.11**	31.81	17.91
Chia-Ta Textile	3.57**	45.32	10.44**	47.03	9.57
Fu-Ta Textile	8.05**	39.29	17.02**	24.05	15.79
Sakura Development	3.46**	27.91	17.06**	13.50	7.09
Silicon Integrated	3.50**	31.67	44.68**	16.19	9.56
Twinhead	2.18**	36.91	27.77**	25.75	9.80
President Chain Store	5.11**	63.05	24.18**	21.94	9.51
Wisher	1.65**	34.38	27.13**	38.91	11.08
Hin Nan Construction	6.08**	34.13	13.72**	8.14	11.12
Fu Sheng	2.06**	32.46	38.37**	53.35	12.05
Tyc Brother	2.06**	28.35	40.30**	30.00	13.25
Ta Yih	2.50**	20.19	24.04**	39.81	10.85
Kye System	0.97**	35.15	42.29**	23.68	13.97
Shan-Loong Trans.	2.22**	18.31	18.01**	27.35	9.72
Askey	2.39**	23.39	36.01**	39.10	11.09
United	4.87**	49.25	24.19**	42.37	7.05
Zyxel Communition	5.49**	14.37	51.69**	3.39	10.61

Table 7. Continued.

Name of Firms	(1) b	T statistic	(3) E	T statistic	(5) Ew
Gold Circuit	0.21**	18.80	25.39**	10.78	6.66
Formosa International Hotel	5.79**	22.61	14.64**	20.53	5.44
Shin Hai Gas	1.15**	21.59	28.95**	18.23	11.23
Lingsen	2.25**	36.82	19.81**	19.04	12.25
Sumagh	3.10**	38.25	10.91**	13.43	10.94
Yung Shin Construction	6.30**	20.06	58.06**	7.22	8.86
Hwa Hsia Leasing	4.23**	31.35	46.94**	42.04	14.64
Broad Electronic	2.39**	30.21	20.41**	29.29	23.80
Kai Yih Inc.	1.05**	19.98	47.10**	18.24	20.35
Teapo Electronic	2.94**	11.44	9.26**	10.81	9.74
Pan Overseas	1.33**	27.76	27.33**	34.28	12.55
Giga-byte	2.22**	49.11	42.51**	26.42	13.80
Universal Microelectronics	0.36**	36.37	43.54**	30.92	15.08
Micro Star	2.82**	29.89	51.95**	44.00	12.42
Shin Kao Gas	9.06**	24.32	8.01**	20.37	9.96
Tyn Tek co.	3.13**	26.35	23.91**	16.97	17.69
Hitron Technology	3.48**	6.97	10.89**	4.19	8.83
Davision Inc.	5.43**	28.46	45.77**	4.47	8.14
Arima Computer	1.55**	14.22	38.41**	15.92	16.50
Mean	2.78		29.47		14.20
Median	2.22		32.07		12.33
Standard deviation	2.00		15.97		6.95
Minimum	0.21		6.96		5.44
Maximum	9.06		81.38		44.60

Source: The data was obtained from the Taipei Securities Dealers Association, the 50 IPOs underwritten via public auction in the Taiwan stock market took place from December 20, 1995 to October 22, 1998.

To verify the presence of large price elasticity, we calculated the constant elasticity (E), and the point elasticity for the quantity weighted winning price (Ew). The results are shown in columns (3) and (5) of Table 4. The average median values of E (Ew) for the U.S. repurchase sample are 15.66 and 17.05 (9.01 and 8.83) respectively. For the sample of Taiwan IPOs, the average and median E (Ew) are 29.47 and 32.07 (14.20 and 12.33). Using auction data, this study finds very large price elasticity for both the demand for and supply of common stocks. Our results indicate that price elasticity of common stock seems to be higher than we had perceived based on the findings of previous studies.

5. CONCLUSIONS

What factors determine price changes when firms repurchase shares or issue new shares is an unsolved issue. Based on a perfect market framework and the assumption of perfect substitutions among assets, finance theory suggests that shareholders' evaluation is homogenous and the price elasticity of demand (or supply) is infinite. As a result, price changes occurring at the time of an important announcement perceived to be caused by the information content embedded in the announcement. However, previous empirical studies do not support this contention. Many studies documented finite price elasticity of demand or supply. In particular, Bagwell (1992) reported that the average arc (point) price elasticity for a sample of 31 firms repurchasing their own shares is about 1.65 (0.68). For a 15% stock repurchase, the stock price should rise by approximately 9%. Comparing this with the actual price change (7.7%), Bagwell claimed that the dominant cause of price change is an upward-sloping supply curve, and his result is aligned with the hypothesis of shareholder heterogeneity. On the demand side, Shleifer (1986) and Loderer et al. (1991) analyzed the announcement of the inclusion into the S&P 500 and of the SEOs. They attributed the resulting positive and negative returns to the finite price elasticity of the demand curve. However, Kandel at al. (1999) reported that the demand schedules in Israeli IPOs under uniform price auction are quite elastic.

This paper illustrates that Bagwell's average price elasticity measure tends to underestimate the actual elasticity. This is mainly due to his elasticity measure's dependence on a normalized unit and his use of an inappropriate averaging procedure. Using Bagwell's data, we re-estimated the average price elasticity based on the concept of arc elasticity, after making an adjustment to the averaging procedure. Our estimate of price elasticity (22.44) is 13 times higher than that reported by Bagwell. In addition, this study examines the different auction samples from Taiwan IPOs that were conducted as discriminatory auctions, and the results document a very large price elasticity of demand curve, consistent with the findings of Kandel et al. (1999) from a non-discriminatory auction. This implies that quite large a price elasticity of demand (or supply) exists in both kinds of auction pricing systems and the shareholders in auctions are more homogeneous than was expected based on Bagwell's (1992) findings.

ACKNOWLEDGMENTS

Authors would like to thank the anonymous referee and conference participants of 1999 Financial Management Association in Orlando for helpful comments.

NOTES

1. As mentioned in Fama and Miller (1971), there is no perfect capital market in the real world. Therefore, empirical research should examine whether price elasticity is large enough rather than whether it is infinite.

2. In addition, according to Bagwell (1992) and Bernardo and Cornell (1997)'s arguments, the fact that bidders bid different prices imply that demand or supply is not perfect elastic.

3. Obviously, different auction mechanisms may lead to different investor behavior. Kandel's (1999) work examined samples from which auctions were conducted as uniform pricing scheme. It is of interest to examine if the result still holds under a different mechanism and to discuss why Bagwell's elasticity measure is so small.

4. Actually, in Bagwel1 (1992), the equation of the elasticity measure is not listed. Equation (1) is induced from the paper's calculation process.

5. When relative position on a supply curve is considered, since elasticity is jointly determined by the relative position on curve and relative quantity change to price change at each point, a constant elasticity curve can still exist.

6. The classification in auction theory has been discussed in Vickrey (1961) and Milgrom and Weber (1982). After investigating the treasury auction, both Reinhart (1992) and Stevens and Dumitru (1992) have proposed that Dutch auction rather than English auction should be adopted in order to avoid the winner's curse effect.

REFERENCES

Bagwell, L. S. (1991a). Share Repurchase and Takeover Deterrence. *Rand Journal of Economics*, *22*, 72–88.

Bagwell, L. S. (1991b). Shareholder Heterogeneity: Evidence and Implications. *American Economic Review*, *81*, 218–221.

Bagwell, L. S. (1992). Dutch Auction Repurchases: An Analysis of Shareholder Heterogeneity. *Journal of Finance*, *47*, 71–106.

Bernardo, A. E., & Cornell, B. (1997). The Valuation of Complex Derivatives by Major Investment Firms: Empirical Evidence. *Journal of Finance*, *52*, 785–798.

Fama, E. F., & Miller, M. H. (1972). *The Theory of Finance*. Hinsdale: Dryen Press.

Harris, L., & Gurel, E. (1986). Price and Volume Effects Associated with Changes in the S&P 500: New Evidence for the Existence of Price Pressures. *Journal of Finance*, *41*, 815–829.

Kandel, S., Sarig, O., & Wohl, A. (1999). The Demand for Stocks: An Analysis Of IPO Auctions. *Review of Financial Studies*, *12*, 227–247.

Loderer, C., Cooney, J., & Drunen, L. V. (1991). The Price Elasticity of Demand for Common Stock. *Journal of Finance*, *46*, 621–651.

Milgrom, P., & Weber, R. (1982). A Theory of Auctions and Competitive Bidding. *Econometrica*, *50*, 1089–1122.

Reinhart, V. (1992). An Analysis of Potential Treasury Auction Techniques. *Federal Reserve Balletin*, (June), 403–413.

Shleifer, A. (1986). Do Demand Curves for Stock Slope Down? *Journal of Finance*, *41*, 579–590.

Stevens, E. J., & Dumitru, D. (1992). Auctioning Treasury Securities. *Economic Commentary*, 15.

Vickrey, W. (1961). Counterspeculation, Auction, and Sealed Tenders. *Journal of Finance*, *16*, 8–37.

THE LONG-RUN PERFORMANCE AND PRE-SELLING INFORMATION OF INITIAL PUBLIC OFFERINGS

Anlin Chen and James F. Cotter

ABSTRACT

We show that private information as well as public information is important in revising the terms of the offer during the pre-selling period (or the waiting period) and that when the revealed private information is positive, the underwriter compensates the investors for this information by underpricing the issue more than when the information is negative. Even though the cost of compensating positive information is quite high, the issuer still benefits from the positive information in that the wealth transferred to the investors is smaller under underwriter's information acquisition activities. Furthermore, IPO long-run performance is negatively related to the positive information revealed during the waiting period and the underwriter prestige. Finally, IPO firms without receiving significant information during the waiting period survive longer after issuance.

1. INTRODUCTION

Theoretical models by Benveniste and Spindt (1989) and Benveniste, Busaba and Wilhelm (1996) examine the initial public offering (IPO) process when the underwriter plays an active role in marketing the offer to investors and setting

---

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 203–232.

ISBN: 0-7623-0887-7

the parameters of the offering. A growing body of empirical literature analyzes the observable activities of the underwriter and documents the factors that affect the relationship between the issuer and investors in an IPO. This paper examines the competing objectives of each party in the IPO process and empirically documents the role the underwriter plays in balancing the differing financial objectives of issuer, and investors. Specifically, we examine the costs and benefits of the underwriter's attempts to gather information to help issue a successful offering. Further, we show how the financial incentives of each participant are satisfied, either wholly or in part, given the other parties to the transaction. Finally, we also show the effects of market fads on the adjustment of pricing terms and on the short-run and long-run performance.

When a private firm is considering an equity offering, it must first interview underwriting firms and select the appropriate investment bank to provide advising and underwriting services. Once selected, the underwriter, in conjunction with other advisers (accounting, legal, etc.), formulates a strategy for the issuer to offer equity securities to the market. The first administrative step in the process involves preparing and filing a registration statement. In addition to detailed discussion of the issuer's business operations and financial condition, the underwriter and the issuer must agree on their first estimate of the offering terms, including the offer price range and number of shares that the issuer plans to offer. The period of time between filing of the registration statement and the actual offering of securities is called "waiting period".

During the waiting period, the underwriter assumes a special responsibility to market the offer to the investment community and attempt to determine the demand for the securities offered. Especially, for firm commitment offerings, the underwriters bear the proceeds risk in case the issuance may fail. Interestingly, the information the underwriter is seeking both from the market and from the investors will lead to changes in the terms of the offer that potentially may have an adverse affect on the investors' financial situation. For example, assume that the terms of the offer filed in the registration statement are attractive to the investors considering the offer. If the investors communicate their strong intent to purchase the offering, the underwriter then knows that the terms of the offer can be adjusted to increase the price and/or the number of shares offered.

In this paper, we examine the benefit of the issuer, and investors in an attempt to document the role of information acquisition activities prior to issuance and IPO long-run performance with respect to the information received during the waiting period. We find that issuers benefit from the underwriter's activities because the benefits of higher proceeds when demand is strong (positive information) are larger than when demand is weak (negative information). We also find that investors benefit from the activities of the underwriter because the initial return for IPOs when demand is strong is larger than otherwise. In addition, we find that investors further benefit when they give positive information because the issuer, on average, sells more shares, allowing the investors to earn a positive return on more shares. Furthermore, IPOs receiving positive information during the waiting period experience poor long-run performance. IPOs without receiving significant information during the waiting period survive longer. Thus, investments in IPOs are subject to market fads or speculative bubbles.

The remainder of this paper is organized as follows. In Section 2, we discuss the statutory scheme of underwriting. Section 3 discusses information acquisition during the waiting period and IPO long-run behavior. We present our measure of IPO long-run performance in Section 4. In Section 5, data description and descriptive statistics are presented. We report the methodologies and empirical results about the decision adjustment processes, the benefits to investors, and to the issuer and IPO long-run performance in Section 6. Finally, Section 7 concludes.

2. STATUTORY SCHEME FOR UNDERWRITING

The Securities Act of 1933 divides the registration process into three periods. The first period, the prefiling period, occurs before the filing of the registration statement with the Securities and Exchanges Commission (SEC). During this period, even though the issuing firm may have an agreement or understanding to issue and sell securities, the issuer cannot actually offer or sell a security, and cannot contact the prospective investors.

The second period, the waiting period, occurs between the filing date of the registration and the declared effective date by the SEC. During this period, offering a security is allowed, but selling is prohibited. However, indications of interest are obtained by the underwriters from the regular investors, and this information is used to price the IPO. In practice, adjustment in the final offer price is accompanied by revision in the number of shares issued. During the waiting period, the underwriter distributes information about the IPOs including the description of the offering and the security, the number of shares, and the expected offer price in the prospectus to the investors. If too long a period elapses in the waiting period, the financial statement will be considered stale, and the SEC or the investors may ask the issuer to provide more current financial statements.[1]

During the final, post-effective period which occurs after the registration statement is in effect, the underwriter begins to distribute the new offering to the public. Securities may only be sold in this period.

3. INFORMATION ACQUISITION AND IPO LONG-RUN PERFORMANCE

Numerous studies document the underpricing of IPOs at issuance.[2] Ibbotson, Sindelar and Ritter (1988) find an average day-one initial return of 16.37% for 8,668 IPOs issued from 1960 to 1987. It is, however, puzzling why the underwriter consistently underprices new issues. Ritter (1987) argues that the underpricing of IPOs is a kind of indirect cost of issuance. It is really costly for the issuers to underprice their IPO shares. Hence, the issuers would prefer to reduce the extent of underpricing at issuance. A way to reduce indirect issuance cost is to raise the offer price by realizing the interests of the investors on the new issues.

The Information Acquisition During the Waiting Period

Benveniste and Spindt (1989) develop a model that explicitly incorporates the information acquisition activities of the lead underwriters. Their model argues that underwriters can reduce the extent of underpricing by acquiring information or indications of interest (demand for shares of the IPO) from the "regular investors" prior to issuance. Investors that typically reveal their true demand will receive a greater number of shares of offerings with excess demand. IPO underpricing is a consequence of compensating informed investors for revealing truthful information. In the model, the optimal rule to establish the offer price and the number of shares issued is set to induce the regular investors to reveal positive information truthfully by premarketing the shares to the regular investors.

Benveniste and Spindt's model has several empirical implications. First, underpricing is directly related to the value of private information. This suggests that IPOs should be more underpriced if the revealed information causes significant changes in offer price, number of shares, or the effective date. Second, underpricing is directly related to the level of preselling activities during the waiting period. Third, underpricing is minimized if the underwriter allocates shares to the investors who reveal positive information. Therefore, shares should be distributed disproportionately to those who show great interest in the offering. Finally, underpricing is directly related to the level of interest in the premarket. That is, the stronger the demand for the IPO in the promotion process during the waiting period, the higher the extent of underpricing for the IPO.

Hanley (1993) examines some empirical implications of Benveniste and Spindt's (1989) model using 1,430 IPOs issued from 1983 to 1987. She investigates the relationship between the final offer price and the range of the offer price disclosed in the preliminary prospectus. Hanley shows that IPOs with final offer prices higher than the high limit of the expected offer price range are more underpriced while IPOs with final offer prices lower than the low-limit of the price range are less underpriced. One interpretation of this result is that underwriters are compensating the regular investors for revealing positive information about the offer during the pre-offer period. Thus, while investors have a larger return, the issuer receives a higher price for each share issued than was initially expected. The empirical results in Hanley (1993) are broadly supportive of the information acquisition model of Benveniste and Spindt (1989). Lowry and Schwert (2000) argue that the issuers would prefer to go public by reducing the initial return. Typically, the information acquisition activities during the waiting period can help reduce the initial return of IPOs at issuance. The public information about market condition should be incorporated into the offer price without any cost. However, the private information learned from the waiting period is only partially incorporated into the offer price and initial return. Lowry and Schwert (2000) show that new information revealed during the waiting period is related to the number of issues and the pricing of issuance. Cornelli and Goldreich (1999) find that the investment bankers distribute more shares to the regular investors who provide information about the demand for the issuance. These results support the justification for use of bookbuilding for IPO issuance.

Basically, issuers should be upset about the underpricing of offerings. Nevertheless, Benveniste and Spindt (1989), Hanley (1993) and Lowry and Schwert (2000) show that the underprcing is part of costs of acquiring information from the investors. Loughran and Ritter (2001) argue that the issuing firms acquiesce the underpricing of the new issues when the unknown demand of the offering becomes apparent during the waiting period. The reason why the issuers do not upset about leaving money on the table in IPOs is to retrieve the information of demand on IPOs.

The private information learned from the investors is not free. The issuers have to compensate the investors to induce truthful information. Has the introduction of information acquisition increased the efficiency of IPO issuance? Ljungqvist, Jenkinson and Welhelm (2000) point out that the direct costs of bookbuilding are twice as large as those for fixed price offers. They, however, also show that bookbuilding leads to less underpricing substantially.

