On the Predictability of Stock Market Returns: Evidence from Industry-Rotation Strategies[1]

Robert R. Grauer
Simon Fraser University

This paper evaluates historic, Bayes-Stein, Capital Asset Pricing Model (CAPM) and dividend-yield riskfree-rate estimators of asset means using statistical and economic criteria. None of the estimators exhibit much in the way of out-of-sample predictive ability when judged by statistical criteria. Yet, when combined with a discrete-time power-utility portfolio selection model, all the estimators generate economically significant returns judged in terms of compound return – standard deviation plots and accumulated wealth. Even so, the portfolios generated from dividend-yield riskfree-rate estimators perform by far the best and portfolios generated from traditional CAPM estimator perform the worst. For the most part, commonly accepted statistical measures of investment performance support these rankings.

How do we judge whether returns are predictable? We could regress returns on past returns or on information variables that might include accounting data, dividend yields, riskfree interest rates and other macroeconomic indicators. Within this framework we could judge predictability in terms of in-sample slope coefficients and R-square values. Of course, out-of-sample measures of statistical significance—R-

[1]Earlier versions of the paper were presented at the VII International Conference on Stochastic Programming at the University of British Columbia in Vancouver, the Northern Finance Association meetings in Calgary, the Pacific Northwest Finance Conference in Seattle, the University of Victoria, the Western Decision Sciences Meetings in Vancouver, and the Decision Sciences Meetings in San Francisco. The author thanks the Social Sciences Research Council of Canada for financial support; seminar participants, Nils Hakansson, Kris Jacobs, Peter Klein, Anthony Lynch, Ieuan Morgan, Mark Rubinstein, William Ziemba and an anonymous reviewer for helpful comments; and Reo Audette, Steven Bovencamp, Poh Chung Fong, Christopher Fong, John Janmaat and William Ting for their most capable assistance.

squares or mean square errors—would lend more credence to any claims of predictability. However, Leitch and Tanner (1991) point out that statistical measures of predictability may not shed much light on the economic value of a forecast. They show that commercial interest rate forecasts do not perform any better than naïve forecasts when evaluated in terms of out-of-sample statistical criteria. But, the commercial (naïve) forecasts generate economically profitable (unprofitable) trading strategies. This paper evaluates historic, Bayes-Stein, CAPM and dividend-yield riskfree-rate estimators of asset means using statistical and economic criteria. None of the estimators exhibit much in the way of out-of-sample predictive ability when judged by statistical criteria. Yet, when combined with a discrete-time power-utility portfolio selection model, all the estimators generate economically significant returns. The results are in agreement with Leitch and Tanner's: evaluating the estimators in terms of out-of-sample statistical criteria sheds little light on their economic value.

Grauer and Hakansson (1986, 1987) and Grauer, Hakansson and Shen (1990), apply a discrete-time power-utility model in conjunction with the empirical probability assessment approach (EPAA) in domestic, global and industry-rotation asset-allocation settings. The results are noteworthy for two reasons. First, the model often generates economically and statistically significant abnormal returns. Second, no attempt is made to correct for estimation error, which is clearly present in the EPAA. It would seem prudent, therefore, to examine the effects of making corrections for estimation error.

One approach adjusts for estimation error by estimating the means based on statistical, financial or forecasting models. The results of applying Stein and CAPM instead of historical estimators of the means are mixed. Evidence based on mean-variance (MV) portfolio selection, simulation analysis and out-of-sample portfolio performance suggests that Stein and CAPM estimators of the means can improve investment performance substantially (Jobson, Korkie & Ratti, 1979; Jobson & Korkie, 1980, 1981; Jorion, 1985, 1986, 1991). On the other hand, Grauer and Hakansson (1995) find that although the Stein estimators outperform the sample (historic) estimator in an industry-rotation setting, the gains are not as great as those reported by others. Moreover, in a global setting just the opposite is true: the sample estimator outperforms the Stein estimators. In all cases, the CAPM estimator exhibits the worst performance, which is just the opposite of what Jorion (1991) finds in an industry setting using MV analysis that allows short sales. In light of these contradictory results, this paper examines the effects of adding two dividend-yield riskfree-rate estimators to the mix.

A second approach adjusts for estimation risk by constraining portfolio weights. The results are again somewhat mixed. Employing simulation and MV analysis, Frost and Savarino (1988) report that imposing upper bounds both reduces estimation bias and improves performance. In a companion paper to this one, Grauer (2007) shows that the portfolios of less risk-averse MV investors generated from dividend-yield riskfree-rate estimators of the means bankrupt in an out-of-sample industry-rotation setting with short sales permitted. Yet, with short sales precluded, the dividend-yield riskfree-rate portfolios of these less risk-averse MV investors exhibit the best performance. Moreover, portfolios generated from the CAPM estimator, which display the best performance with short sales permitted, exhibit the worst performance with short sales precluded. On the other hand, Grauer and Shen (2000), while employing the discrete-time power-utility model in an out-of-sample setting with short sales precluded, reported that constraining the portfolio weights further led to appreciably more diversification and less realized risk. But the cost is a less realized return.

Mean estimators trace their origins to different parts of the statistics and finance literature. The mean square error properties of the historic estimator make it an obvious choice for an estimate of a mean. Stein estimators are based on purely statistical arguments that minimize the mean square error of a vector of means and completely ignore risk-return tradeoffs that may be helpful in predicting stock returns. CAPM estimators fill this void by drawing on the best-known financial model of asset pricing. Dividend-yield riskfree-rate estimators, as well as estimators based on other information variables, trace their origin to the return predictability or weak form efficient markets literature (Fama, 1991). Return predictability is of interest not only because of its fundamental implications for market efficiency, but also because it is steeped in controversy. There is disagreement about whether returns are predictable and, if they are, whether predictability implies market inefficiency or is a result of rational variation in expected returns.

In the early to mid-1990s there was a consensus, based on statistical criteria, that stock market returns could be predicted from informational variables—at least over one-year to four-year decision horizons. (Fama & French, 1988, 1989; Fama, 1991; Hawawini & Keim, 1995). The importance of this evidence extended beyond statistical considerations, as it helped renew the interest in continuous-time portfolio choice, hedging demand and non-myopic investment decisions discussed below. Fama and French (1988) show that the power of dividend yields to forecast stock returns increases with the return horizon. The monthly and quarterly results are unimpressive, with R-squares on the order of 0.01. But, with four-year returns, the R-squares range from 0.13 to 0.64. More impressive, out-of-sample R-squares from forecasts made with coefficients estimated from 30-year rolling regressions are close to the in-sample R-squares for all return horizons.