The less extent of underpricing is the incentive for the issuers to acquire information from the investors during the waiting period. Nevertheless, with the sample of IPOs in New Zealand, Camp and Munro (2000) find that there is no difference in underpricing between fixed price and bookbuilding methods. By 1999, close to 80% of non-U.S. IPOs are issued using bookbuilding. It is really interesting to investigate whether the issuers benefit from information acquisition activities during the waiting period.

IPO Long-run Performance

Aggrawal and Rivoli (1990) examine the one-year holding period returns and one-year aftermarket returns on IPOs and find that the long-run performance of IPOs is worse than the market performance. Aggarwal and Rivoli (1990) argue that the IPO long-run under-performance may be due to fads or speculative bubbles in the early aftermarket stage. Ritter (1991) shows that the underpricing of IPOs appears to be a short-run phenomonen. The average three-year performance of IPOs is worse than the market performance and that of the matching firms. Furthermore, the younger companies that go public in the hot-issue years experience even worse long-run performance than average. Ritter argues that the negative long-run performance of IPOs may be due to fads in IPO market.

With Aggarwal and Rivoli (1990) and Ritter (1991), the long-run underperformance of IPOs implies that investors may be systematically too optimistic about the prospect of IPO firms. Brav and Gompers (1997) show that the venture-back IPOs outperform the nonventure-back IPOs using equally weighted returns. Under several benchmarks, venture-back IPOs do not suffer poor long-run performance while the small nonventure-back IPOs do underperform in the long run. However, Brav and Gompers (1997) also show that similar size and book-to-market firms not issuing equity perform as poor as IPO firms. This is inconsistent with Ritter (1991) who shows that IPOs underperform compared to those with same size and industry. Furthermore, Carter, Dark and Singh (1998) find that the underperformance of IPOs underwritten by prestigious underwriters is less severe than those underwritten by non-prestigious underwriters since prestige of underwriters conveys the credibility of IPO firms to the investors. Basically, IPOs backed by venture capitalist or prestigious underwriters underperform less in the long run.

Even though the long-run underperformance of IPOs is documented as a well-known anomaly, Brav, Geczy and Gompers (2000) find that underperformance is concentrated primarily in small size and low book-to-market offerings. Brav and Gompers (1997), Carter, Dark and Singh (1998) and Brav, Geczy and Gompers (2000) basically argue that small and risky IPOs with little credibility underperform more in the long run.

Aggarwal and Rivoli (1990) and Ritter (1991) basically attribute IPO long-run performance to fads. Ritter (1991) also presents evidence that IPOs with high initial returns tend to suffer poor aftermarket long-run performance. Obviously, the market fads around issuance affect the determination of the final terms of the offerings as well as the transaction prices in the aftermarket. During the waiting period, the issuers and underwriters receive public information about the market condition and acquire private information about the demand for the issuance. If issuers perceive strong market condition or possible fads for the new issues, the issuers will tend to raise the final offer price. When the upward adjusted offer price is due to market fads, the market fads would possibly also drive up the aftermarket prices for the new issues. Therefore, market fads for the issuance would result in the upward adjustment of offer price and the positive initial return even though the offer price is probably higher than the intrinsic value. On the other hand, the issuers have to compensate for the revealed positive information. Hence, partial adjustment implies that IPOs with final offer price adjusted above the expected price range experience high initial return. Basically, market fads and compensation for private information result in the positive relationship between the price adjustment during the waiting period and the initial return. It is difficult to tell market fads from partial adjustment phenomenon with the finding of positive relation between the price adjustment and initial return. However, the long-run performance of IPOs would be able to distinguish the partial adjustment from the market fads. The partial adjustment phenomenon due to compensation for private information does not lead to the connection between the price adjustment and long-run performance. The market fads, however, lead to positive relation between the price adjustment and initial return but negative relation between the price adjustment and long-run performance. Thus, we argue that the investigation of short-run, long-run performance and the price adjustment during the waiting period is important to detect the partial adjustment phenomenon.

4. THE MEASUREMENT OF LONG-RUN PERFORMANCE

IPO long-run underperformance has been well documented. The long-run abnormal return subsequent to a certain event implies that stock prices react to the public information with a long delay, which contradicts to market efficiency. However, the evidence of long-run abnormal return is probably due to the misspecification of the benchmarks used to measure the abnormal returns. Kothari and Warner (1997) and Barber and Lyon (1997) point out biases in the measurement of long-run returns for event studies. Barber and Lyon (1997) identify three reasons for the biases in test statistics in calculating long-run abnormal returns. The three biases include new listing bias, rebalancing bias and skewness bias. Kothari and Warner (1997) present additional source of misspecification for long-run abnormal returns. Kothari and Warner argue that parameter shift and issue of survival affect tests of abnormal return. Lyon, Barber and Tsai (1999) advocate the use of carefully constructed benchmark portfolios that are free of "new listing" and "rebalancing" biases to measure the long-run abnormal returns.

In this paper, we measure the long-run performance of IPOs by market-adjusted cumulative abnormal return (CAR) as follows:

$$AR_t = R_{it} - R_{mt},$$

$$CAR_T = \sum_{t=0}^{T} AR_t,$$

where R_{mt} is the equally-weighted market return on t-th day after issuance.

The market-adjusted CAR is a typical measure for the performance over a certain window period for event studies. The reasons for this measure include:

(1) In our IPO sample, we can avoid the new listing bias mentioned in Barber and Lyon (1997) because the new listing firms are included in our sample.
(2) Buy-and-hold measure may experience less severe rebalancing bias than CAR does. However, Brav, Gezcy and Gompers (2000) show that using buy-and-hold returns tend to magnify the underperformance of IPOs. Therefore, CAR is used to measure the long-run performance of IPOs rather than buy-and-hold return.
(3) This paper does not employ a positive model such as controlling for size and book-to-market. Loughran and Ritter (2000) argue that if a positive model is used, one is merely testing whether any patterns existing are captured by other known patterns. Brav (2000) also find that the Fama-French (1993) model is inconsistent with the observed long-run price performance of IPOs.
(4) Equally-weighted portfolio rather than value-weighted portfolio returns are used to calculate CAR because Loughran and Ritter (2000) show that value-weighted portfolio returns have low power to identify abnormal returns for events.

5. DATA SOURCE AND DESCRIPTION

Data for this study consists of 1,456 initial public offerings issued in the U.S. collected from *Bloomberg Financial Markets* (hereafter, *Bloomberg*) over the period from 1993 to 1994. *Bloomberg* provides detailed data about the offering including the expected offer price range, expected number of shares disclosed in the preliminary prospectus, the final offer price, the final number of shares issued, the lead underwriter and co-underwriters, the date when the registration statement is filed, the date of issuance, and the industry of the IPO. Unit offerings, mutual fund offerings, best efforts offerings[3] and American Depository Receipts (ADRs) are excluded from the data set. As a result, our sample consists of 955 firm commitment offerings issued in 1993 and 1994.[4] However, young and risky IPO firms may die soon after issuance. In our sample, only 802 IPO firms survive three years after issuance.

The closing prices on the first trading days are collected from *OTC, ASE, or NYSE Daily Stock Price Record*.[5] With the closing price on the first trading day and the final offer price, we can calculate the initial return for each IPO. The initial return is used to measure the compensation for the revealed information and issuer's cost of compensating information. We define the initial return of an IPO as follows:

$$R_1 = \frac{p_1 - p_o}{p_o},$$

where p_o is the final offer price and p_1 is the closing price on the first trading day.

The length of the waiting period is equal to the number of days from when the issuer files its first registration to when the registration statement becomes effective and the issue is priced. The relative price range is defined as the high limit of the expected price range divided by the low limit of the range, i.e. p_H/p_L, where p_H is the highest price of the offer price range; p_L is the lowest price of the offer price range. The price adjustment is the final offer price divided by the expected offer price, which is the midpoint of the price range, i.e. p_o/p_E, $p_E = (p_H + p_L)/2$. The number of shares adjustment is equal to the final number of shares divided by the expected number of shares, i.e. q_o/q_E where q_o and q_E are the final and expected number of shares issued, respectively. The proceeds adjustment is measured as the final total proceeds (the product of the final offer price and the number of shares issued) divided by the expected total proceeds (the product of the expected offer price and the expected number of shares), i.e. p_oq_o/p_Eq_E. The adjustment behavior in price and number of shares are the measurements of the revealed information and its effects.

Information about changes in the level of the overall market during the waiting period also may influence the offer price adjustment process and the adjustment in the number of shares. Because most of the IPOs are issued and traded in the National Association of Securities Dealers (NASD) exchange, we employ the National Association of Securities Dealers Automated Quotation (NASDAQ) index as the proxy for the market index. The overall stock market change over an IPO waiting period is defined as:

$$R_{\text{market}} = \frac{\text{NASDAQ index on the day before the offer date}}{\text{NASDAQ ndex on the filing date}} - 1,$$

NASDAQ indices are taken from the CRSP data base. Besides the overall market change, we also calculate an industry index for each individual new offering. The industry index change over the waiting period for each individual IPO is defined as the equally weighted holding period return for the securities in CRSP tapes with the same two-digit SIC code[6] as that of an IPO, from the filing date of the new offering to its offer date. That is,

$$R_{\text{industry}} = \frac{\sum_{i=1}^{n} \prod_{j=\text{filing date}}^{\text{offer date}-1} (1 + R_{ij})}{n} - 1,$$

where n is the number of securities with the same two-digit SIC code as that of an IPO, and R_{ij} is the daily return of security i on day j in CRSP. The overall market index and industry index are employed as the public sources that affect the price and number of shares adjustment processes.

The lead underwriter is the representative for the underwriting syndicate of the new issuance. In this paper, we argue that the prestigious underwriters likely will underwrite more offerings where the issuer proceeds will be larger, than those underwritten by non-prestigious underwriters. Therefore, we measure underwriter prestige by calculating the market share of the lead underwriter based on IPO proceeds in 1993 and 1994. We form a measure that is equal to:

$$\text{Underwriter prestige} = \frac{\sum_{i=1}^{n_j} proceeds_i}{Total\ IPO\ proceeds\ in\ 1993\ and\ 1994},$$

where $proceeds_i$ is the offer price times the number of shares offered for IPO_i; n_j is the number of IPOs in the sample underwritten by underwriter_j. This measure is similar to that used by Beatty and Ritter (1986) and Megginson and Weiss (1991).[7]

Most of the IPO firms are young and risky. They may not be able to survive long after issuance. This paper examines the relation between the surviving rate and the information revealed during the waiting period. The surviving rate at time t of IPOs is defined as:

$$\text{Surviving rate at time t} = \frac{\text{number of IPO firms at time t after issuance}}{\text{number of IPO firms at issuance}}.$$

Table 1 provides the mean, and median of IPO characteristics for the full sample of 955 offerings issued in the years 1993 and 1994 and the subsamples based on the proceeds adjustment. In the sample, we have 300 (31.41%) IPOs with final proceeds below the expected proceeds range, 397 (41.57%) IPOs with the final proceeds within the range, and 258 (27.02%) IPOs with final proceeds above the expected range.[8] For the full sample, panel A indicates that the mean initial return is 9.11% with a median of 4.25%. The higher mean initial return relative to the median implies that the distribution of the initial return is positively skewed, which is consistent with Ruud (1993). Although the mean initial return is positive, it is somewhat lower than found in other studies. For the subsamples, we find that IPOs with final proceeds higher than the expected range are more underpriced, while IPOs with final proceeds below the range are less underpriced. This result implies that when the underwriter receives the information of strong demand for the offerings, the compensation for the information is higher.

The average offer price for the full sample is $12.73 and most of the new offerings are priced between $10 and $20. IPOs with final proceeds higher than the range have higher offer prices. The mean relative price range (p_H/p_L) is equal to 1.15 and is wider for the IPOs with proceeds within the proceeds range. Of course, IPOs with wider proceeds range are more likely to have final proceeds within the range. The average total proceeds of IPOs for the full sample is $65.46 million but the median value is less than $30 million, implying that some IPOs are quite large in size. On average in 1993 and 1994, IPO market share of the lead underwriter is 3.43%. IPOs underwritten by more prestigious underwriters are more likely to experience proceeds adjustment which implies that prestigious underwriters are more capable of acquiring information from the investors under the condition that prestigious underwriters can price offerings more accurately.

In panel B, we outline the variables associated with the waiting period. For the full sample, an IPO is issued to the market 65.6 days after the registration statement is filed. IPOs with final proceeds below the expected range have longer waiting periods while IPOs with final proceeds higher than the range have shorter waiting periods. This phenomenon supports the argument that

Table 1. Descriptive Statistics on the Characteristics of the IPOs by Total Proceeds Adjustment.

Mean descriptive statistics on the initial return, length of the waiting period, expected price range, final offer price, final number of shares, market change, industry change, and underwriter prestige for the entire sample and the subsample classified by the proceeds adjustment behavior. The data for the sample of 955 IPOs from 1993 to 1994 are collected from *Bloomberg*. Medians are reported in the brackets.

Variable	Full sample	Total proceeds below the range	Total proceeds within the range	Total proceeds above the range
Panel A: IPO characteristics				
Number of offerings	955	300	397	258
Percent of sample		31.41%	41.57%	27.02%
Initial return	9.11%	2.54%	8.22%	18.11%
	[4.20%]	[0.60%]	[4.50%]	[13.00%]
Offer price	$12.73	$11.02	$12.32	$15.35
	[$12.00]	[10.00]	[12.00]	[15.00]
Relative offer price range[a]	1.15	1.14	1.17	1.14
	[1.17]	[1.15]	[1.17]	[1.15]
Total proceeds (million)	$65.46	$53.29	$65.19	$80.02
	[$30.00]	[23.41]	[25.20]	[40.15]
Underwriter prestige[b]	3.43%	3.42%	2.84%	4.33%
	[1.24%]	[1.49%]	[0.86%]	[2.09%]
Panel B: Price adjustment process				
Number of days of the waiting period	65.60	74.11	63.60	58.80
	[57.00]	[63.00]	[55.00]	[54.00]
Ratio of price adjustment[c]	0.98	0.82	0.99	1.14
	[1.00]	[0.82]	[1.00]	[1.12]
Ratio of shares adjustment[d]	1.02	0.93	1.01	1.14
	[1.00]	[1.00]	[1.00]	[1.10]
Ratio of proceeds adjustment[e]	1.00	0.75	1.00	1.29
	[1.00]	[0.78]	[1.00]	[1.22]
Panel C: Non-compensated information during the waiting period				
Return of market index[f]	1.10%	1.63%	1.17%	0.49%
	[0.80%]	[1.31%]	[0.60%]	[0.59%]
Return of industry index[g]	1.81%	0.03%	1.35%	4.66%
	[1.84%]	[0.43%]	[1.64]	[4.42%]

Table 1. Continued.

Variable	Full sample	Total proceeds below the range	Total proceeds within the range	Total proceeds above the range
Panel D: Long-run behavior after issuance				
One-year CAR[h]	–7.88%	–6.87%	–9.48%	–6.60%
	[–10.11%]	[–11.37%]	[–11.67%]	[–4.47%]
Two-year CAR	–14.00%	–11.19%	–16.56%	–13.20%
	[–22.88%]	[–24.80%]	[–22.77%]	[–22.72%]
Three-year CAR	–18.77%	–12.28%	–15.69%	–30.86%
	[–26.24%]	[–21.50%]	[–24.56%]	[–43.71%]
Surviving rate three years later[i]	83.96%	77.70%	88.14%	84.85%

[a] The high limit of the price range divided by the low limit.
[b] The IPO market share of the lead underwriter during 1993 and 1994.
[c] The final offer price divided by the midpoint of the price range.
[d] The final number of shares issued divided by the expected number of shares offered.
[e] The final proceeds divided by the expected proceeds.
[f] Return of market index over the waiting period.
[g] Return of industry index over the waiting period.
[h] CAR is measured by the IPO raw return adjusted by the return of equally weighted market index.
[i] Surviving rate is defined as the number of IPOs three years after issuance divided by the number of IPOs in the sample.

once the underwriters receive positive information from the investors, they try to offer the issuance as soon as possible. After the waiting period, the final offer price is about 2% lower than the expected offer price and the number of shares issued is 2% higher than the expected number of shares offered as outlined in the preliminary prospectus. The final total proceeds, however, are similar to the expected total proceeds outlined in the registration statement. For the subsamples, IPOs with final proceeds adjusted above have higher adjustment ratios of offer price and number of shares offered. That is, the offer price and number of shares offered are typically adjusted in the same direction.

In panel C, we report the descriptive statistics for the “non-compensated” public information. For the proxies of the public information sources, the returns on the overall market and on the industry portfolio are equal to 1.10% and 1.81% from the filing date to the offer date, respectively. The proceeds adjustments are not consistent with the overall market return behavior. Rather, that the proceeds adjustments agree with industry return pattern implies that the underwriter adjusts up the offer price and number of shares when the industry booms during the waiting period.

Panel D shows that 83.96% of the IPOs still survive three years after issuance.[9] However, IPOs with final proceeds adjusted outside the expected proceeds range are more likely to die within three years after issuance. Further, the cumulative abnormal return (CAR) patterns indicate that IPO long-run performance is poor. The one-year, two-year, and three-year CARs are –7.88%, –14.00% and –18.77%, respectively. IPOs with final proceeds adjusted upward experience even poorer long-run performance than others. The average three-year CARs for IPOs with final proceeds adjusted below and above the expected range are –12.28% and –30.86%, respectively. These results are consistent with the view that the adjustments of the pricing terms during the waiting period are due to market fads.

Generally speaking, the descriptive statistics provide preliminary evidence about the information acquisition activities during the waiting period, the compensation and strategy of compensation for the revealed information, and the long-run performance behavior of IPOs.

6. METHODOLOGIES AND EMPIRICAL RESULTS

Decision Adjustment and Strategies to Compensate for the Revealed Information

The terms of the offer of an initial public offering outlined in the initial registration statement may be quite different from those at issuance. There may be changes in the final offer price, final number of shares issued, and thus the final total proceeds of the offering.