The consensus began to crack through the 1990s and into the new century. Lo and MacKinlay (1990) and Foster, Smith and Whaley (1997) are concerned with data mining. Others (Hodrick, 1992; Goetzmann & Jorion, 1993; Goyal & Welch, 2003; Ang & Bekaert, 2003) question the long-horizon results on statistical grounds. Goetzmann and Jorion (1993) and Stambaugh (1999) study the biases due to dependent stochastic regressors. Ferson, Sarkissian and Simin (2003) question whether there is a spurious regression bias in predictive regressions. Bossaerts and Hillion (1999) examine the statistical significance of a variety of informational variables using monthly data in an international setting. They confirm the presence of in-sample predictability, but report that even the best prediction models have no out-of-sample forecasting power. Goyal and Welch (2003) confirm Bossaerts and Hillion's evidence, while Pesaran and Timmermann (1995) report contradictory results. Perhaps surprisingly in light of the early evidence on long-horizon predictability and the out-of-sample evidence in Bossaerts and Hillion, Goyal and Welch and this paper, Ang and Bekaert (2003) and Torous, Valkanov and Yan (2005) report that the predictive power of dividend yields is best visible at short horizons—with the short rate as an additional regressor in Ang and Bekaert's case.

A number of authors (Breen, Glosten & Jagannathan, 1989; Fuller & Kling, 1990; Pesaran & Timmermann, 1994; Larsen & Wozniak, 1995; Pelaez, 1998; Schwert, 2003) investigate the economic value of trading rules based on predictive regressions, when the predictions are not combined with a MV or power-utility portfolio selection model. The results are mixed and sample-period dependent. Schwert, for example, examines a strategy of investing in short-term bonds when a dividend yield model predicts stock returns are lower than interest rates. The model predicts poorly during the 1990s. Schwert (2003: 953) concludes: "In short, the out-of-sample prediction performance of this model would have been disastrous." But, this paper shows that when dividend-yield riskfree-rate forecasts are combined with the discrete-time power-utility model, the results through the 1990s are anything but disastrous.

The single-period MV model, the discrete-time power-utility model and the continuous-time power-utility model either are or have been combined with forecasts based on information variables in order to determine the economic value of the forecasts. Contributors to the MV literature include: (Solnik, 1993; Klemkosky & Bharati, 1995; Connor, 1997; Beller, Kling & Levinson, 1998; Ferson & Seigel, 2001; Marquering & Verbeek, 2001; Fletcher & Hillier, 2002; Avramov, 2004; Avramov & Chordia, 2006). They examine the portfolio returns of MV investors who exhibit "average" degrees of risk aversion and revise their portfolios monthly. While these papers report economically significant returns in U.S. bond-stock, U.S. industries and international settings, none report results from before 1960. The benefits of the MV model include familiarity, ease of estimation—only the means, variances and covariances need to be estimated—and ease of computation.

This paper employs a discrete-time power-utility model that embodies a broad range of risk-aversion characteristics, quarterly decision horizons, borrowing and lending at different rates and an industry dataset that spans the 1934-99 period. The primary benefit of this model is the formal justification of a myopic decision rule. If returns are independent (but not necessarily stationary) from period to period, the use of a stationary myopic power-utility decision rule in each period is optimal. That is, the optimal policy only depends on next period's joint return distribution. (The single-period MV model simply assumes myopic policies are optimal.) The costs include increased complexity in estimation—the entire joint return distribution must be specified—and in computation.

Merton (1971, 1973) introduced a stochastically changing opportunity set that leads to hedging demand and non-myopic investment decisions. Recently, the (weak) evidence of in-sample return predictability based on information variables led to a resurgence of interest in the continuous-time model. A rich literature investigates hedging, the question of whether a long-horizon investor should allocate his wealth differently from a short-horizon investor, the effects of parameter and model uncertainty, the effects of transactions costs and the effects of conditioning on asset pricing models when returns are predictable, see (Kandel & Stambaugh, 1996; Kim & Omberg, 1996; Brennan, Schwartz & Lagnado, 1997; Balduzzi & Lynch, 1999; Brandt, 1999; Campbell & Viceira, 1999; Barberis, 2000; Pastor, 2000; Pastor & Stambaugh, 2000; Lynch, 2001; Lynch & Balduzzi, 2001; Avramov, 2002; Brennan & Xia, 2002). Much of the computational analysis calibrates the importance of hedging demand in simulated settings where the stochastic process is consistent with a regression of returns on informational variables. And, with the exception of Brennan, Schwartz and Lagnado (1997), there is little in the way of out-of-sample results. The approach provides a great deal of insight into the multiperiod investment problem, but does not come without the added costs of predicting returns beyond the current period and still further computational complexity. In many instances, specific distributional assumptions are needed to make the model tractable, which causes problems in computing expected utility. See, for example, Barberis (2000) and Kandel and Stambaugh (1996), who constrain investors from short selling and from buying on margin to insure that the expected utility problem has a feasible solution.

Clearly, the original in-sample evidence of return predictability generated from information variables calls into question the assumption that returns are inter-temporally independent and explains the resurgence of interest in hedging and the continuous-time model. But, Bossaerts and Hillion (1999) and Goyal and Welch (2003) find no evidence of out-of-sample predictability based on information variables and in this paper, little, or no, evidence of out-of-sample predictability for *any* of the mean estimators is found. In light of this evidence and questions about in-sample predictability involving regressions of returns on information variables, this study employs inter-temporal independence and the myopic behavior of the discrete-time power-utility model as working assumptions in this paper.

The paper proceeds as follows. The next section outlines the basic multiperiod investment model and the method employed to make it operational. The data, the estimators of the means, and the statistical measures employed to evaluate the investment performance of the portfolios generated from five mean estimators are described in the following three sections. The results based on statistical criteria, economic criteria, and statistical measures of investment performance, and the robustness of the results are discussed in the second to last section. The final section contains a summary and conclusions.