In Fig. 1, we report the mean initial return (the compensation for the revealed information) of the groups of IPOs based on the price and number of shares adjustments. From Fig. 1, we can see that IPOs with final offer price above the offer price range are more underpriced (have a larger initial return), while IPOs with final offer price within or below the offer price range are less or not underpriced. These results imply that the positive information is compensated by the underpricing. Figure 1 also shows that the effect of the number of shares adjustment combined with the price adjustment has an even stronger effect on underpricing. The group of IPOs with final offer price below the offer price range and the number of shares downward adjusted is the only group where the initial return is insignificantly different from zero. This is consistent with the hypothesis that negative information needs no compensation.[10] For all other

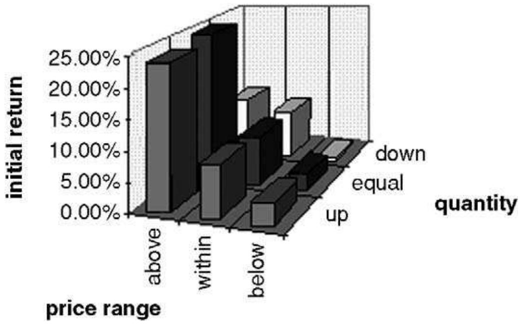


Fig. 1. Initial returns of IPOs.

The initial return of IPOs based on the price and number of shares adjustment process. The 955 offerings are broken into 9 groups according to the final offer price above, within or below the expected offer price range and the final number of shares issued higher than, equal to or lower than the expected number of shares offered. The numbers of IPOs in each category are 88, 127, 64, 63, 326, 118, 13, 86, 70 for (below, down), (below, equal), (below, up), (within, down), (within, equal), (within, up), (above, down), (above, equal) and (above, up), respectively.

groups, the mean initial returns are significantly larger than zero. If the underwriter decreases the offer price and raises the number of shares issued after receiving negative information, investors will have an incentive to exaggerate the negative information or to reveal negative information more frequently than the truth. Therefore, IPOs with final offer price below the offer price range and the number of shares higher than expected still encounter positive initial returns. These results imply that the underwriter increases the offer price and number of shares when receiving positive information and decreases offer price and number of shares when receiving negative information.

Changes in the terms of the offer during the waiting period may also be affected by other factors. The reasons for these changes can be attributed to public information about the overall market return or the return of firms in the same industry, as well as private information related to investors' demand for the new offering. Additionally, factors such as the riskiness of the offer and the characteristics of the underwriter may also lead to adjustments in the terms of the offer. Offers with higher *ex ante* uncertainty are more likely to encounter decision adjustments after the waiting period. The offer size of an IPO is frequently used as a proxy for *ex ante* uncertainty. For instance, Beatty and

Ritter (1986) use the reciprocal of the proceeds as a proxy for *ex ante* uncertainty of an IPO.

In addition, the reputation of the underwriter may have an effect on changes in the terms of the offer during the waiting period. The prestige of the lead underwriter may have two effects on the terms of the offer during the waiting period. Generally, Carter and Manaster (1990) show that prestigious underwriters have lower underpricing, suggesting that they can measure the offer terms more precisely on the filing date and decrease the likelihood of changes in the offer. On the other hand, a prestigious lead underwriter may have more relationships with regular investors that enhance the information acquisition activities, leading to an increase in the likelihood of changes in the terms of the offer. In this section, we present the following multivariate regression model to investigate the adjustments in the IPO terms.

Adjustment in the terms of IPO = f(market return, industry return, offer size, underwriter prestige)

where

Adjustment in the terms of IPO	:	price adjustment, number of shares adjustment, and total proceeds adjustment
Market return	:	return of market index over the waiting period
Industry return	:	return of industry index over the waiting period
Offer size	:	the logarithm of the expected total proceeds (product of the expected offer price and expected number of shares issued)
Underwriter prestige	:	the market share of the lead underwriter over the period of 1993 and 1994

From the regression results presented in Table 2, the return of the industry portfolio is positively related to adjustments in offer price, number of shares offered, and total proceeds. The return of the overall market portfolio, however, does not enter the regressions significantly. These results imply that the overall market return is not an important factor for adjustments in the terms of the IPOs, while industry return is quite important.[11]

Table 2 also shows that the expected offer size is negatively related to adjustments of the terms of the IPO during the waiting period. Therefore, smaller offers are more likely to have increases in the offer price and number

Table 2. Regression Analyses of Decisions Adjustment.

Regression analyses of price adjustment, number of shares adjustment and proceeds adjustment on the public information to measure the effect of public on the final decision terms. The market index and the industry index are employed as proxies for public information revealed during the waiting period. The data for the sample of 955 IPOs issued in 1993 and 1994 are collected from Bloomberg. The p-values are reported in the parentheses. ***, ** and * represent the significance test at 1%, 5%, and 10% levels, respectively.

	Dependent variables		
Explanatory variables	Model 1: Price adjustment[a]	Model 2: Quantity adjustment[b]	Model 3: Proceeds adjustment[c]
Intercept	0.516	0.554	0.093
	(0.001)***	(0.000)***	(0.673)
Return of market index[d]	–0.067	0.022	–0.048
	(0.517)	(0.821)	(0.724)
Return of industry index[e]	0.577	0.489	1.061
	(0.000)***	(0.000)***	(0.000)***
Size of offer[f]	–0.018	–0.014	–0.049
	(0.001)***	(0.002)***	(0.000)***
Underwriter prestige[g]	0.505	0.457	1.003
	(0.001)***	(0.001)***	(0.000)***
Pr > F	0.0001***	0.0001***	0.0001***
R-square	6.35%	5.50%	9.64%
Correlation coefficient between price adjustment and number of shares adjustment = 0.142 (0.000)***			
Correlation coefficient between the residual terms of model 1 and model 2 = 0.151 (0.000)***			

[a] Price adjustment is measured as the final offer price divided by the expected offer price.
[b] Quantity adjustment is the final number of shares issued divided by the expected number of shares offered.
[c] Proceeds adjustment is defined as the final total proceeds divided by the expected total proceeds.
[d] Return of market index is the holding period return for NASDAQ index over the waiting period.
[e] Return of industry index is the holding period return of the average return of firms in the same industry over the waiting period.
[f] Size of offer is measured by the logarithm of the expected total proceeds.
[g] Underwriter prestige is the market share of the lead underwriter in 1993 and 1994.

of shares issued. Table 2 outlines a positive relation between the adjustment process and underwriter prestige, suggesting that prestigious underwriters are more likely to revise the terms of the IPO upward, or that the terms outlined in the registration statement are more conservative.

The correlation coefficient between the price adjustment and number of shares adjustment is positive and significant (correlation coefficient = 0.142, p-value = 0.000). However, the positive relation between the price adjustment and number of shares adjustment may be attributed to either the public information and/or private information. To measure possible correlation between the price adjustment model and the quantity adjustment model, we also measure the correlation coefficient between the residual terms of the regression models with price adjustment and number of shares adjustment as dependent variables.[12] These residual terms are independent of the public information and might be considered as something related to the private information plus firm specific characteristics. The correlation coefficient of the residual terms reported in Table 2 is significantly positive (correlation coefficient = 0.151, and p-value = 0.000). We assume that the firm specific characteristics of IPOs are independent with one another. Furthermore, the residual terms should be positively related to private information.[13] The positive correlation coefficient between the error terms of the decision adjustment regression implies that the price adjustment and number of shares adjustment are affected by the private information in the same direction. The positive relation between the price adjustment and number of shares adjustment as well as the positive relation between the residuals from the regression models are consistent with the hypothesis that underwriters are pursuing strategies to compensate investors for the revealed information.

Investors' Compensation for Providing Information

The information provided by investors will cause the underwriter to adjust the offer terms of an IPO. If the information revealed during the waiting period causes significant changes in offer terms, the investors who provide this information should receive compensation consistent with its value. Since the effect of information cannot be measured explicitly, we employ three proxies for the effect of information: one is the time length of the waiting period; another is the relative offer price range disclosed in the preliminary prospectus; the third is the absolute value of the total proceeds adjustment. For the time length of the waiting period, we expect that if the underwriter receives positive information from the investors or the market, he will issue the offering as soon as possible, i.e. the waiting period is shorter. For the relative offer price range, we argue that if the underwriter is more confident on the offer price of an IPO, the price range will become narrower, and thus the outside information will be less important. For the proceeds adjustment, obviously, we expect that important information will cause the underwriter to adjust the pricing terms significantly and hence should be compensated more. In addition, four control variables are used to control for factors known to affect IPO pricing, including return of market portfolio, return of industry portfolio, size of offer, and underwriter prestige.

Hence, the relation between compensation to investors and the effect of information is measured by the following regression model:

Investor compensation = f(time length of waiting period, relative offer price range, proceeds adjustment, industry return, market return, offer size, underwriter prestige),

where investor compensation is the initial return; time length of waiting period is the number of days from the filing date to the pricing date; relative offer price range is the upper limit of the offer price range divided by the lower limit (p_H/p_L); magnitude of the proceeds adjustment equal to ($|p_o q_o - p_E q_E|/p_E q_E$); offer size is the logarithm of the final total proceeds. Results from this model are presented in Table 3.

In Table 3, we find that each proxy for the effect of information itself is significant at least at 10% level. The time length of the waiting period is significantly negatively related to the compensation (coeffieicnt = –0.030, p-value = 0.009 in column one and coefficient = –0.027, p-value = 0.020 in column four), suggesting that the longer the waiting period, the more likely that the revealed information is negative and that compensation for IPOs with longer waiting period is small. The relative offer price range is significantly positively related to compensation (coefficient = 0.118, p-value = 0.080 in column two). This suggests that when the underwriter is unsure about the intrinsic value of the offering, outside information becomes more important. As a consequence, the compensation for outside information is higher. The third proxy, the magnitude of the total proceeds adjustment, is significantly positively related to compensation (coefficient = 0.067, p-value = 0.019 in column three; coefficient = 0.059, p-value = 0.020 in column four). The information provided by investors causes the final total proceeds to deviate from the expected total proceeds. The larger the deviation, the stronger the effect of the information. When we include all the proxies for the effect of information into the regression, we still find that the compensation for information is negatively related to the time length of the waiting period,

Table 3. Regression Analyses of the Effect of Information.

Regression analyses of the effect of revealed information on the initial return controlling for the market index, industry index, offer size and underwriter prestige. The proxies for the revealed information are the length of the waiting period, the relative offer price range, and the magnitude of proceeds change. The data for the sample of 955 IPOs issued in 1993 and 1994 are collected from Bloomberg. The p-values are reported in parentheses. ***, ** and * represent the significance test at 1%, 5%, and 10% levels, respectively.

Explanatory variables	Dependent variable: initial return[a]			
Intercept	–0.460	–0.674	–0.543	–0.562
	(0.002)***	(0.000)***	(0.000)***	(0.001)***
Number of days of the waiting period[b]	–0.030			–0.027
	(0.009)***			(0.020)**
Relative offer price range[c]		0.118		0.085
		(0.080)*		(0.240)
Magnitude of proceeds adjustment[d]			0.067	0.059
			(0.019)**	(0.020)**
Return of market index[e]	0.179	0.162	0.161	0.175
	(0.054)*	(0.084)*	(0.074)*	(0.056)*
Return of industry index[f]	0.521	0.496	0.487	0.515
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Offer size[g]	–0.015	–0.007	–0.013	–0.014
	(0.041)**	(0.092)*	(0.078)*	(0.071)*
Underwriter prestige[h]	0.223	0.241	0.195	0.235
	(0.105)	(0.100)	(0.173)	(0.101)
Pr > F	0.0001***	0.0001***	0.0001***	0.0001
R-square	4.51%	4.14%	4.43%	5.22%

[a] The initial return is defined as the

$$\frac{p_1 - p_o}{p_o},$$

where p_1 and p_o are the closing price on the first trading day and offer price, respectively.

[b] The waiting period is from the date of filing registration to the effective date.

[c] The high limit of the expected offer price range divided by the low limit of the price range.

[d] The absolute value of the difference of final proceeds and the expected proceeds divided by the expected proceeds.

[e] The return of NASDAQ index over the waiting period.

[f] The average return of the firms of the same industry over the waiting period.

[g] The logarithm of the total proceeds (the product of offer price and number of shares issued).

[h] The market share of the lead underwriter based on the IPO proceeds in 1993 and 1994.

positively to relative offer price range, and positively to the magnitude of proceeds change. However, the relation between the initial return and relative offer price range is not significant.

For the control variables in the regression model, we can see that when the market and/or the industry grow up, the initial return becomes higher. This suggests that when the market return and/or the industry return is high over the waiting period, the offer is more likely to be traded at a price higher than the offer price. The size of offer is negatively related to the initial return. This result is consistent with the argument that the initial return is directly related to the uncertainty of the offering.[14] The underwriter prestige is positive but not significant to the initial return.[15]

Issuer's Benefit from Information Acquisition

As discussed earlier, investors have an incentive not to reveal positive information to the underwriter unless they are properly compensated for it. Figure 1 shows that IPOs with final offer price above the price range and final number of shares adjusted upward are highly underpriced ($\bar{R}_1 = 23.38\%$). That is, it is very costly to acquire information from the investors. With such high compensation, the information acquisition from the investors during the waiting period may not be worthwhile. To investigate if the issuer benefits from underwriter information acquisition activities, we compare the wealth transferred to investors when the underwriter acquires information and when underwriter does not. The wealth transferred to the investors is measured by the extent of underpricing and the number of shares offered.

To measure the issuer's benefit from information acquisition, we use the following equation:

$$\text{issuer's benefit} = [(p_o - p_1) \times q_o] - [(p_E - p_1) \times q_E].$$

where p_1 is the closing price on the first trading day; p_o and q_o are the final offer price and number of shares issued, respectively, and p_E and q_E are the expected offer price and number of shares outlined in the preliminary prospectus. In the equation, we assume that the final offer price and number of shares offered are determined based on the information acquired during the waiting period. However, if the underwriter does not acquire information from the investors, we assume that the underwriter sets the final offer price equal to the expected offer price and the final number of shares offered equal to the expected one.

We divide the sample into nine groups based on the price and number of shares adjustments. Figure 2 indicates that even though the compensation for the revealed positive information is high, the amount of the issuer's benefit still,

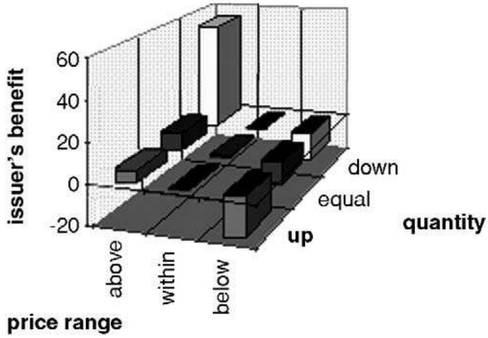


Fig. 2. Amount of Issuer's Benefit from Information Acquisition (unit: million).

Amount of issuer's benefit based on the price and number of shares adjustment process. The 955 IPOs are broken into 9 groups according to the final offer price above, within or below expected price range and the final number of shares higher than, equal to or lower than the expected number of shares.

is higher when its underwriter acquires positive information from the investors. For positive information, the issuer's benefit is maximized by increasing the price and decreasing the number of shares issued. Since increasing offer price and decreasing number of shares is simply for the issuer's benefit, the investors will tend to reserve their positive information rather than to reveal it.

For negative information, Fig. 2 shows that issuer's benefit is lower when the underwriter acquires information. The issuer prefers the underwriter not acquire information from investors during the waiting period if the expected offer price is set higher than the aftermarket price. In firm commitment offerings, the issuer will receive the total proceeds whether the offer is successful or not. The revealed negative information implies that the expected offer price is too high. Hence, if the underwriter sets the final offer price at the expected offer price, the underwriter loses money and the issuer gains profit from overpricing of the offering. In this case, acquiring information is for the underwriter's own benefit by helping avoid failure.

IPO Long-Run Performance with Respect to Information Revealed During the Waiting Period

Aggarwal and Rivoli (1990) and Ritter (1991) show that the long-run performance of IPOs is worse than the market performance. They argue that the poor long-run performance of IPOs may be attributed to fads or speculative bubbles in the IPO market. Basically, these findings are consistent with the overreaction phenomenon in the IPO market. Therefore, the overreaction phenomenon in the IPO market may shed light on the adjustment of IPO pricing terms during the waiting period. We thus argue that the IPO long-run performance should be negatively related to the proceeds adjustment during the waiting period.

Figure 3 indicates that the long-run performance of IPOs based on the proceeds adjustment. In Fig. 3, we can see that IPOs experience poor long-run performance regardless of the proceeds adjustment during the waiting period. However, if we argue that the market overreacts in the IPO market, IPOs with inferior information revealed during the waiting period should experience less poor long-run performance. We thus set up the following regression model to examine the IPO long-run performance with respect to the adjustments of the pricing terms during the waiting period.

CAR = f(time length of waiting period, relative offer price range, proceeds adjustment, offer size, initial return, underwriter prestige).

The results of this regression model are reported in Table 4. Table 4 shows that proceeds adjustment is not significantly related to one-year or two-year CARs. Nevertheless, proceeds adjustment is significantly negatively related to the

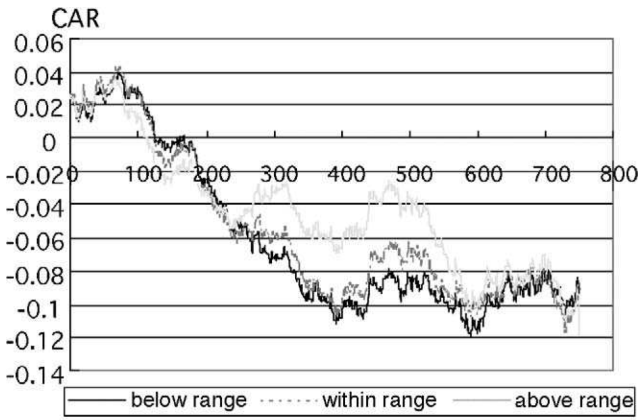


Fig. 3. The CARs of IPOs with Respect to the Proceeds Adjustment.