The Discrete-Time Power-Utility Model

The discrete-time model is the same as the one employed in Grauer and Hakansson (1986) and the reader is referred to that paper (specifically pages 288-291) for details. It is based on the pure reinvestment version of dynamic investment theory. In particular, if $U_n(w_n)$ is the *induced* utility of wealth w with n periods to go (to the horizon) and r is the single-period return on the portfolio, the important convergence result: $U_n(w_n) \rightarrow (1/\gamma)w^\gamma$ for some $\gamma < 1$, holds for a very broad class of terminal utility functions $U_o(w_o)$ when returns are independent (but non-stationary) from period to period. (See Hakansson, 1974; Mossin, 1968; Hakansson, 1971; Leland, 1972; Ross, 1974; Huberman and Ross, 1983). Convergence implies that the use of the stationary *myopic* decision rule: max $E(1/\gamma)(1 + \mathrm{r})^\gamma$, for some $\gamma < 1$, in each period is optimal.

At the beginning of each period t, the investor chooses a vector of portfolio weights

$\mathbf{x}_t$ on the basis of some member γ of the family of utility functions for returns r given by

$$\max_{\mathbf{x}_t} E\,\pi\left[\frac{1}{\gamma}(1 + m_t(\mathbf{x}_t))^{\gamma}\right] = \max_{\mathbf{x}_t} \sum_s \pi_{ts}\frac{1}{\gamma}(1 + (\mathbf{x}_t))^{\gamma} \tag{1}$$

subject to

$$x_{it} \geq 0, \text{ all } i,\ x_{Lt} \geq 0,\ x_{Bt} \leq 0, \tag{2}$$

$$\sum_i x_{it} + x_{Lt} + x_{Bt} = 1, \tag{3}$$

$$\sum_i m_{it}x_{it} \leq 1, \tag{4}$$

$$1 + r_{ts}(\mathbf{x}_t) > 0, \text{ for all s}, \tag{5}$$

where:

$r_{ts}(\mathbf{x}_t) = \sum x_{it}r_{its} + x_{Lt}r_{Lt} + x_{Bt}r^d_{Bt}$ is the *(ex ante)* return on the portfolio in period t if state s occurs,

$\gamma \leq 1$ = a parameter that remains fixed over time,
x_{it} = the amount invested in risky asset category i in period t as a fraction of own capital,
x_{Lt} = the amount lent in period t as a fraction of own capital,
x_{Bt} = the amount borrowed in period t as a fraction of own capital,
$\mathbf{x}_t$ = $(x_{1t},\ldots,x_{nt},x_{Kt},x_{Bt})'$,
r_{it} = the anticipated total return (dividend yield plus capital gains or losses) on asset category i in period t,
r_{Lt} = the return on the riskfree asset in period t,
r^d_{Bt} = the interest rate on borrowing at the time of the decision at the beginning of period t,
m_{it} = the initial margin requirement for asset category i in period t expressed as a fraction,
π_{ts} = the probability of state s at the end of period t, in which case the random return r_{it} will assume the value r_{its}.

Constraint (2) rules out short sales and ensures that lending (borrowing) is a positive (negative) fraction of capital. Constraint (3) is the budget constraint. Constraint (4) serves to limit borrowing (when desired) to the maximum permissible under the margin requirements that apply to the various asset categories. Constraint (5) rules out any *ex ante* probability of bankruptcy. The solvency constraint is not binding for the power functions, with $\gamma < 0$ and discrete probability distributions with a finite number of outcomes, because the marginal utility of zero wealth is infinite.

Nonetheless, it is convenient to explicitly consider equation (5) so that the nonlinear programming algorithm used to solve the investment problem does not attempt to evaluate an infeasible policy as it searches for the optimum.

The inputs to the model are based on the "empirical probability assessment approach" (EPAA) with quarterly revisions. At the beginning of quarter t, the portfolio problem consisting of equations (1)-(5) for that quarter uses the following inputs: the (observable) riskfree return for quarter t, the (observable) call money rate +1% at the beginning of quarter t and the (observable) realized returns for the risky asset categories for the previous k quarters. Each joint realization in quarters $t-k$ through $t-1$ is given probability $1/k$ of occurring in quarter t. Thus, under the EPAA, estimates are obtained on a moving basis and used in raw form without adjustment of any kind. On the other hand, since the whole joint distribution is specified and used, there is no information loss; all moments and correlations are implicitly taken into account. It may be noted that the empirical distribution of the past k periods is optimal if the investor has no information about the form and parameters of the true distribution, but believes that this distribution went into effect k periods ago.

With these inputs in place, the portfolio weights $\mathbf{x}_t$ for the various asset categories and the proportion of assets borrowed are calculated by solving equations (1)-(5) via nonlinear programming methods, (see Best (1975)). At the end of quarter t, the realized returns on the risky assets are observed, along with the realized borrowing rate r^r_{Bt} (which is calculated as a monthly average and may differ from the decision borrowing rate r^d_{Bt}). Then, using the weights selected at the beginning of the quarter, the realized return on the portfolio chosen for quarter t is recorded. The cycle is repeated in all subsequent quarters. Note that if $k = 32$ under quarterly revision, then the first quarter for which a portfolio can be selected is b+32, where b is the first quarter for which data is available.

All reported returns are gross of transaction costs and taxes and assume that the investor in question had no influence on prices. There are several reasons for this approach. First, as in previous studies, we wish to keep the complications to a minimum. Second, the return series used as inputs and for comparisons also exclude transaction costs (for reinvestment of interest and dividends) and taxes. Third, many investors are tax-exempt and various techniques are available for keeping transaction costs low. Finally, since the proper treatment of these items is nontrivial, they are better left to a later study.