The cumulative abnormal returns of IPOs with final proceeds adjusted below, within, and above the expected proceeds range. CARs are measured by the IPO raw returns adjusted by the return of equally weighted market index.

Table 4. Regression Analyses of the IPO Long-run Performance.

Regression analyses of the one-year, two-year and three-year CARs on the number of days during the waiting period, relative offer price range and proceeds adjustment with the proceeds size, initial return and underwriter prestige as control variables. The CARs are measures by the IPO raw return adjusted by equally-weighted market return. The data for the sample of 853 IPOs issued in 1993 and 1994 are collected from Bloomberg and CRSP. The p-values are reported in parentheses. ***, ** and * represent the significance test at 1%, 5%, and 10% levels, respectively.

	Dependent variable: CAR		
Explanatory variables	One-year	Two-year	Three-year
Intercept	–0.216	–0.566	–0.038
	(0.484)	(0.226)	(0.948)
Number of days during the waiting period[a]	–0.028	0.002	0.018
	(0.446)	(0.976)	(0.798)
Relative offer price range[b]	0.286	0.521	0.170
	(0.164)	(0.097)*	(0.657)
Proceeds adjustment[c]	–0.053	–0.120	–0.281
	(0.419)	(0.233)	(0.023)**
Proceeds size[d]	–0.000	–0.001	–0.001
	(0.073)*	(0.021)**	(0.006)***
Initial return[e]	0.060	0.165	–0.251
	(0.568)	(0.296)	(0.196)
Underwriter prestige[f]	–0.003	–0.012	–0.018
	(0.383)	(0.046)**	(0.013)**
Sample size	885	853	802
Pr > F	0.024**	0.000***	0.000***
R-square	1.70%	3.60%	5.41%

[a] The waiting period is from the date of filing registration to the effective date.
[b] The high limit of the expected offer price range divided by the low limit of the price range.
[c] The final proceeds divided by the expected proceeds.
[d] The return of NASDAQ index over the waiting period.
[e] The logarithm of the total proceeds (the product of offer price and number of shares issued).
[f] The market share of the lead underwriter based on the IPO proceeds in 1993 and 1994.

three-year CAR. These results imply that the market overreacts in the IPO market and corrects the overreaction three years after issuance. Two years is not long enough for the market to revise overreaction phenomenon in the issuance market. However, the revision is getting stronger as time passes by (p-values = 0.419 and 0.233 for one-year and two-year CARs, respectively). Further, the p-value of the proceeds adjustment with three-year CAR as the dependent variable is 0.023. Thus, the more significant the proceeds adjustment in the issuance market, the poorer the IPO long-run performance. Moreover, the IPO long-run performance is negatively related to underwriter prestige (p-values = 0.046 and 0.013 for two-year CAR and three-year CAR, respectively). These results are inconsistent with Carter, Dark and Singh (1998) showing that IPOs underwritten by more prestigious underwriters experience better long-run performance.

The Survival of IPO Firms after Issuance

Ibbotson (1975) argues that IPOs are typically young and risky firms. Therefore, IPO investment needs more risk premium leading to the underpricing of IPO firms. In this paper, we show that the issuing firms acquire information from the investors to adjust the IPO pricing terms during the waiting period. The healthy IPO firms may receive positive information (strong demand) and survive long after issuance. However, if the positive information revealed by the investors is subject to fads, the IPOs receiving positive information may not be able to survive after the alleviation of fads. On the other hand, bad IPOs that receive negative information during the waiting period may die soon after issuance.

Figure 4 shows the surviving rate of IPOs with the final proceeds adjusted below, within, and above the expected proceeds range. We find that not many

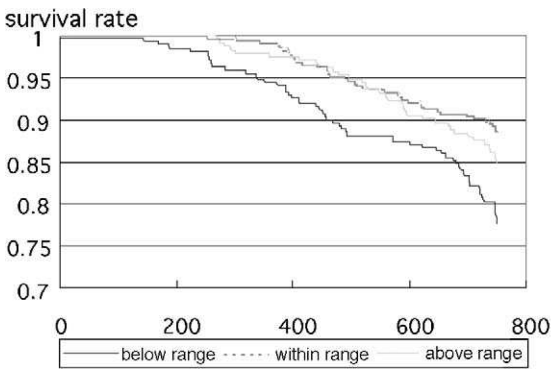


Fig. 4. The Surviving Rate of IPOs with Respect to the Proceeds Adjustment.

The surviving rate of IPOs with final proceeds adjusted below, within, and above the expected range. The surviving rate is defined as the number of IPO firms in the market at a certain time point divided by the number of IPOs in the sample.

IPO firms disappear from the market one year after issuance. Two or three years after issuance, a lot of IPOs fail to survive. Furthermore, the surviving rate of IPOs receiving negative information during the waiting period is the worst. More than 20% of the IPO firms receiving negative information during the waiting period die three years after issuance. The surviving rate of IPOs without significant adjustment on the pricing terms is highest among the sample while the surviving rate of IPOs receiving positive information is in the middle. Therefore, part of IPO proceeds adjustment can be attributed to the market fads. Ordinary IPOs without particular notice from the investors survive longer.

7. CONCLUSION

This paper examines the information acquisition activity by the underwriter of an IPO during the waiting period. As do Benveniste and Spindt (1989), we argue that the information acquired in the waiting period can help the underwriter measure the demand for the new offering and more accurately set the price and number of shares issued. For positive information, the underwriter increases the offer price to lessen the extent of underpricing. For negative information, the underwriter decreases the offer price to avoid the possibility of failure. The issuer has to compensate investors for their information to induce them to reveal their information truthfully.

Our results indicate that the industry change but not the market change does influence decisions on pricing term adjustments. However, the pricing term adjustments cannot be fully attributed to industry change. Information from investors plays an important role in explaining the behaviors of offer price and number of shares offered. The positive information acquired from the investors during the waiting period is costly. However, the issuer still benefits from the information gathering activities. The issuer benefits if the underwriter acquires positive information from investors and receives a larger positive initial return. Nevertheless, the investors cannot force the issuer's benefit to zero by requesting higher compensation because the investors may not always be able to receive an allocation of the offering. This allocation discrimination causes the issuer to have positive benefit from the information acquisition activities.

This paper supports the validity of the hypothesis of underwritten hypothesis of underwriter's information acquisition proposed by Benveniste and Spindt (1989) during the IPO waiting period. The information acquisition helps IPO pricing become more efficient. The issuer benefits from the information acquisition activities because a lesser amount of wealth is transferred to the investors. The investors who provide positive information benefit from compensation in the form of underpricing and a number of shares offered for their revealing information, and may keep a long-term relationship with the underwriter to receive allocations of future IPOs.

Besides, this paper finds a connection between the information revealed during the waiting period and IPO long-run performance. Basically, IPOs receiving positive information about strong demand for the offerings during the waiting period experience poor long-run performance. This finding is consistent with fads or overreaction in the IPO market. Furthermore, we also show that IPOs without receiving significant information during the waiting period survive longer.

ACKNOWLEDGMENTS

We would like to thank Tim Loughran, Forrest Nelson, J. SaAadu, Theo Termaelen, Anand Vijh, the seminar participants at the University of Iowa, National Sun Yat-Sen University (Taiwan), NTU International Conference on Finance (Taiwan), and the 12th PACAP-FMA Cinference for their helpful comments and suggestions. All the remaining errors are our own responsibility.

NOTES

1. There is a minimum 20-day waiting period requirement from the SEC. However, no maximum number of days limit for the waiting period exists. Nevertheless, the longer the waiting period, the less informative the information revealed in the preliminary prospectus.
2. For example, see Ibbotson (1975), Smith (1986), Ibbotson, Sindelar and Ritter (1988), Ibbotson and Ritter (1995) and Chen (1997), among others.
3. Bloomberg does not specify the "firm commitment" offerings and "best efforts" offerings. We confirm that the IPOs in the sample are firm commitment offerings by checking the offerings listed in Corporate Finance: *The IDD Review of Investment Banking*.
4. Our sample covers data over the period from 1993 to 1997. Since we are also measuring the long-run performance and the survival rate of IPOs three years after issuance, IPOs issued after 1994 will not have long enough life for the measure of long-run performance and the survival rate.
5. Since Bloomberg seldom reports CU.S.IP for IPOs in the sample, the benefit of using the closing price from *OTC, ASE or NYSE Daily Stock Price Record* rather than Center for Research in Security Prices (CRSP) data base is to help detect whether an IPO is a unit offering or not.
6. Since few securities in CRSP have the same three-digit SIC code as many of the IPOs, the two-digit SIC code is used to define an industry.

7. Beatty and Ritter (1986) and Megginson and Weiss (1991) employ the market share of an underwriter as the proxy for underwriter prestige. We do not use Carter-Manaster (1990) measure for underwriter prestige because Carter-Manaster sample starts in 1979 and ends in 1983 which does not overlap with our sample. If we used Carter-Manaster measure, we assumed the underwriter prestige does not change over time.

8. Some may argue that our sample suffers from truncation bias: There are occasions where instead of reducing the number of shares or lowing the offer price, the IPO is just withdrawn from the market. In our sample, we still have 31.41% of the IPOs with final offer proceeds adjusted below the expected proceeds range implying that we have large enough sample to examine the behavior of IPOs with final proceeds adjusted.

9. The average survival rate of non-IPOs in 1993 and 1994 is 86.25%. Hence, IPOs are more likely to die than non-IPOs. Furthermore, one third of delisted IPOs are due to bankruptcy while two thirds are merged.

10. Since investors always have incentives to reveal negative information to the issuers and underwriters, therefore negative information needs not be compensated.

11. If the industry change is not included as an independent variable of the regression, the effect of overall market change on the decision change is significant. Consequently, the effect of market change on the decision adjustment is subsumed by the effect of industry change.

12. We would like to thank Forrest Nelson for suggesting this idea.

13. This method is similar to Beatty (1989) dealing with the reputation of auditors.

14. See Beatty and Ritter (1986). They argue that the size of offer is a proxy for the ex ante uncertainty for the IPO.

15. The impact of underwriter prestige on IPO underpricing is two-folded: (1) Prestigious underwriter can always price the offerings more precisely. Therefore, IPOs underwritten by prestigious underwriters should be less underpriced; (2) From the point of view of information acquisition, prestigious underwriters typically acquire more information from the regular investors to adjust the pricing terms. Therefore, compensation for IPOs underwritten by more prestigious underwriters is higher.

REFERENCES

Aggarwal, R., & Rivoli, P. (1990). Fads in the Initial Public Offering Market? *Financial Management*, *19*, 45–57.

Barber, B. M., & Lyon, J. D. (1997). Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics. *Journal of Financial Economics*, *43*, 341–372.

Beatty, R. P. (1989). Auditor Reputation and the Pricing of Initial Public Offerings. *Accounting Review*, *64*, 693–709.

Beatty, R. P., & Ritter, J. R. (1986). Investment Banking, Reputation, and Underpricing of Initial Public Offerings. *Journal of Financial Economics*, *15*, 213–232.

Benveniste, L. M., Busaba, W. Y., & Wilhelm, W. J. (1996). Price Stabilization as a Bonding Mechanism in New Equity Issues. *Journal of Financial Economics*, *42*, 223–255.

Benveniste, L. M., & Spindt, P. A. (1989). How Investment Bankers Determine the Offer Price and Allocation of New Issues. *Journal of Financial Economics*, *24*, 343–361.

Brav, A., (2000). Inference in Long-Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings. *Journal of Finance*, *55*, 1979–2016.

Brav, A., & Gompers, P. A. (1997). Myth or Reality? The Long-Run Underperformance of Initial Public Offerings: Evidence from Venture and Nonventure Capital-Backed Companies. *Journal of Finance*, *52*, 1791–1821.

Brav, A., Geczy, C., & Gompers, P. A. (2000). Is the Abnormal Return Following Equity Issuance Anomalous? *Journal of Financial Economics*, *56*, 209–249.

Camp, G., & Munro, R. (2000). Underpricing of Initial Public Offerings in New Zealand: A Comparison of the Fixed Price and Book-Building Methods. Working paper. University of Melbourne.

Carter, R., Dark, F. H., & Singh, A. K. (1998). Underwriter Reputation, Initial Returns, and the Long-Run Performance of IPO Stocks. *Journal of Finance*, *53*, 285–311.

Carter, R., & Manaster, S. (1990). Initial Public Offerings and Underwriter Reputation. *Journal of Finance*, *45*, 1045–1067.

Chen, A. (1997). A Survey on IPO Related Theories and Evidence. *Journal of Management (Taiwan)*, *14*, 403–436.

Cornelli, F., & Goldreich, D. (1999). Bookbuilding and Strategic Allocation. Working paper. London Business School.

Fama, E., & French, K. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, *33*, 3–56.

Hanley, K. W. (1993). The Underpricing of Initial Public Offerings and the Partial Adjustment Phenomenon. *Journal of Financial Economics*, *34*, 231–250.

Ibbotson, R. G., & Sindelar, J. L., & Ritter, J. R. (1988). Initial Public Offerings. *Journal of Applied Corporate Finance*, *1*, 37–45.

Ibbotson, R. G., & Ritter, J. R. (1995). Initial Public Offerings. In: R. A. Jarrow, V. Maksimovic & W. T. Ziemba (Eds), *Handbooks of Operations Research and Management Science*. North-Holland.

Ibbotson, R. G. (1975). Price Performance of Common Stock New Issues. *Journal of Financial Economics*, *2*, 235–272.

Kothari, S. P., & Warner, J. B. (1997). Measuring Long-Horizon Security Price Performance. *Journal of Financial Economics*, *43*, 301–339.

Ljungqvist, A. P., Jenkinson, T., & Wilhelm, W. J. (2000). Has the Introduction of Bookbuilding Increased the Efficiency of International IPOs? Working paper. Oxford University.

Loughran, T., & Ritter, J. R. (2001). Why Don't Issuers Get Upset about Leaving Money on the Table in IPOs? *Review of Financial Studies*, forthcoming.

Loughran, T., & Ritter, J. R. (2000). Uniformly Least Powerful Tests of Market Efficiency. *Journal of Financial Economics*, *55*, 361–389.

Lowry, M., & Schwert, G. W. (2000). IPO Market Cycles: An Exploratory Investigation. Working paper. University of Rochester and NBER.

Lyon, J. D., Barber, B. M., & Tsai, C. (1999). Improved Methods for Tests of Long-Run Abnormal Stock Returns. *Journal of Finance*, *54*, 165–201.

Megginson, W. L., & Weiss, K. A. (1991). Venture Capitalist Certification in Initial Public Offerings. *Journal of Finance*, *46*, 879–903.

Ritter, J. R. (1987). The Costs of Going Public. *Journal of Financial Economics*, *19*, 269–281.

Ritter, J. R. (1991). The Long-Run Performance of Initial Public Offerings. *Journal of Finance*, *46*, 3–27.

Rock, K. (1986). Why New Issues Are Underpriced. *Journal of Financial Economics*, *15*, 187–212.

Ruud, J. S. (1993). Underwriter Price Support and the IPO Underpricing Puzzle. *Journal of Financial Economics*, *34*, 135–151.

Smith, C. W. (1986). Investment Banking and the Capital Acquisition Process. *Journal of Financial Economics*, *15*, 3–29.

THE TERM STRUCTURE OF RETURN CORRELATIONS: THE U.S. AND PACIFIC-BASIN STOCK MARKETS

Ming-Shiun Pan and Y. Angela Liu

ABSTRACT

This paper examines the term structure of correlations of weekly returns for six national stock markets namely, Australia, Hong Kong, Japan, Malaysia, Singapore, and the U.S. We decompose stock indexes into permanent and temporary components using a canonical correlation analysis and then calculate short- and long-horizon return correlations from these two price components. The empirical results for the sample period of January 1988 to December 1994 reveal that the relationships of return correlations among these stock markets are not stable across return horizons. While correlations, in general, tend to increase with return horizons, there are several cases showing that correlations decline when investment horizons increase.

1. INTRODUCTION

Portfolio theory suggests that investing in less correlated assets will result in greater diversification effect. The recent attention in investing internationally from exposure to foreign equity markets seems to suggest that there is a trend toward international diversification. While the growing interest of investing internationally implies the general belief of that the interdependence of stock

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 233–249.

ISBN: 0-7623-0887-7

price movements be less pronounced across countries than within a country, the gradual liberalization and changing investment environments in most markets in recent years propose that diversification gains lean upon the continuous exploration to the correlation structure of national stock markets.

Considerable amount of work has been devoted to investigating the degree of interdependence between national stock markets.[1] The evidence revealed in these studies suggest that the degree of correlations in international stock markets has increased substantially for the past decade. Several notable factors, such as expanding links between national economies and increasing intra-regional trade, the trend of deregulations (e.g. an easing of rules on foreign investment and capital control), and the explosive growth in cross-border investing, could result in a world in which stock markets become more synchronized and hence more correlated.[2] Such an increase in correlations might reduce the international diversification benefits.

Although the increased correlation in international stock markets seems to suggest that diversification across national borders may be only marginally beneficial, a further investigation on the degree of correlations in international stock markets is needed for several reasons. First, the correlations reported in previous studies mainly are short-horizon correlations, as they are calculated from relatively high frequency data (e.g. daily returns). Clearly, a knowledge of long-horizon return correlations between different national stock markets is important to many international investors since their holding periods of investment portfolios might be very long. Second, the structures of correlations among international stock markets for short horizons could be different from those of long horizons. For instance, Meric and Meric (1989) document that the correlation relationships of international stock returns are relatively stable in the long run, but less so in the short run. Finally, the correlation estimates for long-horizon returns computed from non-overlapping observations might be biased because of the lack of sufficient data points.

This paper, in the spirit of a study by Chou and Ng (1995), uses a canonical correlation analysis to examine the term structure of return correlations for the stock markets in Australia, Hong Kong, Japan, Malaysia, Singapore, and the U.S. Specifically, we employ the canonical correlation analysis, developed by Tsay and Tiao (1990), to decompose stock prices into permanent and temporary components. The entire term structure of correlations is then derived from the permanent and temporary return components. The canonical correlation approach has obvious advantages. First, the method takes account of possible existence of cointegration among stock markets into the decomposition process. Second, it allows us to compute robust estimates of long-horizon return correlations without facing the problem of not having enough non-overlapping data. The results from this study should shed additional light on the case of diversifying in the Pacific Basin stock markets as explored by Bailey and Stulz (1990) and Solnik (1991).

The rest of the paper is organized as follows. Section 2 describes the canonical correlation approach of decomposing stock prices into permanent and temporary components and the construction of the term structure of return correlations. Section 3 describes the data used and presents the empirical results. The conclusion is in the final section.

2. RESEARCH METHODOLOGY

Decomposition of Stock Prices

An in-depth analysis of the term structure of correlations among international stock markets can be conducted using the canonical correlation analysis of Tsay and Tiao (1990). The Tsay and Tiao canonical correlation approach is appealing for not only it permits us to decompose stock prices into permanent and temporary components, but also it takes account of possible presence of cointegration among a set of stock prices into the decomposition procedure. Given the permanent and temporary price components, a term structure of correlations can be constructed by calculating the variance-covariance matrices of the permanent and temporary returns components for different horizons.

Formally, consider the following process of the natural logarithm of stock price, p_t, such that

$$p_t = x_{1,t} + x_{2,t} \tag{1}$$

where $x_{1,t}$ is referred as the permanent price component and it satisfies

$$x_{1,t} = \mu_1 + x_{1,t-1} + e_{1,t} \tag{2}$$

and $x_{2,t}$ is called the temporary component that satisfies

$$x_{2,t} = \mu_2 + \rho x_{2,t-1} + e_{2,t} \tag{3}$$

for $0 < \rho < 1$. The disturbance terms, $e_{1,t}$ and $e_{2,t}$, are assumed to be stationary, independent and distributed $N(0, \sigma_i^2; i = 1, 2)$. As discussed in Fama and French (1988) and Poterba and Summers (1988), Eq. (1) is used to model p_t as the sum of a random walk, $x_{1,t}$, and a stationary component, $x_{2,t}$.[3]

Let the $(n \times 1)$ vector P_t denote an n-dimensional time series of the natural logarithms of stock prices with a sample of T observations. Following Tsay and Tiao's (1990) canonical correlation procedure, we decompose each element of P_t into a permanent and a temporary component by solving the eigenvalues of matrix A:

$$A = \beta_1 \beta_2 \tag{4}$$

where

$$\beta_1 \equiv \left(\sum_{t=2}^{T} P_t P'_t\right)^{-1} \left(\sum_{t=2}^{T} P_t P'_{t-1}\right) \tag{4.1}$$

$$\beta_2 \equiv \left(\sum_{t=2}^{T} P_{t-1} P'_{t-1}\right)^{-1} \left(\sum_{t=2}^{T} P_{t-1} P'_t\right) \tag{4.2}$$

Note that β_1 is simply the coefficient matrix from a matrix regression of P_{t-1} on P_t and β_2 defines the coefficient matrix from a matrix regression of P_t on P_{t-1}. The eigenvalues of the matrix A, denoted as $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and ordered $(\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n)$, turn out to be the squares of the sample canonical correlations.

Let $(k_1, k_2, \ldots, k_n)$ be the associated $(n \times 1)$ eigenvectors under the normalization constraint and $K \equiv [k_1, \ldots, k_n]$. Define a new $(n \times 1)$ vector η_t and calculate its elements as linear combinations of P_t using K as the matrix of weights; i.e. $\eta_t = K'P_t$. The individual elements of η_t are called canonical variates and can be shown uncorrelated with one another. The first canonical variate η_{1t} can be interpreted as the linear combination of P_t that yields the maximum first order serial correlation from the set of all linear combinations of the time series. The variate η_{2t} gives the linear combination of P_t that is uncorrelated with η_{1t} and yet obtains the largest remaining first order serial correlation, and so on.

Given the constructed canonical variates, the price vector P_t then can be expressed as:

$$\begin{aligned} P_t &= (K')^{-1} K' P_t \\ &= (K')^{-1} \eta_t \\ &= W \eta_t \end{aligned} \tag{5}$$

where $W \equiv (K')^{-1}$. That is, each individual element of P_t can be written as a linear combinations of the canonical variates as

$$p_{it} = \sum_{j=1}^{n} w_{ij} \eta_{jt}, \text{ for } i = 1, \ldots, n, \tag{6}$$

where w_{ij} is the (i, j)*th* element of W. Assuming the first k canonical variates are random walks and the remaining n-k canonical variates are stationary, p_{it} can be further decomposed into the sum of a permanent and a temporary component as follows:

$$\begin{aligned} p_{it} &= \sum_{j=1}^{n} w_{ij}\eta_{jt} \\ &= \sum_{j=1}^{K} w_{ij}\eta_{jt} + \sum_{j=k+1}^{n} w_{ij}\eta_{jt} \\ &= x_{1it} + x_{2it}, \text{ for } i = 1, \ldots, n \end{aligned} \tag{7}$$

where

$$x_{1it} \equiv \sum_{j=1}^{K} w_{ij}\eta_{jt}$$

and

$$x_{2it} \equiv \sum_{j=k+1}^{n} w_{ij}\eta_{jt}.$$

The variable x_{1it} is referred as the permanent component of p_{it} and it is a random walk since $\eta_{1t}, \ldots, \eta_{kt}$ are random walks. The variable x_{2it} is the stationary, temporary component of p_{it} as $\eta_{k+1,t}, \ldots, \eta_{nt}$ are stationary. Thus, the Tsay and Tiao (1990) canonical correlation analysis enables us to decompose each of the multivariate time series into the sum of permanent and temporary components as Eq. (1) shows.

From Eq. (7), the one period return of i*th* stock price series, Δp_{it}, can be expressed as the sum of the permanent return component and the temporary return component. That is,

$$\Delta p_{it} = \Delta x_{1it} + \Delta x_{2it}, \tag{8}$$

where Δ denotes the difference operator. Consequently, we can examine the relationships among the different return components of different stock market indexes.

Calculate the Term Structure of Stock Return Correlations

As demonstrated in Chou and Ng (1995), the decomposition analysis presented above allows us to derive a term structure of return correlations for a set of stock markets. Particularly, the derived term structure of correlations permits us to differentiate short-horizon return correlations that are driven by the temporary components from long-horizon return correlations, which are affected mainly by the permanent components.

According to Chou and Ng (1995), the unconditional variance of the q-period return of stock index i can be calculated as:

$$\mathrm{Var}(\Delta^q p_{it}) = \sum_{h=1}^{k} w_{ih}^2 \, \mathrm{Var}(\Delta\eta_{h,t})q + \sum_{h=k+1}^{n} w_{ih}^2 \, \mathrm{Var}(\Delta^q \eta_{h,t}) \tag{9}$$

where Δ^q denotes the q*th* difference operator and Var(.) is the unconditional variance of (.). Under the assumption that the first k canonical variates are random walks, the first item in the right hand side of Eq. (9) is a result that the variance of the increments in a random walk is linear in the sampling interval. Further, the unconditional covariance between the q-period returns of stock indexes i and j can be calculated as:

$$\mathrm{Cov}(\Delta^q p_{it}, \Delta^q p_{jt}) = \sum_{h=1}^{k} w_{ih} w_{jh} \, \mathrm{Var}(\Delta\eta_{h,t})q + \sum_{h=k+1}^{n} w_{ih} w_{jh} \, \mathrm{Var}(\Delta^q \eta_{h,t}) \tag{10}$$

From Eqs (9) and (10), one can easily derive the entire term structure of correlations of different stock index returns by changing q. Chou and Ng (1995) also show that the limiting correlation between two stock index returns is the correlation between the two corresponding permanent return components as the return horizon, q, increases. Furthermore, in the extreme case when $k = n$ so that there are no temporary components, the correlation will be invariant to the investment horizon q. That is, when they are no temporary components, short-horizon correlations are unbiased estimates of long-horizon correlations.

3. DATA AND EMPIRICAL RESULTS

Data

Data employed in this study are weekly stock market indexes from six countries: Australia, Hong Kong, Japan, Malaysia, Singapore, and the U.S.[4] The data, retrieved from a data bank maintained by the Ministry of Education in Taiwan, are the Goldman Sachs/FT-Actuaries indexes as reported in the Financial Times. The sample data consist of 365 weekly stock indexes covering the period from January 6, 1988 to December 28, 1994.

Table 1 reports summary statistics of weekly returns for the six stock market indexes both in local currency and in U.S. dollar. The Hong Kong and

Malaysian markets have the best performances during the sample period, though the two markets are also the most volatile as the standard deviations show. The stock indexes of Australia, Japan, and Singapore exhibit comparable volatility, while the U.S. index has the lowest volatility. Because of the added fluctuation in exchange rates, the standard deviations of the Australia and Japanese indexes are slightly increased when expressed in U.S. dollar.

Table 1 also shows the correlations for the return series between the six stock markets. The correlation coefficients range from 0.774, for Malaysia and Singapore, to 0.215 for Hong Kong and Japan. All the Asian stock markets have

Table 1. Descriptive Statistics of Weekly Returns for the U.S. and Five Asian Stock Market Indexes.

	Australia	Hong Kong	Japan	Malaysia	Singapore	U.S.
Panel A: In Local Currency						
Mean (%)	0.120	0.342	–0.042	0.389	0.254	0.160
Std. Dev.	0.022	0.035	0.027	0.032	0.029	0.018
Maximum (%)	7.404	7.693	13.473	14.167	10.638	5.524
Minimum (%)	–6.703	–22.557	–11.292	–19.460	–12.486	–7.258
Correlation Coefficients						
Australia	1.000					
Hong Kong	0.388	1.000				
Japan	0.261	0.215	1.000			
Malaysia	0.367	0.490	0.279	1.000		
Singapore	0.404	0.502	0.377	0.774	1.000	
U.S.	0.414	0.253	0.300	0.332	0.391	1.000
Panel B: In US$						
Mean (%)	0.147	0.343	0.028	0.386	0.345	0.160
Std. Dev.	0.026	0.035	0.032	0.033	0.030	0.018
Maximum (%)	10.320	7.978	14.590	13.964	10.900	5.524
Minimum (%)	–7.767	–22.915	–9.978	–23.358	–12.858	–7.258
Correlation Coefficients						
Australia	1.000					
Hong Kong	0.356	1.000				
Japan	0.184	0.173	1.000			
Malaysia	0.290	0.496	0.240	1.000		
Singapore	0.297	0.492	0.319	0.769	1.000	
U.S.	0.325	0.253	0.212	0.316	0.385	1.000

a correlation above 0.3 with the U.S., except Hong Kong. In addition, the correlations are, in general, higher when local currency returns are used. The correlations for Hong Kong and the U.S., however, are the same for local currency and U.S. dollar returns. This result may reflect the fact that Hong Kong established a fixed official rate of HK$7.80 per Greenback on October 17, 1983.

Preliminary Analysis of Data

We first employ unit root tests to determine whether each stock index series has a unit root (i.e. a permanent component). Two unit root tests, the Augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests, are used to test the null hypothesis of a unit root in each stock index series. Table 2 reports the ADF and PP test statistics. The ADF test statistics range from –3.208 (for Australia in U.S. dollar) to –0.897 (for Japan in U.S. dollar) and all of the tests, except for Australia in U.S. dollar, fail to reject the null hypothesis of a unit root in the stock index series at the 5% significance level. Moreover, based on the PP test statistics, none of the six stock index series rejects the null hypothesis of a unit root at the 5% significance level. Therefore, the results strongly suggest that each stock index series contains a permanent component.

Table 2. Tests for Unit Roots in the Logarithms of the Six Stock Market Indexes.

Statistic	Australia	Hong Kong	Japan	Malaysia	Singapore	U.S.
Panel A: In Local Currency						
ADF	–2.359	–0.900	–1.264	–1.125	–1.290	–1.625
PP	–1.956	–0.814	–1.119	–1.090	–1.140	–1.446
Panel B: In US$						
ADF	–3.208	–0.897	–1.896	–0.921	–0.911	–1.625
PP	–2.670	–0.806	–1.709	–0.885	–0.810	–1.446

Notes: The ADF and PP statistics are, respectively, the Augmented Dickey-Fuller and the Phillips-Perron adjusted t-statistics. The truncation lag used is four for the ADF test and is three for the PP test. Values less than the critical value reject the unit root null hypothesis. The critical values for both test statistics are –2.86 (5%) and –3.43 (1%).

Decomposition Results

Given that each of the six stock market indexes contains a permanent component, we utilize the canonical correlation procedure described in pervious section to decompose each of the six stock indexes into a permanent and a temporary component. Table 3 gives the ordered eigenvalues for the matrix A, in local currency as well as in U.S. dollar. For both cases, the first four eigenvalues are very close to 1, while the fifth eigenvalue for the U.S.-dollar case seems to approach 1. To determine statistically how many of the canonical variates follow random walks, we apply the ADF and the PP unit root tests on the canonical variates, and the test results are also given in Table 3. As can be seen, both the ADF and the PP test statistics clearly cannot reject the unit root null hypothesis at the 5% significance level for the first four canonical variates for both cases. The fifth canonical variate in U.S. dollar seems to be stationary based on the ADF statistic, though the PP statistic suggests that it follows a random walk. Nevertheless, the empirical results reported in what follows would be based on that the first four canonical variates are random walks and the last two are stationary.[5]

Table 3. Tests for Unit Roots in the Canonical Variates.

	Component					
Statistic	1	2	3	4	5	6
Panel A: In Local Currency						
Eigenvalue	0.999	0.997	0.976	0.948	0.899	0.823
ADF	–1.936	–0.688	–1.387	–2.014	–3.798	–3.884
PP	–1.732	–0.443	–1.581	–2.296	–3.312	–4.185
Panel B: In US$						
Eigenvalue	0.999	0.996	0.975	0.959	0.930	0.852
ADF	–1.695	–0.900	–1.323	–1.997	–3.019	–4.362
PP	–1.526	–0.683	–1.393	–2.065	–2.711	–3.883

Notes: Eigenvalues are the eigenvalues of matrix A in Eq. (4) in descending order. The ADF and PP statistics are, respectively, the Augmented Dickey-Fuller and the Phillips-Perron adjusted *t*-statistics. The truncation lag used is four for the ADF test and is three for the PP test. Values less than the critical value reject the unit root null hypothesis. The critical values for both test statistics are –2.86 (5%) and –3.43 (1%).

The presence of four common unit roots in the six stock price series suggests that there are common forces driving the long-run movements among them. In other words, the result implies that the six stock index series are cointegrated.[6]

Once the number of common unit roots, k, is determined, each stock index series can be decomposed into a permanent component and a temporary component series based on Eq. (7). Further, we can calculate the return for each component. Table 4 contains summary statistics of the permanent and temporary return components for the six stock indexes. It appears that most of the stock index returns are primarily contributed by the permanent return components since most of the sample means of the temporary return components are close to zero. Also, the standard deviations of the permanent return components are generally greater than those of the temporary return components. As expected, the correlations between the permanent and temporary

Table 4. Descriptive Statistics of the Returns of the Permanent and Temporary Components.

	Permanent component		Temporary component			
Country	Mean (%)	Std. Dev.	Mean (%)	Std. Dev.	ρ	R^2
Panel A: In Local Currency						
Australia	0.113	0.018	0.007	0.012	0.021	0.710
Hong Kong	0.344	0.032	–0.002	0.014	–0.008	0.838
Japan	–0.036	0.024	–0.006	0.014	–0.025	0.753
Malaysia	0.387	0.032	0.002	0.005	–0.001	0.979
Singapore	0.245	0.021	0.010	0.020	0.001	0.512
U.S.	0.160	0.018	0.000	0.001	–0.022	0.994
Panel B: In US$						
Australia	0.116	0.017	0.031	0.019	0.013	0.458
Hong Kong	0.316	0.032	0.028	0.014	–0.003	0.843
Japan	0.036	0.028	–0.008	0.015	–0.003	0.779
Malaysia	0.392	0.033	–0.006	0.004	0.011	0.987
Singapore	0.330	0.026	0.015	0.015	–0.018	0.758
U.S.	0.168	0.017	–0.009	0.004	–0.009	0.940

Notes: ρ is the correlation between the permanent return component and the temporary return component. R^2 is the R-squared value from regressing stock returns on the permanent return components. It measures the percentage of variation in stock returns explained by variation in the permanent return components.

returns components are quite small, as indicated by the ρ values. Moreover, the rather high R^2's indicate that a large proportion of variations in returns for most of the markets can be explained by the permanent component, except for Singapore in local currency case and Australia in U.S. dollar case. For these two exceptions, the temporary component seems to be as important as the permanent component in explaining the variation in stock index returns.

To further investigate the stochastic process of the temporary components, we fit an ARMA model to each temporary price series. The results, contained in Table 5, show that an AR(1) model is reasonable in describing the temporary price series. None of the Ljung-Box Q statistics is significant, suggesting support for the AR(1) model. The estimated AR(1) coefficient is useful in determining the speed of mean reversion in the temporary components. In specific, the half life of a shock to the temporary component, defined as m, can be computed as $\alpha^m = 0.5$, where α is the estimated AR(1) coefficient.

The speeds of mean reversion in the temporary price components, as measured by m, are shown in Table 5. The U.S. stock market appears to have the slowest mean reversion with the AR(1) coefficients of 0.941 and 0.967,

Table 5. The Results of Estimating the Temporary Components as an AR(1) Model.