Data

The data used to estimate the probabilities of the next period's returns on risky assets and to calculate each period's realized returns on risky assets come from several sources. The returns for Standard and Poor's 500 Index come from the Ibbotson Associates database. The returns for the value-weighted industry groups are constructed from the returns on individual New York Stock Exchange firms contained in the Center for Research in Security Prices' (CRSP) monthly returns database. The firms are combined into twelve industry groups on the basis of the first two digits of the firms' SIC codes. (Grauer, Hakansson & Shen (1990) contains a detailed description of the industry data.) The riskfree asset is assumed to be 90-day U.S. Treasury bills maturing at the end of the quarter. *The Survey of Current Business* and the *Wall Street Journal* are the sources. The borrowing rate is assumed to be the call money rate +1% for decision purposes (but not for rate of return calculations). The applicable beginning of period decision rate, r_{Bt}^{d}, is viewed as persisting throughout the period and thus as riskfree. For 1934-76, the call money rates are obtained from the *Survey of Current Business*. For later periods, the *Wall Street Journal* is the source. Finally, margin requirements for stocks are obtained from the *Federal Reserve Bulletin*. There is no practical way to take maintenance margins into account in our programs. In any case, it is evident from the results that they would come into play only for the more risk-tolerant strategies and for them only occasionally and that the net effect would be relatively neutral.

Estimators of the Means

Under the historic approach means are not used directly but are implicitly computed from the realized returns in the estimation period. The *n*-vector of historic means at the beginning of period *t* is

$$\boldsymbol{\mu}_{Ht} = (\overline{r}_{1t},\ldots,\overline{r}_{nt})', \tag{6}$$

where $\overline{r}_{it} = \dfrac{1}{k}\sum_{\tau=t-k}^{t-1} r_{i\tau}$. This EPAA approach implicitly estimates the means one at a time, relying exclusively on information contained in each of the time series.

Stein's (1955) suggestion that the efficiency of the estimate of the means could be improved by pooling the information across series leads to a number of so-called "shrinkage" estimators that shrink the historical means to some grand mean. A classic example is the James-Stein estimator (Efron & Morris 1973, 1975, 1977). It was first employed in the portfolio selection literature by Jobson, Korkie and Ratti (1979). However, this paper focuses on a Bayes-Stein (BS) estimator (Jorion, 1985, 1986, 1991)

$$\boldsymbol{\mu}_{BSt} = (1 - w_t)\, \boldsymbol{\mu}_{Ht} + w_t\overline{r}_{Gt}\boldsymbol{\iota}, \tag{7}$$

where $w_t = \lambda_t / (\lambda_t + k)$ is the shrinking factor, $\lambda_t = (n + 2)/((\boldsymbol{\mu}_{Ht} - \overline{r}_{Gt}\boldsymbol{\iota})\text{'}S_t^{-1}(\boldsymbol{\mu}_{Ht} - \overline{r}_{Gt}\boldsymbol{\iota}))$, *n* is the number of risky assets, S_t is the sample covariance matrix calculated from the *k* periods in the estimation period, $\overline{r}_{Gt} = \boldsymbol{\iota}\text{'}S_t^{-1}\boldsymbol{\mu}_{Ht} / (\iota\text{'}S_t^{-1}\boldsymbol{\iota})$ is the grand mean and $\boldsymbol{\iota}$ is a vector of ones. The λ does not contain an adjustment for degrees of freedom in estimating the covariance matrix as in Jorion (1985, 1986, 1991). We chose to model the problem this way to allow for the possibility of combining a Stein estimator with a set of non-equal probabilities for the states of nature used to estimate the joint distribution of security returns. In this case, the grand mean is the mean of the global minimum-variance portfolio generated from the historical data. Having calculated the Bayes-Stein and historic means for asset *i*, we add the difference $(\overline{r}_{BSit} - \overline{r}_{it})$, where $\overline{r}_{BSit}$ and $\overline{r}_{it}$ are the Bayes-Stein and historic means for asset *i* at time *t*, to each actual return on asset i in the estimation period. That is, in each estimation period, we replace the raw return series with the adjusted return series $r_{i\tau}^{A} = r_{it} + (\bar{r}_{BSit} - \bar{r}_{it})$, for all i and τ. No adjustment is made to the EPAA variance-covariance structure or to the other moments. Thus, the mean vector of the adjusted series is equal to the Bayes-Stein means of the original series; all other moments are unchanged. The same procedure is followed for the CAPM and dividend-yield riskfree-rate estimators discussed below.

A third estimator of the means is based on the CAPM (Sharpe, 1964; Lintner, 1965). The CAPM estimator is

$$\boldsymbol{\mu}_{CAPMt} = r_{Lt}\boldsymbol{\iota} + (\bar{r}_{mt} - \bar{r}_{Lt})\hat{\boldsymbol{\beta}}_t, \tag{8}$$

where $\bar{r}_{mt} = \frac{1}{k}\sum_{\tau=t-k}^{t-1} r_{m\tau}$, $\bar{r}_{Lt} = \frac{1}{k}\sum_{\tau=t-k}^{t-1} r_{L\tau}$, and $\bar{r}_{mt} - \bar{r}_{Lt}$ is an estimate of the expected excess return on the "market" portfolio and $\hat{\boldsymbol{\beta}}_t$ is the vector of estimated betas or systematic risk coefficients. At each time t, $\hat{\boldsymbol{\beta}}_t$ is estimated from the market model regressions

$$r_{i\tau} = \alpha_{it} + \beta_{it} r_{m\tau} + e_{i\tau}, \text{ for all } i \text{ and } \tau, \tag{9}$$

in the t–k to t–l estimation period, where the CRSP value-weighted index is employed as the proxy for the market portfolio. This method of estimating CAPM means, employed by Jorion (1991) and Grauer and Hakansson (1995), assumes the excess return on the market is constant over the estimation period. Alternatively, the ratio of the excess return on the market to the market's standard deviation or variance might be assumed to be constant as in Merton (1980) and Best and Grauer (1985).

The next two estimators use dividend yields and riskfree interest rates to forecast the means. To construct the dividend-yield riskfree-rate estimators, the following regression is run at each time *t:*

$$r_{i\tau} = a_{0i} + a_{1i} dy_{\tau-1} + a_{2i} r_{L\tau} + e_{i\tau}, \text{ for all } i \text{ and } \tau, \tag{10}$$

in the *t–k* to *t–1* estimation period, where the i subscript denotes an industry, $dy_{\tau-1}$ is the annual dividend yield on the CRSP value-weighted index lagged one month so that it is observable at the beginning of quarter t, and r_{Lt} is the (observable) beginning-of-quarter Treasury bill rate. Both independent variables are "de-meaned." Hence, a_{0i} is the historic average rate of return on asset (industry) i. The traditional one-period ahead forecast of the mean of industry i is

$$\bar{r}_{DRit} = \hat{a}_{0i} + a_{1i} dy_{t-1} + \hat{a}_{2i} r_{Lt}, \tag{11}$$

where $\hat{a}_{0i}$, $\hat{a}_{1i}$ and $\hat{a}_{2i}$ are the estimated coefficients and dy_{t-1} and r_{Lt} are observable at the beginning of period t+1. That is, the quarterly variable dy_{t-1} is lagged one month and there is no need to lag r_{Lt} as it is observable at the beginning of the quarter. The vector of dividend-yield riskfree-rate (DR) estimators is

$$\boldsymbol{\mu}_{DRt} = (\bar{r}_{DR1t},\ldots,\bar{r}_{DRnt})'. \tag{12}$$

However, this forecast is extremely variable. Therefore, we adopt the Bayesian framework advocated by Black and Litterman (1992) and shrink these mean forecasts to CAPM means. The Black-Litterman framework attempts to overcome three problems: the difficulty in estimating means; the extreme sensitivity of the solutions to slight perturbations in equilibrium means documented by Best and Grauer (1991, 1992) and Green and Hollifield (1992) and need for a set of means that would clear the market in an equilibrium setting. When the DR and CAPM means are assumed to be equally likely, the vector of dividend-yield riskfree-rate - CAPM (DRCAPM) mean estimators is

$$\boldsymbol{\mu}_{DRCAPMt} = (\boldsymbol{\mu}_{DRt} + \boldsymbol{\mu}_{CAPMt})/2. \tag{13}$$

The CAPM means are estimated from equation (8). Black and Litterman (1992) estimate the means in a slightly different way. They estimate what they call equilibrium (and Best & Grauer (1985) call $(\Sigma,\mathbf{x})$ - compatible) means from the equation $\boldsymbol{\mu} = r_L\boldsymbol{\iota} + \delta\Sigma\mathbf{x}$ where Σ is the covariance matrix of asset returns and $\mathbf{x}$ is a vector of portfolio weights. When these means and Σ are inputs to a MV problem, subject only to a budget constraint, $\mathbf{x}$ is the optimal solution.

Statistical Measures of Investment Performance

In the results section we will see that the compound return – standard deviation plots and cumulative wealth values provide convincing evidence that: (1) all the mean estimators provide economic value when combined with the discrete-time power-utility model and (2) the two dividend-yield riskfree-rate estimators perform better than the traditional CAPM estimator. Although the figures and wealth values get to the heart of the matter, they do not give us a sense of how much of the difference can be attributed to randomness. In order to shed light on this issue the results from a number of commonly accepted statistical measures of performance are reported. Unfortunately, none is without problems. First, the industry-rotation strategies examined here are neither the pure selectivity strategies implicit in Jensen's (1968) test, nor the pure market-timing strategies embodied in Treynor and Mazuy's (1966) and Henriksson and Merton's (1981) tests of market timing. Second, Roll (1978) argues that Jensen's test is ambiguous because the choice of the benchmark (market) portfolio affects both systematic risk (beta) and abnormal return (alpha), also see (Dybvig & Ross, 1985; Grauer, 1991; Green, 1986). Third, expected returns and risk measures may vary with economic conditions.

In light of these problems this study employs an eclectic mix of performance measures that include conditional and unconditional versions of the Jensen, Henriksson-Merton and Treynor-Mazuy tests as well as the portfolio change measure, see Grinblatt and Titman (1993) which gauges performance without reference to a proxy for the market portfolio. For each of the measures the null hypothesis is that there is no superior investment performance. The alternative hypothesis is a one-tailed test that there is superior performance. The Jensen, Henriksson-Merton and Treynor-Mazuy regressions are corrected for heteroskedasticity using White's (1980) correction.

The unconditional Jensen (1968) test is based on the regression

$$R_{pt} = \alpha_p + \beta_p R_{mt} + u_{pt}, \tag{14}$$

where $R_{pt} = r_{pt} - r_{Lt}$ is the excess return on portfolio p over the Treasury bill rate, $R_{mt} = r_{mt} - r_{Lt}$ is the excess return on the CRSP value-weighted index, α_p is the unconditional measure of performance and β_p is the unconditional measure of risk.

However, expected returns and betas almost certainly change over time. Therefore, Ferson and Schadt (1996) and Ferson and Warther (1996) among others, building on the earlier work of Shanken (1990) advocate conditional performance measures. This study follows their suggestion that a portfolio's risk is related to dividend yields and short-term Treasury yields postulating that

$$\beta_p = b_{0p} + b_{1p} dy_{t-1} + b_{2p} r_{Lt}, \tag{15}$$

where dy_{t-1} is the CRSP value-weighted index annual dividend yield at the beginning of period t and r_{Lt} is the (observable) beginning-of-quarter Treasury bill rate, both measured as deviations from their estimation-period means. Substituting equation (15) into equation (14), yields the conditional Jensen test

$$R_{pt} = \alpha_{cp} + b_{0p} R_{mt} + b_{1p}[dy_{t-1} R_{mt}] + b_{2p}[r_{Lt} R_{mt}] + e_{pt}, \tag{16}$$

where α_p is the conditional measure of performance, and b_{1p} and b_{2p} measure how the conditional beta varies with dividend yields and Treasury bill rates.

The unconditional regression specification for the Treynor and Mazuy (1966) test is

$$R_{pt} = \alpha_{cp} + \beta_p R_{mt} + \gamma_p R_{mt}^2 + u_{pt}, \tag{17}$$

where α_p is the measure of selectivity, β_p is the unconditional beta and γ_p is the market-timing coefficient. Substituting for β_p, the conditional regression specification is

$$R_{pt} = \alpha_{cp} + b_{0p} R_{mt} + b_{1p}[dy_{t-1} R_{mt}] + b_{2p}[r_{Lt} R_{mt}] + e_{pt}, \tag{18}$$

where α_{cp}, b_{1p}, b_{2p} and γ_p are defined above.