Statistic	Australia	Hong Kong	Japan	Malaysia	Singapore	U.S.
Panel A: In Local Currency						
Coefficient	0.939	0.909	0.929	0.927	0.933	0.941
Q(4)	5.887	2.273	6.007	3.692	5.073	5.661
Q(8)	7.777	11.596	9.194	7.719	7.818	7.651
m	11.013	7.265	9.412	9.144	9.995	11.398
Panel B: In US$						
Coefficient	0.961	0.966	0.934	0.953	0.940	0.967
Q(4)	7.709	4.261	6.482	4.119	4.467	5.380
Q(8)	9.916	7.950	9.491	7.381	9.006	9.506
m	17.424	20.038	10.152	14.398	11.202	20.656

Notes: Coefficient is the estimated AR(1) coefficient. Q(k) statistic is Ljung-Box Q statistic, which is distributed as χ^2 distribution with k degree of freedoms, and is to test the hypothesis that all autocorrelations up to lag k are jointly zero. The critical values for Q(4) and Q(8) are 7.779 (10%) and 9.488 (5%), and 13.362 (10%) and 15.507 (5%), respectively. m is the half life of a shock to the temporary component and is computed as $\alpha^m = 0.5$, where α is the estimated AR(1) coefficient for the temporary component series.

respectively, for the local currency and U.S. dollar cases. These AR(1) coefficients represent a shock half life of 11.398 weeks and 20.656 weeks respectively. Hong Kong, with a shock half life of 7.265 weeks, has the fastest mean reversion for the local currency case, while it is Japan that has the highest speed of mean reversion (i.e. 10.152 weeks) for the U.S. dollar case.

Term Structure of Correlations of Stock Returns

As Eqs (9) and (10) show, the variance and covariance of the stock returns for different horizons are a function of the variances of the last n-k stationary canonical variates and the return horizon q. Technically, the computations of the variance and covariance of the stationary canonical variates can be done more efficiently by utilizing the properties of their stochastic processes. To this end, we fit an ARMA model to each of the two stationary canonical variates. The results of the estimations, reported in Table 6, indicate that an AR(1) model is appropriate, based on the Ljung-Box statistics.

Since both the last two canonical variates follow an AR(1) process, the variance of the qth difference of the series can be calculated as:

$$\mathrm{Var}(\Delta^q \eta_{h,t}) = [(1 - \alpha_h^q)/(1 - \alpha_h)]\mathrm{Var}(\Delta\eta_{h,t}),\ h = 5, 6. \tag{11}$$

Table 6. The Results of Estimating the Last Two Canonical Variates as an AR(1) Model.

Statistic	η_5	η_6
Panel A: In Local Currency		
Coefficient	0.950	0.908
Q(4)	5.806	1.230
Q(8)	10.450	11.009
Panel B: In US$		
Coefficient	0.967	0.925
Q(4)	5.229	5.196
Q(8)	9.367	10.453

Notes: Coefficient is the estimated AR(1) coefficient. Q(k) statistic is Ljung-Box Q statistic, which is distributed as χ^2 distribution with k degree of freedoms, and is to test the hypothesis that all autocorrelations up to lag k are jointly zero. The critical values for Q(4) and Q(8) are 7.779 (10%) and 9.488 (5%), and 13.362 (10%) and 15.507 (5%), respectively.

where α_h is the estimated AR(1) coefficient. Using Eqs (9), (10), and (11), we can derive a term structure of return correlations by changing return horizon q.

We derive a term structure of return correlations for the six stock markets for $q = 1$, 4, 13, 52, 104, and ∞.[7] These q values are roughly corresponding to investment horizons for one week, one month, one quarter, one year, two years, and the infinite holding-period. The compiled term structure of correlations results are contained in Table 7. For the local currency case, the return correlations between different markets generally become larger when investment horizons increase, implying a deduction in diversification benefits from cross-border investing when investment horizons increase. However, a notable exception is the term structure of correlations between Hong Kong, the second best performing market over the period, and the other markets. The term structure of correlations between Hong Kong and Japan is a downwarding-slope curve, while it is relatively flat for those between Hong Kong and the markets in Malaysia, Singapore, and the U.S. The results suggest that the diversification gains could be substantial for Japanese and Hong Kong investors by investing in the perspective markets for a long investment horizon. It is also noteworthy that the correction coefficient between Malaysia and Singapore is as high as 0.95 for a long horizon, which could be due to that the two countries are geographically close to each other as well as economically interdependent.

For the U.S. dollar case, most of the term structures of correlations have positive slopes, i.e. longer horizons, higher correlations. The exceptions are for those of Australia/Hong Kong, Hong Kong/Singapore, and Japan/Singapore. In addition, several pairs of countries, including Australia/Hong Kong, Hong Kong/Japan, Japan/Singapore, and Japan/the U.S., exhibit a correlation as low as about 0.20 even for a very long investment horizon (i.e. the infinite holding period). Thus, our results suggest that significant diversification benefits are available for investors who invest in those markets that are less correlated with the domestic markets. Furthermore, such diversification gains could be more pronounced when investment horizons are relatively long.

4. CONCLUSIONS

This study uses a canonical correlation analysis to decompose weekly stock indexes into permanent (random walk) and temporary (stationary) components for six stock markets (i.e. Australia, Hong Kong, Japan, Malaysia, Singapore, and the U.S.) during the period of 1988 to 1994. The term structure of return correlations is then derived from the decomposed permanent and temporary price components for different investment horizons.

Table 7. Term Structure of Return Correlations for the Six Stock Markets.

q		Australia	Hong Kong	Japan	Malaysia	Singapore
Panel A: In Local Currency						
1	Australia	1.000				
	Hong Kong	0.401	1.000			
	Japan	0.259	0.226	1.000		
	Malaysia	0.358	0.505	0.270	1.000	
	Singapore	0.399	0.518	0.371	0.770	1.000
	U.S.	0.405	0.261	0.301	0.323	0.386
4	Australia	1.000				
	Hong Kong	0.430	1.000			
	Japan	0.292	0.210	1.000		
	Malaysia	0.366	0.504	0.270	1.000	
	Singapore	0.424	0.511	0.369	0.781	1.000
	U.S.	0.414	0.261	0.302	0.323	0.396
13	Australia	1.000				
	Hong Kong	0.497	1.000			
	Japan	0.374	0.173	1.000		
	Malaysia	0.384	0.504	0.270	1.000	
	Singapore	0.482	0.498	0.370	0.809	1.000
	U.S.	0.437	0.262	0.303	0.324	0.422
52	Australia	1.000				
	Hong Kong	0.622	1.000			
	Japan	0.547	0.111	1.000		
	Malaysia	0.415	0.506	0.276	1.000	
	Singapore	0.574	0.491	0.411	0.876	1.000
	U.S.	0.486	0.263	0.305	0.325	0.486
104	Australia	1.000				
	Hong Kong	0.667	1.000			
	Japan	0.619	0.091	1.000		
	Malaysia	0.424	0.507	0.281	1.000	
	Singapore	0.600	0.498	0.445	0.908	1.000
	U.S.	0.507	0.264	0.306	0.326	0.516
∞	Australia	1.000				
	Hong Kong	0.716	1.000			
	Japan	0.707	0.063	1.000		
	Malaysia	0.443	0.498	0.304	1.000	
	Singapore	0.630	0.498	0.507	0.951	1.000
	U.S.	0.535	0.262	0.316	0.338	0.561

Table 7. Continued.

q		Australia	Hong Kong	Japan	Malaysia	Singapore
Panel B: In US$						
1	Australia	1.000				
	Hong Kong	0.359	1.000			
	Japan	0.188	0.182	1.000		
	Malaysia	0.269	0.497	0.236	1.000	
	Singapore	0.280	0.493	0.318	0.765	1.000
	U.S.	0.317	0.253	0.218	0.305	0.381
4	Australia	1.000				
	Hong Kong	0.359	1.000			
	Japan	0.217	0.183	1.000		
	Malaysia	0.273	0.502	0.242	1.000	
	Singapore	0.300	0.491	0.306	0.779	1.000
	U.S.	0.331	0.259	0.219	0.304	0.389
13	Australia	1.000				
	Hong Kong	0.356	1.000			
	Japan	0.292	0.186	1.000		
	Malaysia	0.285	0.514	0.256	1.000	
	Singapore	0.349	0.488	0.277	0.814	1.000
	U.S.	0.370	0.274	0.219	0.304	0.407
52	Australia	1.000				
	Hong Kong	0.330	1.000			
	Japan	0.483	0.206	1.000		
	Malaysia	0.329	0.544	0.281	1.000	
	Singapore	0.445	0.478	0.228	0.884	1.000
	U.S.	0.491	0.318	0.215	0.301	0.452
104	Australia	1.000				
	Hong Kong	0.302	1.000			
	Japan	0.592	0.222	1.000		
	Malaysia	0.364	0.561	0.289	1.000	
	Singapore	0.481	0.470	0.216	0.912	1.000
	U.S.	0.576	0.345	0.211	0.300	0.476
∞	Australia	1.000				
	Hong Kong	0.244	1.000			
	Japan	0.781	0.242	1.000		
	Malaysia	0.448	0.584	0.310	1.000	
	Singapore	0.547	0.456	0.222	0.949	1.000
	U.S.	0.728	0.385	0.205	0.310	0.516

The findings indicate that the structure of correlations is not stable over investment horizons with long-horizon return correlations in general larger than those of short-horizons. However, several exceptions that have an inverted curve are also observed. These exceptions represent the cases that the correlation coefficients decline when investment horizons increase, implying that substantial diversification benefits can be gained by having a relatively long investment horizon. In short, our results suggest that investing in Pacific Basin stock markets for diversification benefits should not only consider short-run return correlations but also long-run return correlations.

NOTES

1. See, for example, Panton, Lessig, and Joy (1976), Finnerty and Schneeweis (1979), Hillard (1979), Schollhammer and Sand (1985), Von Furstenberg and Jeon (1989), and Koch and Koch (1991), among many others.
2. For example, Gultekin, Gultekin, and Penati (1989) implicate governments' policies, such as different styles of political organization and trade barriers, as the source of segmentation in international capital markets.
3. To simplify the presentations, we assume the temporary component follows a stationary first-order autoregressive (AR(1)) process. The results derived in what follows also hold if the temporary component follows a higher order ARMA process.
4. All stock indexes are closing indexes quoted on Wednesday. If Wednesday's price index is missing, then Thursday's closing index is used. There are only seven cases in which the Thursday's close is used.
5. We also conduct the study by treating the first five canonical variates in U.S. dollar as random walks. The results are available upon request.
6. This cointegration result appears to be in contrast with that of Chan, Gup, and Pan (1992), in which they use Engle and Granger's (1987) two-step cointegration analysis and find no evidence of cointegration among the U.S. and five Asian stock markets. The contradiction might be due to different testing methods used or to the increase of international stock markets integration after the October 1987 stock market crash as documented in Arshanapalli and Doukas (1993).
7. Notice that the correlation between two stock index returns when the return horizon approaches infinity is simply the correlation between the two permanent stock return components.

ACKNOWLEDGMENTS

We are grateful to Kate Phylaktis and participants at the 1996 APFA/PACAP Finance Conference in Taipei for helpful comments.

REFERENCES

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.

Bailey, W., & Stulz, R. M. (1990). Benefits of International Diversification: The Case of Pacific Basin Stock Markets. *Journal of Portfolio Management*, *17*, 57–61.

Chan, K. C., Gup, B. E., & Pan, M.-S. (1992). An Empirical Analysis of Stock Prices in Major Asian Markets and the United States. *Financial Review*, *27*(2), 289–307.

Chou, R. Y., & Ng, V. K. (1995). Correlation Structure of the Permanent and Temporary Components of International Stock Market Prices. Unpublished working paper. Georgia Institute of Technology.

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Corrections: Presentation, Estimation, and Testing. *Econometrica*, *55*, 251–276.

Fama, E., & French, K. (1988). Permanent and Temporary Components of Stock Prices. *Journal of Political Economy*, *96*(2), 246–273.

Finnerty, J., & Schneeweis, T. (1979). The Co-movement of International Asset Returns. *Journal of International Business Studies*, *10*, 66–78.

Gultekin, M. N., Gultekin, N. B., & Penati, A. (1989). Capital Controls and International Capital Market Segmentation: The Evidence from the Japanese and American Stock Markets. *Journal of Finance*, *44*, 849–869.

Hillard, J. (1979). The Relationship Between Equity Indices on World Exchange. *Journal of Finance*, *34*, 103–114.

Koch, P. D., & Koch, T. W. (1991). Evolution in Dynamic Linkages Across Daily National Stock Indexes. *Journal of International Money and Finance*, *10*, 231–251.

Meric, I., & Meric, G. (1989). Potential Gains from International Portfolio Diversification and Inter-Temporal Stability and Seasonality in International Stock Market Relationships. *Journal of Banking and Finance*, *13*, 627–640.

Panton, D., Lessig, V., & Joy, O. (1976). Comovements of International Equity Markets: A Taxonomic Approach. *Journal of Financial and Quantitative Analysis*, *11*, 415–432.

Poterba, J., & Summers, L. (1988). Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics*, *22*(1), 27–59.

Schollhammer, H., & Sand, O. (1985). The Interdependence Among the Stock Markets of Major European Countries and the United States: An Empirical Investigation of Interrelationships Among National Stock Price Movements. *Management International Review*, *25*, 17–26.

Solnik, B. (1991). Pacific Basin Stock Markets and International Diversification. In: S. G. Rhee & R. P. Chang (Ed.), *Pacific-Basin Capital Market Research*, *II*, 309–321.

Tsay, R., & Tiao, G. (1990). Asymptotic Properties of Multivariate Nonstationary Processes with Applications to Autoregressions. *Annals of Statistics*, *18*, 220–250.

Von Furstenberg, G. M., & Jeon, B. N. (1989). International Stock Price Movements: Links and Messages. *Brookings Papers on Economic Activity*, *1*, 125–179.

This Page Intentionally Left Blank

CHARACTERISTICS VERSUS COVARIANCES: AN EXAMINATION OF DOMESTIC ASSET ALLOCATION STRATEGIES

Jonathan Fletcher

ABSTRACT

This paper examines the out of sample performance of monthly asset allocation strategies within U.K. industry portfolios using linear asset pricing models and a characteristic-based model of stock returns to forecast expected returns. We find that strategies that use conditional versions of the asset pricing models outperforms the strategy that uses the characteristics-based model in terms of higher Sharpe performance and more positive abnormal returns. In addition, these strategies provide significant positive Jensen (1968) and Ferson and Schadt (1996) performance measures even with binding investment constraints. Our results support the usefulness of conditional asset pricing models in mean-variance analysis.

1. INTRODUCTION

There has been a recent debate in the empirical asset pricing literature as to whether stock characteristics or covariances drive cross-sectional stock returns. This has centred on the interpretation of the size (SMB) and book-to-market

Advances in Investment Analysis and Portfolio Management, Volume 9, pages 251–271.

ISBN: 0-7623-0887-7

(HML) factors in the Fama and French (1993) factor model. Proponents in this debate include studies by Fama and French (1993, 1995, 1996), Davis, Fama and French (2000), Daniel and Titman (1997), and Daniel, Titman and Wei (2001).

The issue of whether characteristics or covariances drive expected returns has a number of important implications. Daniel and Titman (1998) argue that a characteristics model of stock returns has major impact in the estimation of expected returns and in the performance evaluation of managed funds. Daniel, Grinblatt, Titman and Wermers (1997) develop a measure of fund performance where the performance of the fund is evaluated relative to a portfolio which has similar characteristics in terms of size, book-to-market ratio and momentum[1] of the stocks held by the mutual fund. Daniel et al find that this approach to fund performance yields a more favorable picture of U.S. mutual fund performance compared to measures based on linear factor models. Another implication is in the area of optimal portfolio choice for a mean-variance investor. If expected returns are explained by an asset pricing model, then the optimal portfolio to hold is some combination of the factor portfolios (Grinblatt & Titman, 1987). However this does not hold under a characteristics model of stock returns.

This paper examines the importance of the characteristics versus covariances debate from a different perspective. We examine the out of sample performance of domestic U.K. industry asset allocation strategies where the investor uses either a characteristics model or linear asset pricing model to forecast expected returns[2]. We explore whether a strategy that uses a characteristics model to forecast expected returns outperforms strategies that use linear asset pricing models. We include a number of linear asset pricing models in our analysis, such as the capital asset pricing model (CAPM), arbitrage pricing theory (APT), and empirical factor models similar to Fama and French (1993) and Elton, Gruber and Blake (1996). We evaluate the out of sample performance of the strategies using measures based on Jensen (1968) and Ferson and Schadt (1996).

Our study complements a number of recent studies that examine the practical implications of the characteristics versus covariances debate for investors (see Haugen & Baker, 1996; Pastor & Stambaugh, 2000; Chou, Li & Zhou, 2001; among others). The second contribution of the paper is that we provide evidence as to whether using conditional versions of the asset pricing models has any impact on the out of sample performance of strategies that use linear asset pricing models. Empirical research suggests that conditional versions of asset pricing models are more able to explain the cross-section of stock returns (see Hodrick & Zhang, 2001; among others). However the use of conditional asset pricing models is not without its' critics. Ghysels (1998) argues that if the assumed dynamics of the conditional model is mis-specified, then they may have poorer out of sample performance compared to unconditional versions. Our study provides additional evidence on this issue.

We report three main findings in the paper. First, the strategy that uses the characteristic model exhibits better performance than the strategies that use linear asset pricing models whenever unconditional versions of the models are used. Second, we find the performance of the strategies that use linear asset pricing models improves sharply whenever conditional versions of the models are used. Third, strategies that use conditional asset pricing models outperforms the strategy that uses the characteristics model. Our results support the usefulness of conditional asset pricing models in mean-variance analysis.