The unconditional Henriksson and Merton (1981) test is given by

$$R_{pt} = \alpha_p + b_{dp} R_{mt} + \gamma_p \max(0, R_{mt}) + u_{pt}, \tag{19}$$

where α_p is the measure of selectivity, β_p is the down-market beta, γ_p is the market-timing coefficient, in this case the difference between the up- and down-market beta, and $\max(0,R_{mt})$ is the payoff on a call option on the market with an exercise price equal to the riskfree rate of interest. Following Ferson and Schadt (1996) the conditional Henriksson-Merton test is

$$R_{pt} = \alpha_{cp} + b_{dp}R_{mt} + b_{1p}[dy_{t-1}R_{mt}] + b_{2p}[r_{Lt}R_{mt}] + \gamma_p R^*_{mt} + b^*_{1p}[dy_{t-1}R^*_{mt}] + b^*_{2p}[r_{Lt}R^*_{mt}] + e_{pt}, \quad (20)$$

where R^*_{mt} is the product of the excess return on the CRSP value-weighted index and an indicator dummy for positive values of the difference between the excess return on the index and the conditional mean of the excess return. (The conditional mean is estimated by a linear regression of the excess return of the CRSP value-weighted index on dy_{t-1} and r_{Lt}.) The most important coefficients is γ_p, the market-timing coefficient, which in this case is the difference between the up- and down-market conditional betas.

In contrast to most other performance measures, Grinblatt and Titman's (1993) portfolio change measure employs portfolio holdings as well as rates of return and does not require an external benchmark (market) portfolio. In order to motivate the portfolio change measure, assume that uninformed investors perceive that the vector of expected returns is constant, while informed investors can predict whether expected returns vary over time. Informed investors can profit from changing expected returns by increasing (decreasing) their holdings of assets whose expected returns have increased (decreased). The holding of an asset that increases with an increase in its conditional expected rate of return will exhibit a positive unconditional covariance with the asset's returns. The portfolio change measure is constructed from an aggregation of these covariances. For evaluation purposes, let

$$\text{PCM}_t = \sum_i r_{it}\,(x_{it} - x_{i,t-j}),$$

where r_{it} is the quarterly rate of return on asset i time t, x_{it} and $x_{i,t-j}$ are the holdings of asset i at time t and time $t–j$, respectively. This expression provides an estimate of the covariance between returns and weights at a point in time. Alternatively, it may be viewed as the return on a zero-weight portfolio. The portfolio change measure is an average of the PCM_t's

$$\overline{\text{PCM}} = \sum_t \sum_i [r_{it}(x_{it} - x_{i,t-j})/\text{T}], \quad (21)$$

where T is the number of time-series observations. The portfolio change measure test itself is a simple t-test based on the time series of zero-weight portfolio returns, i.e.,

$$t = (\overline{\text{PCM}} / \sigma(\text{PCM}))\sqrt{T}, \quad (22)$$

where σ(PCM) is the standard deviation of the time series of PCM_t's.

In their empirical analysis of mutual fund performance, Grinblatt and Titman work with two values of j that represent one- and four-quarter lags. They report that four-quarter lag portfolio change measures are statistically significant. Hence, this paper will focus on the four-quarter lag portfolio change measure. The portfolio change measure is particularly apropos in the present study because the portfolio weights are chosen according to a pre-specified set of rules over the same quarterly time interval as performance is measured. Thus, one does not have to worry about possible gaming or window-dressing problems that face researchers trying to gauge the performance of mutual funds.

Results

Results Based on Statistical Criteria

Table 1 depicts the labels used in the figures and tables. The out-of-sample statistical forecasting results are reported in Table 2. The findings confirm Bossaerts and Hillion's (1999) and Goyal and Welch's (2003) results that there is little, or no, out-of-sample short-horizon forecasting ability with information variables. There is little to distinguish between any of the forecasts and even some of the small differences are counterintuitive. The total mean square error of the DR means is the largest even though the results in the next section show that the performance of the portfolios generated from them dominates the performance of all but the DRCAPM portfolios. The average out-of-sample R-squares are below 0.01 except for the CAPM estimator. Yet, when this estimator is combined with the power-utility model the resulting portfolios exhibit the worst economic performance. It is also somewhat surprising that the Bayes-Stein estimator, noted for minimizing the total mean square error, exhibits the smallest average out-of-sample R-square.

Table 1: *Definitions of the Labels in Tables and Figures*

	Benchmarks
RL	Riskfree lending at the three month Treasury bill rate
VW	Market (value-weighted CRSP index)
V5	50% in VW, 50% in lending
V15	150% in VW, 50% in borrowing at the call money rate plus 1%
V20	200% in VW, 100% in borrowing at the call money rate plus 1%
	Estimators of the Menu
Historic	Historic means
Bayes-Stein	Bayes-Stein means
CAPM	CAPM means
DR	Dividend – yield riskfree – rate means
DRCAPM	Dividend – yield riskfree – rate means shrunk to CAPM means

There are 264 quarterly forecasts of the means for each industry in the 1934-1999 period. At a point in time, each estimator bases its forecast on data from the previous 32-quarters. R-square values are formed by squaring the correlation coefficient between the time series of 264 mean forecasts and the 264 realized return. Mean squared errors are in units of percent squared per quarter. See Table 1 for definitions of the labels.

Table 2: *Out-of-Sample R-squares and Mean Squared Errors for Five Estimators of Industry Returns*

	Historic		Bayes-Stein		CAPM		DR		DRCAPM	
	R^2	MSE	R^2	MSE	R^2	MSE	R^2	MSE	R^2	MSE
Petroleum	0.00	92	0.00	92	0.00	92	0.00	99	0.00	91
Finance & Real Estate	0.01	102	0.01	102	0.02	103	0.00	121	0.00	104
Consumer Durables	0.00	124	0.00	124	0.02	125	0.02	131	0.01	120
Basic Industries	0.00	85	0.00	85	0.01	85	0.01	93	0.00	84
Food & Tobacco	0.00	65	0.00	65	0.00	65	0.01	73	0.01	65
Construction	0.00	142	0.00	139	0.02	142	0.02	149	0.01	137
Capital Goods	0.01	98	0.00	98	0.01	98	0.01	107	0.01	97
Transportation	0.02	142	0.00	140	0.01	142	0.00	163	0.00	142
Utilities	0.00	52	0.00	53	0.01	53	0.00	64	0.00	55
Textiles & Trade	0.00	112	0.00	111	0.00	111	0.01	125	0.01	112
Services	0.00	214	0.00	213	0.00	211	0.01	237	0.01	212
Leisure	0.03	179	0.00	176	0.03	178	0.01	201	0.00	177
Average R^2	0.01		0.00		0.01		0.01		0.00	
Sum MSE		1407		1398		1405		1563		1397

Results Based on Economic Criteria

Figure 1 plots the annual compound return (obtained by compounding the quarterly realized returns) and standard deviation of the realized returns for five sets of ten power-utility strategies, based on γs in equation (1) ranging from -50 (extremely risk averse) to 1 (risk-neutral), for the 66-year period from 1934-99. (For consistency with the compound return the standard deviation is based on the log of one plus the rate of return. This quantity is very similar to the standard deviation of the rate of return for levels less than 25 percent.) Portfolios are chosen each quarter, employing a 32-quarter estimation period, from an investment universe that includes the twelve value-weighted U.S. industry indices, lending and borrowing. The first set of strategies (black circles) shows the portfolio returns generated from historic means. The second set (open squares) displays the portfolio returns based on Bayes-Stein means. The third set (black triangles) depicts those based on CAPM means. The fourth set (black circles) presents the portfolio returns obtained employing the DR forecasts of the means. The fifth set (open circles) presents the portfolio returns obtained employing the DRCAPM forecasts. The figure also shows the benchmarks: RL, V5, VW, V15 and

V20 as black diamonds. Figure 2 plots the corresponding results for the 30-year sub-period from 1966-99. Finally, Figure 3 displays the results for the (inflationary) 16-year sub-period from 1966-81, a period which experienced a one-half percent per year *negative* realized risk premium on the value-weighted portfolio of risky assets.

The figures show three main results. The portfolios generated from the DR and DRCAPM estimators "outperform" portfolios that employ the historic, Bayes-Stein and CAPM estimators. The portfolios based on the CAPM estimator are "dominated" by the other four and by the benchmark portfolios. In the 1966-81 period, however, portfolios employing historic and Bayes-Stein estimators do not so obviously dominate portfolios based on the CAPM estimator.

Figure 1: *Annual compound return versus the standard deviation (of the log of one plus return) for five benchmarks and five sets of power-utility portfolios constructed from historic, Bayes-Stein, CAPM, dividend – yield riskfree – rate (DR), and dividend – yield riskfree – rate-CAPM (DRCAPM) estimators of the means in the 1934-1999 period*

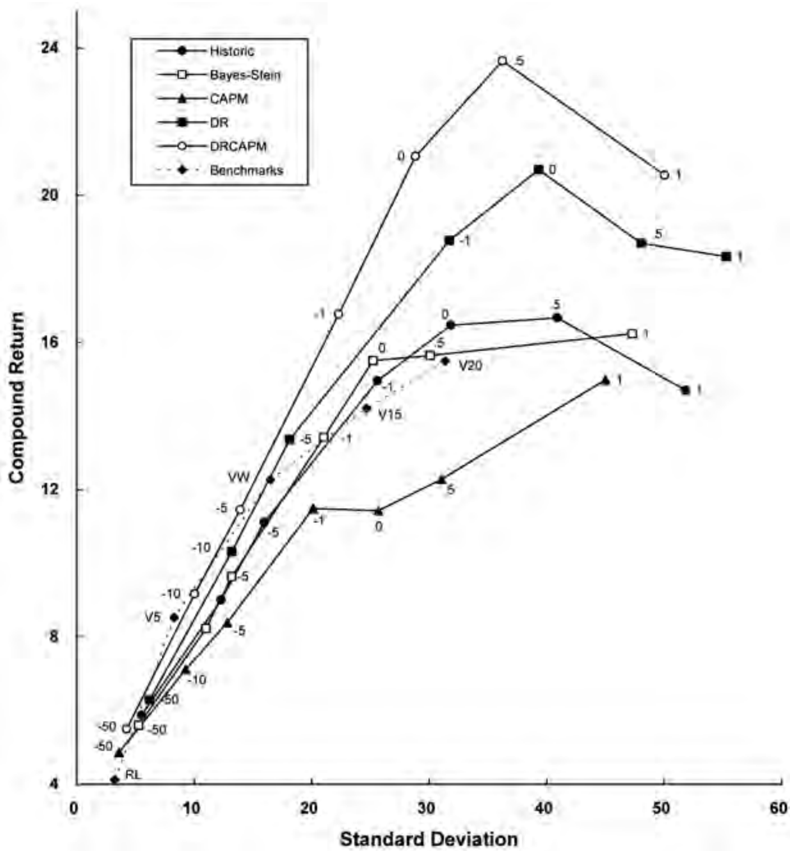


Figure 2: *Annual compound return versus the standard deviation (of the log of one plus return) for five benchmarks and five sets of power-utility portfolios constructed from historic, Bayes-Stein, CAPM, dividend – yield riskfree – rate (DR), and dividend – yield riskfree – rate-CAPM (DRCAPM) estimators of the means in the 1966-1999 period*