The paper is outlined as follows. Section 2 describes the method used in the paper. The data and factor models are discussed in Section 3. The empirical results are reported in Section 4. The final section contains concluding comments.

2. METHOD

Models of Expected Return

Linear Asset Pricing Models

The first group of models we use are linear asset pricing models. We estimate the expected excess returns on the assets using unconditional and conditional versions of the asset pricing models. Using the unconditional versions of the models, we estimate the expected excess return by:

$$E(r_i) = \sum_{k=1}^{K} \beta_{ik} E(r_k) \tag{1}$$

where $E(r_i)$ is the expected excess return on asset i, β_{ik} is the beta of asset i relative to factor k, $E(r_k)$ is the expected excess return on the mimicking portfolio that captures factor k and K is the number of factors in the model.

To estimate expected excess returns from Eq. (1), we require estimates of the asset betas and factor risk premiums. We estimate the asset betas by the following multiple regression over the prior 60 months:

$$r_{it} = \alpha_i + \sum_{k=1}^{K} b_{ik} r_{kt} + u_{it} \tag{2}$$

where r_{it} and r_{kt} is the excess return in period t on asset i and factor k respectively and the b_{ik}'s are the estimated factor betas. We use the sample mean excess returns on the factor over the prior 60 months as the estimate of the factor risk premium. The use of a 60 month estimation window is similar to Chou, Li and Zhou (2001) among others.

We also estimate expected excess returns using the conditional versions of the models. Using the conditional versions of the models, we estimate expected excess returns by:

$$E(r_{it} | Z_{t-1}) = \sum_{k=1}^{K} \beta_{ik}(Z_{t-1})\lambda_k(Z_{t-1}) \tag{3}$$

where $\beta_{ik}(Z_{t-1})$ is the conditional beta of asset i on the kth factor, $\lambda_k(Z_{t-1})$ is the conditional risk premium of factor k, and Z_{t-1} is the assumed information set of investors. We assume that the betas are constant and allow the factor risk premiums to change over time. We assume the betas are constant because Ferson and Harvey (1991, 1993) show that the predictability in stock returns that can be explained by an asset pricing model is due to changing factor risk premiums and not changing betas (see also Ferson & Locke, 1998). We use the betas from Eq. (2).

We estimate the factor risk premiums by following the approach of Ferson and Harvey (1993), Ferson and Korajcyzk (1995) and Ferson and Locke (1998). We assume that the factor risk premium is a linear function of a common set of information variables used by investors. The information variables are represented by a set of L information variables Z_1 to Z_L. We estimate the expected factor risk premiums at a given point in time in two steps. First, we estimate the coefficients in the factor risk premium function by the following regression using data from the prior 60 months:

$$r_{kt} = \gamma_{k0} + \sum_{l=1}^{L} \gamma_{kl} Z_{lt-1} + u_{it} \tag{4}$$

where Z_{lt-1} is the value of the lth information variable at time $t-1$ for $l = 1, \ldots, L$ and u_{it} is a random error term. The coefficients γ_{k0} and γ_{kl}'s are the estimated coefficients in the factor risk premium. Second, we estimate the conditional factor risk premium at time t by multiplying the γ coefficients by the current values of the information variables. This approach of estimating the factor risk premium takes account of the information that is currently available. We use the estimates of betas and factor risk premiums as inputs to Eq. (3), to get the expected excess returns.

Characteristics Model

The second category of models we use is a characteristics model of stock returns. We follow Haugen and Baker (1996) and assume that the expected excess returns are a linear function of the characteristics. We estimate the expected excess returns from the characteristics model of stock returns by:

$$E(r_i) = \gamma_0 + \sum_{k=1}^{K} \gamma_k c_{ik} \tag{5}$$

where c_{ik}'s are the value of the characteristics of asset i, γ_k are the payoffs of the K stock characteristics, γ_o is the payoff that is unrelated to the stock characteristics and K is the number of characteristics.

We use a three step approach to estimate expected excess returns. First, we estimate cross-sectional regressions of the excess asset returns on a constant and asset characteristics each month over the prior 60 months. Second, we use the average values of the coefficients on the different characteristics as the expected payoffs of the characteristics (Fama & MacBeth, 1973; Haugen & Baker, 1996). Third, we calculate the expected excess return of asset i at time t by the estimated intercept plus the sum of the expected payoffs of the characteristics multiplied by the current values of the characteristics which are known at the start of period t.

Mean-Variance Analysis

We examine the out of sample performance of domestic asset allocation strategies where the investor allocates their wealth across N risky assets and a domestic risk-free asset. Using the notation of Best and Grauer (1990), the goal of the investor is to choose the portfolio that:

$$\text{Max } (t\mathbf{u}'\mathbf{x} - \mathbf{x}'\Sigma\mathbf{x}) \tag{6}$$

Subject to

$$x_i \geq 0 \text{ for } i = 1, \ldots, N$$

$$x_i \leq 0.2 \text{ for } i = 1, \ldots, N$$

where **u** is a N vector of expected excess returns of the N risky assets, Σ is the (N*N) covariance matrix of excess returns of the N risky assets and **x** is a N vector of portfolio weights in the risky assets. The framework in (6) assumes that investors are allowed unrestricted lending and borrowing in the risk-free asset. The coefficient t captures the "mean-variance" risk tolerance of the investor. Given a value of t, the optimal efficient portfolio can be identified.

With unrestricted lending and borrowing, the efficient frontier is a straight line (see Elton & Gruber, 1995). The investment constraints in (6) do not allow short selling in the N risky assets and impose an upper bound constraint of 20% on each risky asset. We also estimate the strategies without any investment constraints.

We construct the domestic asset allocation strategies as follows. At the start of February 1981, we estimate the expected excess return vector and covariance matrix over the prior 60 months (i.e. between February 1976 and January 1981). We use different models to estimate expected excess returns. We use the sample covariance matrix estimated over the prior 60 months as the estimate of the covariance matrix. We use the same covariance matrix across models.[3] Given estimates of expected excess returns and the covariance matrix, we solve the mean-variance optimization for a given level of risk tolerance.[4] We calculate the actual excess returns of the optimal portfolio in February 1981. We repeat this process each month until April 2000.

Performance Measures

We evaluate the out of sample performance of the asset allocation strategies using three different performance measures. This includes the Sharpe (1966) measure, the unconditional Jensen (1968) measure and the conditional Ferson and Schadt (1996) measure. We calculate the Sharpe (1966) measure as the mean excess return divided by the standard deviation of excess returns. We use the Sharpe measure to rank the performance of the strategies against one another. The Jensen (1968) and Ferson and Schadt (1996) performance measures evaluate the performance of the strategy relative to a benchmark model. The Jensen (1968) measure is an unconditional performance measure and assumes that expected returns and betas are constant through time. The Ferson and Schadt (1996) measure is a conditional performance measure and allows expected returns and betas to vary through time.

We estimate the Jensen (1968) measure by the following regression:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it} \tag{7}$$

where r_{mt} is the excess returns on the domestic market index, β_i is the factor exposure of strategy i relative to the market index, and ε_{it} is a random error term with $E(\varepsilon_{it}) = 0$ and $E(\varepsilon_{it} r_{mt}) = 0$. The intercept α_i is the Jensen (1968) performance measure that equals zero under the null hypothesis that strategy i exhibits zero abnormal performance. A positive α_i is usually interpreted as superior performance and a negative α_i as inferior performance. The interpretation of the Jensen measure is controversial because we do not know if the benchmark model is the true asset pricing model that is able to correctly price the cross-section of asset returns.[5] However the Jensen measure can be viewed as the abnormal returns of the strategy compared to an alternative passive strategy that invests in the risk-free asset and the market index that have the same risk characteristics as the strategy (see Elton & Gruber, 1995). We refer to performance from Eq. (7) as the Jensen measure.

We estimate the Ferson and Schadt (1996) performance measure by the following regression:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \sum_{l=1}^{L} \delta_{il} r_{mt} z_{lt-1} + \varepsilon_{it} \tag{8}$$

z_{lt-1} is the de-meaned lth information variable ($l = 1, \ldots, L$) realization at time $t - 1$ and β_i is the average conditional beta of strategy i with respect to each of the benchmark portfolios. The δ_{ilk} coefficients are the estimated coefficients of strategy i in the conditional beta function with respect to each of the L information variables. The intercept α_i is the Ferson and Schadt (1996) measure which equals zero under the null hypothesis of the strategy exhibiting no abnormal performance. The conditional performance measure allows the strategy betas and factor risk premia to vary through time. Ferson and Schadt (1996) use the approximation that the portfolio beta in period t is a linear function of common information variables that are known at that time. The additional terms, $r_t^k z_{lt-1}$ ($l = 1, \ldots, L$), captures the covariance between the conditional betas and the market risk premium. We refer to performance from Eq. (8) as the FS performance measure.

We correct the test statistics of the performance measures for the effects of heteroskedasticity using White (1980).

DATA

Investment Universe

We explore the impact of the different models on domestic U.K. asset allocation strategies. We follow Fletcher (1997) and Grauer (2000) and use industry portfolios as our investment universe. An alternative approach is to use large portfolios of individual securities as in Chou, Li and Zhou (2001). However, the focus on this study is on asset allocation strategies rather than stock selection strategies. We leave the issue of stock selection strategies to future research.

We use 10 U.K. industry portfolios and the domestic risk-free asset as the investment universe. We collect the data from Datastream unless otherwise specified. The industry portfolios are the industry portfolios constructed by Datastream. We choose the following industrial sectors:[6]

(1) Resources
(2) Basic industries
(3) General industrials
(4) Cyclical consumer goods
(5) Non cyclical consumer goods
(6) Cyclical services
(7) Non cyclical services
(8) Information technology
(9) Financials
(10) Telecom, media and IT

We collect monthly excess returns on the 10 industry portfolios between February 1976 and April 2000. The industry indices are value-weighted. We use the monthly return on a one month U.K. Treasury Bill as the risk-free return. In addition to monthly excess returns, we collect various characteristics of the industry portfolios. This includes the market value[7], dividend yield and price-earnings ratio. We collect the values of these characteristics at the end of each month between January 1976 and March 2000. Table 1 reports summary statistics of the industry portfolios. Panel A of the table includes the mean and standard deviation of monthly excess returns and the average value of the three characteristics. Panel B reports the correlations between the 10 industry portfolios.

The mean monthly excess returns in panel A of Table 1 ranges between 0.073% (Cyclical consumer goods) and 1.107% (Non cyclical services). The standard deviations range between 5.344% (Non cyclical consumer goods) and 9.321% (Information technology). The summary statistics suggest that return and total risk tradeoff varies considerably across industry sectors. The cyclical consumer goods sector has the smallest mean excess return but also the second highest standard deviation. In contrast the non cyclical services sector has the highest mean excess returns and is in the middle of the range of standard deviations.

The average values of the characteristics in panel A of Table 1 also show a great deal of cross-sectional variation across sectors. The cyclical consumer goods sector has the smallest average market value and highest average dividend yield. The information technology sector has one of the smallest average dividend yield and highest average PE ratio. The average PE ratio of the information technology sector stands out sharply from the other groups.

The financials sector has the highest average market value and the lowest average PE ratio.

The correlations in panel B of Table 1 range between 0.222 and 0.860. Most of the correlations are in excess of 0.6 but there are some interesting observations. The information technology sector stands out sharply from the other sectors in that it has a small positive correlation with virtually all other

Table 1. Summary Statistics of Industry Portfolios.

Panel A

	Mean	Std Deviation	Size	DY	PE
1 Resources	0.643	6.457	44,280	5.145	16.90
2 Basic	0.248	6.119	30,031	5.243	12.931
3 General	0.494	6.379	23,018	3.991	14.172
4 Cyc CG	0.073	7.913	3,003	5.762	14.073
5 Ncyc CG	0.676	5.344	69,252	4.422	13.941
6 Cyc Ser	0.541	5.611	73,082	4.039	15.95
7 Ncyc Ser	1.107	6.148	43,952	3.385	17.554
8 Info Tech	0.807	9.321	3,974	3.443	28.01
9 Financials	0.596	5.736	84,226	4.827	11.981
10 Telecom, Media	0.832	6.533	55,038	4.159	16.296

Panel B Correlations

	1	2	3	4	5	6	7	8	9
2	0.600								
3	0.544	0.848							
4	0.460	0.754	0.669						
5	0.506	0.759	0.707	0.622					
6	0.545	0.860	0.847	0.699	0.776				
7	0.356	0.600	0.615	0.419	0.644	0.741			
8	0.222	0.294	0.386	0.323	0.266	0.389	0.348		
9	0.599	0.798	0.738	0.652	0.800	0.799	0.632	0.273	
10	0.440	0.726	0.710	0.618	0.680	0.838	0.838	0.433	0.698

Note: Summary statistics of the excess returns and characteristics of 10 U.K. industry portfolios are estimated between February 1976 and April 2000. Panel A includes the mean and standard deviation (monthly %) of monthly excess returns as well as the average value of the industry characteristics in terms of market value (Size – £m), dividend yield (DY) and price-earnings ratio (PE). Panel B reports the correlations between the 10 portfolios.

sectors. The financials sectors tends to be highly positively correlated with most sectors.

Models of Expected Return

We use four linear asset pricing models to estimate expected excess returns. This includes:

CAPM
This is a single factor model. We use the excess returns on the Financial Times All Share index (FTA)[8] as the market index.

Fama and French (1993) (FF)
This is a three factor model similar to Fama and French (1993). The first factor is the excess return on the FTA index. The second factor is the excess return on a small stock index (Size). Up until February 1993, this factor is the excess return on an equally-weighted portfolio of the bottom decile of stocks on the London Business School Share Price Database which is reformed each year. From February 1993 onwards, the FTSE small stock index is used as the small stock index. The third factor captures the Value/Growth differential in stock returns. This factor is the monthly return difference between the Morgan Stanley Capital International (MSCI) U.K. value equity index and growth equity index (V-G).

Elton, Gruber and Blake (1996) (EGB)
This is a four factor model similar to Elton, Gruber and Blake (1996). The first three factors are the same as the FF model. The fourth factor is the excess return on the FT U.K. government bonds (all stocks) index (bonds).

APT
This is a four factor model based on the APT. We use four economic risk factors:

(i) Market – excess return on the FTA index.
(ii) Term structure (Term) – excess return on long-term U.K. government bonds (greater than 15 years).
(iii) Industrial production (IP) – log difference in the U.K. industrial production index.
(iv) Inflation (Inf) – log difference in the U.K. retail price index.

Since factors (i) and (ii) are already portfolio returns, we do not require to construct mimicking portfolios (see Shanken, 1992). We construct mimicking portfolios of the factors (iii) and (iv) using an approach similar to Connor and Korajcyzk (1991). First, we regress the demeaned values of the factors on a constant and the demeaned values of the 10 industry portfolio excess returns.[9] Second, we multiply the coefficients from the first regression by the actual excess returns on the industry portfolios to get the mimicking portfolio for that factor.

Table 2 reports summary statistics of the different factors in the models. This table includes the mean and standard deviation of excess returns and correlations between the factors for each of the four models.

Table 2 shows that the FTA market index and the small stock index have the highest average excess returns over the sample period. These factors are both significantly positive at the 5% significance level. The other average excess returns are insignificantly different from zero and are considerably smaller in magnitude. The magnitude of the Value/Growth differential in U.K. stock returns is fairly small over this sample period. The correlations between the factors in most cases tends to be small and close to zero.

Table 2. Summary Statistics on Factors.

Panel A

	FTA	Size	V-G	Bonds	Term	IP	Inf
Mean	0.563	0.569	0.128	0.147	0.263	0.012	–0.008
Std Dev	5.043	4.935	2.696	2.232	3.250	0.194	0.162

Panel B

FF	FTA	Size		**EGB**	FTA	Size	V-G
Size	0.652			Size	0.652		
V-G	–0.044	0.101		V-G	–0.044	0.101	
				Bonds	0.443	0.183	–0.054
APT	FTA	Term	IP				
Term	0.422						
IP	–0.032	–0.031					
Inf	–0.048	–0.091	0.004				

Note: Summary statistics on the excess returns on the FTA market index, small stock index (Size), Value-Growth index (V-G), excess returns of FT government bond index (Bonds), long-term government bond index (Term), U.K. industrial production (IP) and inflation (Inf) are estimated between February 1976 and April 2000. Panel A includes the mean and standard deviation (monthly %) of excess returns for the factors. Panel B reports the correlations between the factors for the three multifactor models. The models include Fama and French (1993) (FF), Elton, Gruber and Blake (1996) (EGB) and Arbitrage Pricing Theory (APT).

Our final model of expected returns is based on the characteristics of the industry portfolios.

Characteristics Model
This is a four factor model. We use the dividend yield, price-earnings ratio, size and prior year excess return (lagged one month) for the industry portfolios. Chan, Karceksi and Lakonishok (1998) show that each of the characteristics have strong predictive power in U.K. stock returns. A notable exception to the list is the book-to-market ratio. We do not include the industry book-to-market rations as they were not available on Datastream. However, Dimson, Nagel and Quigley (2001) find that the book-to-market and dividend yield effects are very highly correlated and dividend yield is a good proxy of the value effect in U.K. stock returns.

Information Variables and Benchmark

We use the FTA index as the benchmark portfolio for the Jensen and FS performance measures. To estimate the conditional versions of the linear asset pricing models, we require to specify the information set of investors. We use instruments that studies have found to be helpful in predicting U.K. stock returns (Solnik, 1993; Fletcher, 1997).[10] The instruments include the lagged dividend yield on the FTA index; lagged risk-free return; lagged excess return on the FTA index; January dummy which equals one in the month of January and zero otherwise.

EMPIRICAL RESULTS

We initially examine the performance of the asset allocation strategies that use linear asset pricing models using the unconditional versions of the models. Tables 3 and 4 report the performance results for the cases of no investment restrictions (Table 3) and where restrictions (Table 4) are imposed. Panel A of each table contains summary statistics of the performance of the five asset allocation strategies and the FTA index. Panel B reports the estimated performance measures and corresponding *t*-statistics of the five strategies.

Panel A of Table 3 shows that the strategy which uses the characteristics model to forecast expected excess returns has the second highest Sharpe performance across the five strategies when no investment restrictions are imposed. However the Sharpe performance of the strategy which uses the characteristics model underperforms that of the market index. This strategy is characterised by high mean excess returns and standard deviations. This differs

Table 3. Performance of Asset Allocation Strategies: Unrestricted.

Panel A

	Mean	Std Deviation	Sharpe	Minimum	Maximum
Char	0.567	15.77	0.036	–58.98	71.00
CAPM	0.138	2.68	0.051	–28.29	7.65
FF	–0.233	3.59	–0.065	–39.57	8.85
EGB	–0.322	3.69	–0.087	–39.59	9.04
APT	0.051	3.88	0.013	–28.10	7.49
Market	0.599	4.89	0.123	–31.56	12.01

Panel B

	Char	CAPM	FF	EGB	APT
Jensen	0.233	–0.113	–0.504	–0.598	–0.175
	(0.23)	(–0.69)	(–1.90)	(–2.24)*	(–0.69)
FS	0.238	0.103	–0.222	–0.322	0.020
	(0.23)	(1.39)	(–1.64)	(–2.19)*	(0.10)

* Significant at 5%.

Note: The out of sample performance of monthly asset allocation strategies is evaluated between February 1981 and April 2000. Expected returns are estimated from a characteristics model of stock returns (Char) and unconditional versions of linear asset pricing models of the CAPM, Fama and French (1993) (FF), Elton, Gruber and Blake (1996) (EGB) and APT models. Panel A includes summary statistics of performance of the five strategies and FTA market index that also includes the Sharpe (1966) measure. Panel B reports the performance measures of Jensen (1968) measure and the conditional measure of Ferson and Schadt (1996) (FS). The *t*-statistics (in parentheses) are corrected for heteroskedasticity using White (1980). The asset allocation strategies are estimated without any investment constraints. All performance numbers are monthly %.

from the other strategies where the standard deviations are lower. However in spite of the low Sharpe performance, the strategy that uses the characteristics model generates positive abnormal returns using either the Jensen or FS performance measures. The lack of statistical significance is due to the high volatility in the portfolio excess returns.

In contrast to the strategy which uses the characteristics model of stock returns, the strategies that use the multifactor models to forecast expected excess returns tend to perform poorly. The strategy that uses the CAPM has the highest Sharpe performance across all five strategies and the best abnormal returns across the four strategies that use linear asset pricing models. All of the

Table 4. Performance of Asset Allocation Strategies: Restricted.

Panel A

	Mean	Std Deviation	Sharpe	Minimum	Maximum
Char	0.335	3.26	0.103	–25.01	14.73
CAPM	0.143	2.67	0.054	–28.34	7.58
FF	0.204	2.82	0.072	–30.13	6.43
EGB	0.194	2.85	0.068	–30.12	6.70
APT	0.241	2.82	0.085	–28.16	6.88

Panel B

	Char	CAPM	FF	EGB	APT
Jensen	0.005	–0.111	–0.066	–0.082	–0.033
	(0.27)	(–0.69)	(–0.39)	(–0.49)	(–0.22)
FS	0.204	0.097	0.142	0.123	0.153
	(1.69)	(1.37)	(1.96)	(1.67)	(1.69)

* Significant at 5%.

The out of sample performance of monthly asset allocation strategies is evaluated between February 1981 and April 2000. Expected returns are estimated from a characteristics model of stock returns (Char) and unconditional versions of linear asset pricing models of the CAPM, Fama and French (1993) (FF), Elton, Gruber and Blake (1996) (EGB) and APT models. Panel A includes summary statistics of performance of the five strategies that also includes the Sharpe (1966) measure. Panel B reports the performance measures of Jensen (1968) measure and the conditional measure of Ferson and Schadt (1996) (FS). The *t*-statistics (in parentheses) are corrected for heteroskedasticity using White (1980). The asset allocation strategies are estimated with no short selling allowed in the risky assets and an upper bound constraint of 20% in each risky asset. All performance numbers are monthly %.

strategies that use linear asset pricing models tend to underperform passive combinations of the risk-free asset and the domestic market index. There is an improvement in the out of sample performance using the FS measure. The underperformance is substantial for the strategies which use the FF and EGB models. These latter two strategies also have negative Sharpe performance.

When investors face binding investment constraints as in Table 4, the Sharpe performance of all five asset allocation strategies increases. The strategies now have considerably lower volatility. The strategy that uses the characteristics model has the highest Sharpe performance across the five strategies. However it still underperforms the market Sharpe measure. This strategy also generates significant positive performance (at the 10% significance level) using the FS measure.

The impact of investment restrictions in Table 4 has little impact on the performance of the strategy which uses the CAPM model. However the effect is dramatic on the performance of the strategies which use multifactor models. This is especially the case for the strategies that use the FF and EGB models. The strategies which use the FF and EGB models no longer exhibit negative Sharpe performance but now have a higher Sharpe performance than the strategy that uses the CAPM. Furthermore the abnormal returns of these two strategies are now close to zero for the Jensen measure and significantly positive using the FS measure (at the 10% significance level).

The evidence that imposing portfolio constraints can improve performance for some strategies is consistent with Frost and Savarino (1988). This improvement appears here for those strategies that exhibit poor out of sample performance when weights are unrestricted. It is well known that a problem with traditional mean-variance analysis that uses sample estimates as the inputs can be highly unstable (Michaud, 1989, 1998). This problem arises because mean-variance optimizers tend to maximize estimation risk. Mean-variance analysis favors assets with high expected returns, low variance and lower correlations. Michaud (1989) argues that these assets are likely to have the highest estimation risk which can lead to extreme portfolios with poor out of sample performance. One solution is to constrain the portfolio weights (Frost & Savarino, 1988). The imposition appears to work here as the performance of some of the five asset allocation strategies improves. However in spite of the improved performance, the evidence in Tables 3 and 4 is mixed as to whether active asset allocation strategies outperform alternative passive strategies whatever approach is followed.

The next issue addressed is to examine the performance of the asset allocation strategies that use linear asset pricing models whenever conditional versions of the models are used. Table 5 reports the out of sample performance of the four asset allocation strategies that use linear asset pricing models. Panel A refers to the case where there are no investment constraints and panel B refers to the case where investment restrictions are imposed.

Table 5 shows that there is a dramatic improvement in the performance of the strategies that use linear asset pricing models whenever conditional versions of the models are used. All four strategies now outperform the strategy that uses the characteristics model in terms of higher Sharpe performance and more positive abnormal returns. When investors face no binding investment constraints, as in panel A of Table 5, all four strategies that use linear factor models have a higher Sharpe performance than the FTA market index and the

Table 5. Performance of Asset Allocation Strategies: Conditional Models.

Panel A

	Mean	Std Deviation	Sharpe	Minimum	Maximum
CAPM	1.132	3.27	0.346	–6.89	24.56
FF	2.938	5.94	0.495	–9.49	28.16
EGB	2.791	6.39	0.437	–13.56	28.86
APT	2.965	9.85	0.301	–29.82	92.09
	CAPM	FF	EGB	APT	
Jensen	1.155	2.957	2.840	3.132	
	(4.15)*	(7.15)*	(6.28)*	(4.62)*	
FS	0.831	2.798	2.664	2.695	
	(3.34)*	(5.89)*	(5.32)*	(3.99)*	

Panel B

	Mean	Std Deviation	Sharpe	Minimum	Maximum
CAPM	0.845	2.27	0.372	–6.89	16.41
FF	1.052	2.70	0.389	–6.13	15.83
EGB	1.063	2.73	0.389	–7.28	16.25
APT	0.920	2.51	0.366	–5.61	12.01
	CAPM	FF	EGB	APT	
Jensen	0.687	0.852	0.864	0.743	
	(5.34)*	(5.53)*	(5.48)*	(5.19)*	
FS	0.505	0.673	0.675	0.578	
	(4.31)*	(5.26)*	(5.55)*	(4.36)*	

* Significant at 5%.

The out of sample performance of monthly asset allocation strategies is evaluated between February 1981 and April 2000. Expected returns are estimated from the conditional versions of linear asset pricing models of the CAPM, Fama and French (1993) (FF), Elton, Gruber and Blake (1996) (EGB) and APT models. Each panel includes summary statistics of performance and two performance measures with *t*-statistics in parentheses. The two performance measures are the Jensen (1968) measure and the conditional measure of Ferson and Schadt (1996) (FS). The *t*-statistics are corrected for heteroskedasticity using White (1980). Panel A refers to the case where there are no investment constraints and panel B refers to the case where there are short selling and 20% upper bound constraints in the risky assets. All performance numbers are monthly %.

strategy that uses the characteristics model. The strategy that uses the FF model has the highest Sharpe performance across the four strategies in panel A of Table 5. This is more than three times larger than the corresponding Sharpe performance of the FTA market index. This stands in sharp contrast to the

Sharpe performance in Table 3 where this strategy has a negative Sharpe performance.

The abnormal returns of the four strategies in panel A are highly statistically significant and economically large. The strategy that uses the CAPM model has the lowest abnormal returns across the four strategies, because this strategy has a lower mean excess return compared to the other three strategies. All four strategies have a negative beta on the FTA index when performance is estimated using the Jensen measure. The strategy that uses the APT model has the highest performance with the Jensen measure and the FF model has the highest performance with the FS measure.

When investors face binding investment constraints as in panel B of Table 5, all four strategies still deliver a higher Sharpe performance than either the FTA market index or the strategy that uses the characteristics model. The performance across the four strategies manifests less variation compared to that in panel A. High mean excess returns and low standard deviation characterize all four strategies. The smaller means leads to lower abnormal returns in panel B compared to those in panel A. However the Jensen and FS performance of the four strategies are still highly significantly positive and economically large. The strategies that use the EGB and FF models exhibit the highest positive performance across the four strategies.

The evidence in Table 5 suggests that strategies which use linear asset pricing models to forecast expected excess returns are able to significantly outperform alternative passive strategies whenever conditional versions of the models are used. Furthermore the strategies which rely on linear asset pricing models are able to outperform the strategy that uses the characteristics model in terms of higher Sharpe performance and more positive abnormal returns. These findings support the evidence in Jagannathan and Wang (1996), Hodrick and Zhang (2001), and Lettau and Ludvigson (2001) among others that conditional asset pricing models are more able to explain the cross-section of stock returns.

CONCLUSIONS

This paper examines the out of sample performance of domestic asset allocation strategies in U.K. stock returns between February 1981 and April 2000 where linear asset pricing models and a characteristics model of stock returns are used to forecast expected excess returns. Our findings suggest that when unconditional versions of asset pricing models are used, the performance of the strategies that use multifactor asset pricing models is poor. However this all changes when we use conditional versions of the asset pricing models.

Active asset allocation strategies that use linear asset pricing models now outperform the strategy based on the characteristics model in terms of a higher Sharpe performance and more positive abnormal returns. Furthermore, all strategies that use linear asset pricing models are able to significantly outperform alternative passive strategies even when the investor faces binding investment constraints. The improved performance of linear asset pricing models when we use conditional versions of the models supports a number of studies from the empirical asset pricing literature that document that conditional versions of asset pricing models are more able to explain cross-sectional patterns in stock returns (see Cochrane (1996), Jagannathan and Wang (1996), Hodrick and Zhang (1996) and Lettau and Ludvigson (2001)). Our findings support the usefulness of conditional asset pricing models in forecasting expected excess returns in domestic asset allocation strategies and provides some support for the use of these models relative to a characteristic based model of stock returns.

NOTES

1. The momentum effect stems from Jegadeesh and Titman (1993).
2. We focus on expected returns because Merton (1980) points out that estimates of expected returns are more unstable than the covariance matrix and Best and Grauer (1991) document the sensitivity of optimal portfolio weights to even small changes in expected returns.
3. Jagannathan and Ma (2001) find that the sample covariance matrix performs just as well as other estimators of the covariance matrix when investors face binding investment constraints.
4. This is set equal to 0.1. Using alternative values of t has no impact on the analysis for the performance measures used in this study.
5. See Grinblatt and Titman (1989), Chen and Knez (1996) for a discussion of these points.
6. We do not include the Utilities sector because data is not available for the whole period.
7. When the characteristics model is estimated, the ln (market value) is used.
8. The FTA index is a value-weighted index of the largest companies on the London Stock Exchange.
9. Connor and Korajcyzk (1991) regress the demeaned values of the factors on statistical factors derived from asymptotic principal components analysis.
10. Keim and Stambaugh (1986) and Fama and French (1988) among others also show that stock and bond returns are partly predictable through time. Fama (1991) and Cochrane (1999) provide a review of stock return predictability.

ACKNOWLEDGEMENT

Helpful comments received from an anonymous reviewer.

REFERENCES

Best, M. J., & Grauer, R. A. (1990). The efficient set mathematics when mean-variance problems are subject to general linear constraints. *Journal of Economics and Business*, *42*, 105–120.

Best, M. J., & Grauer, R. A. (1991). On the sensitivity of mean-variance portfolios to changes in asset means: Some analytical and computational results. *Review of Financial Studies*, *4*, 315–342.

Chan, L. K. C., Karceski, J., & Lakonishok, J. (1998). The risk and return from factors. *Journal of Financial and Quantitative Analysis*, *33*, 159–188.

Chen, Z., & Knez, P. J. (1996). Portfolio performance measurement: Theory and applications. *Review of Financial Studies*, *9*, 511–555.

Chou, P. H., Li, W. S., & Zhou, G. (2001). Portfolio optimization under asset pricing anomalies. Working Paper. Washington University in St Louis.

Cochrane, J. H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy*, *104*, 572–621.

Cochrane, J. H. (1999). New facts in Finance. *Economic Perspectives*, *23*, 36–58.

Connor, G., & Korajcyzk, R. A. (1991). The attributes, behavior and performance of U.S. mutual funds. *Review of Quantitative Finance and Accounting*, *1*, 5–26.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross-sectional variation in stock returns. *Journal of Finance*, *52*, 1–33.

Daniel, K., & Titman, S. (1998). Characteristics or covariances? *Journal of Portfolio Management*, *24*, 24–33.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic based benchmarks. *Journal of Finance*, *52*, 1035–1058.

Daniel, K., Titman, S., & Wei, K. C. J. (2001). Explaining the cross-section of stock returns in Japan: Factors or Characteristics? *Journal of Finance*, *56*, 743–766.

Davies, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, covariances and average returns: 1929 to 1997. *Journal of Finance*, *55*, 389–406.

Dimson, E., Nagel, S., & Quigley, G. (2001). Value versus growth in the U.K. stock market, 1955 to 2000. Working Paper. London Business School.

Elton, E. J., & Gruber, M. J. (1995). *Modern portfolio theory and investment analysis* (5th ed.). Wiley.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). The persistence of risk-adjusted mutual fund performance. *Journal of Business*, *69*, 133–157.

Fama, E. F. (1991). Efficient capital markets II. *Journal of Finance*, *46*, 1575–1617.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*, 3–25.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, *33*, 3–56.

Fama, E. F., & French, K. R. (1995). Size and book-to-market effects in earnings and returns. *Journal of Finance*, *50*, 131–155.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance*, *51*, 55–84.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return and equilibrium. *Journal of Political Economy*, *81*, 607–636.

Ferson, W. E., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, *99*, 385–415.

Ferson, W. E., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of Financial Studies*, *6*, 527–566.

Ferson, W. E., & Korajcyzk, R. A. (1995). Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business*, *68*, 309–350.

Ferson, W. E., & Locke, D. H. (1998). Estimating the cost of capital through time: An analysis of the sources of error. *Management Science*, *44*, 485–500.

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, *51*, 425–461.

Fletcher, J. (1997). An investigation of alternative estimators of expected returns in mean-variance analysis. *Journal of Financial Research*, *20*, 129–143.

Frost, P. A., & Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *Journal of Portfolio Management*, 29–34.

Ghysels, E. (1998). On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance*, *53*, 549–573.

Grauer, R. R. (2000). On the predictability of stock market returns: Evidence from industry rotation strategies. Working Paper. Simon Fraser University.

Grinblatt, M., & Titman, S. (1987). The relation between mean-variance efficiency and arbitrage pricing. *Journal of Business*, *60*, 97–112.

Grinblatt, M., & Titman, S. (1989). Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies*, *2*, 393–422.

Haugen, R. A., & Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, *41*, 401–439.

Hodrick, R. J., & Zhang, X. (2001). Evaluating the specification errors of asset pricing models. *Journal of Financial Economics*, *62*, 327–376.

Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, *51*, 3–53.

Jagannathan, R., & Ma, T. (2001). Risk reduction in large portfolios: A role for portfolio weight constraints. Working Paper. Northwestern University.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, *48*, 65–91.

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945–1964. *Journal of Finance*, *23*, 389–416.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting return in the bond and stock market. *Journal of Financial Economics*, *17*, 357–390.

Lettau, M., & Ludvigson, S. (2001). Resurrecting the C (CAPM): A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, forthcoming.

Merton, R. C. (1980). On estimating the expected return on the market. *Journal of Financial Economics*, *8*, 323–361.

Michaud, R. O. (1989). The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, *45*, 31–42.

Michaud, R. O. (1998). *Efficient Asset Management*. Harvard Business School Press.

Pastor, L., & Stambaugh, R. F. (2000). Comparing asset pricing models: An investment perspective. *Journal of Financial Economics*, *56*, 335–381.

Shanken, J. (1992). On the estimation of beta-pricing benchmark specifications. *Review of Financial Studies*, *5*, 1–55.

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

Solnik, B. (1993). The performance of international asset allocation strategies using conditioning information. *Journal of Empirical Finance*, *1*, 33–56.

White, H. (1980). A heteroskedasticity consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica*, *48*, 817–838.

This Page Intentionally Left Blank