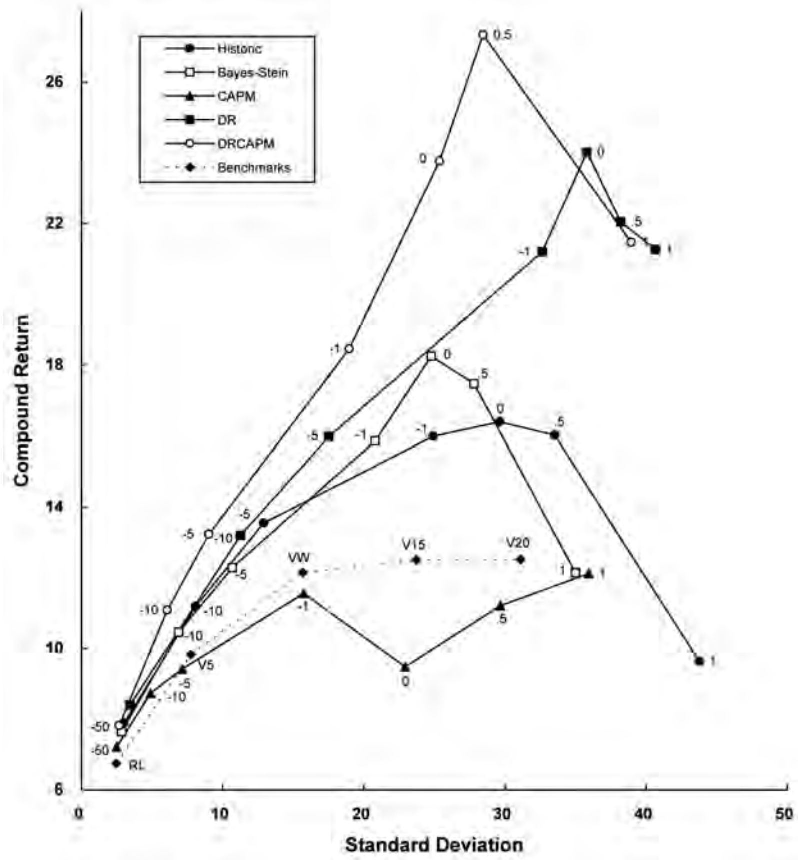
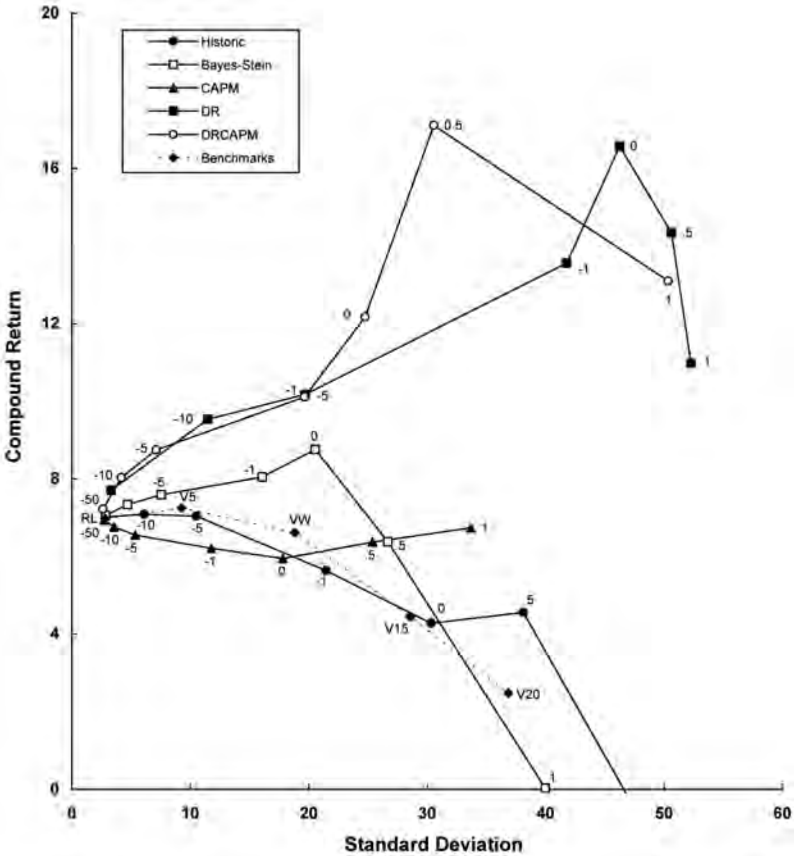


Figure 3: *Annual compound return versus the standard deviation (of the log of one plus return) for five benchmarks and five sets of power-utility portfolios constructed from historic, Bayes-Stein, CAPM, dividend – yield riskfree – rate (DR), and dividend – yield riskfree – rate-CAPM (DRCAPM) estimators of the means in the 1966-1981 period*



Results Based on Statistical Measures of Investment Performance

Table 3 summarizes the results obtained from the conditional and unconditional Jensen tests and from the Grinblatt-Titman portfolio change measure tests in the same three time-periods examined in the figures. For the most part, the results of the tests are consistent with conclusions drawn from the figures and the cumulative wealth values. The alphas and portfolio change measures of the portfolios generated from the DR and DRCAPM estimators are much larger and for the most part more statistically significant than those of the other three mean estimators, particularly in the 1934-99 and 1966-99 periods. The conditional Jensen test uniformly ranks the performance of the historic, Bayes-Stein and CAPM portfolios higher than the unconditional test, which is consistent with Ferson and Schadt's (1996) and Ferson and Warther's (1996) mutual fund results. But this pattern is less obvious for the DR and DRCAPM portfolios. Furthermore, in the 1966-81 period the DR and DRCAPM portfolios perform poorly according to the portfolio change measure test, which is just the opposite of what is shown in Figure 3.

The alphas and portfolio change measures are averages calculated over ten power portfolios. Both are measured in units of percent per quarter. the portfolio change measures are based on a four-quarter lags. For both measures the null hypothesis is that there is no superior investment performance. The alternative hypothesis is a one-tailed test that there is superior performance. The Jensen regressions are corrected for heteroskedasticity using White's (1980) correction Number ≤ 0.05 refers to the number of portfolios out of ten whose coefficients are statistically significant at the 5% level. See Table 1 for label definitions.

Table 3: *Unconditional and Conditional Jensen Alphas and Grinblatt-Titman Portfolio Change Measures for Ten Power Portfolios Estimated from Five Sets of Means*

	Unconditional Jensen			Conditional Jensen			Grinblatt-Titman		
	Alpha	Number Negative	Number ≤ 0.05	Alpha	Number Negative	Number ≤ 0.05	PCM	Number Negative	Number ≤ 0.05
Panel A: Historic Means									
1934-1999	0.50	0	3	0.48	1	8	0.82	0	10
1966-1999	0.57	1	5	0.76	0	7	1.00	0	10
1966-1981	-0.20	4	0	0.62	0	0	0.72	0	6
Panel B: Bayes-Stein Means									
1934-1999	0.54	0	4	0.54	1	9	0.75	0	9
1966-1999	0.70	0	6	0.81	0	9	1.03	0	10
1966-1981	0.06	1	0	0.62	0	3	0.96	0	9
Panel C: CAPM Means									
1934-1999	0.17	0	0	0.13	1	0	0.05	8	1
1966-1999	0.11	1	0	0.20	1	0	0.21	1	0
1966-1981	0.06	8	0	0.86	0	3	0.26	1	0
Panel D: DR Means									
1934-1999	0.80	0	4	0.73	0	6	1.55	0	10
1966-1999	1.40	0	8	0.90	0	1	2.16	0	9
1966-1981	1.75	0	4	0.53	0	0	1.18	4	1
Panel E: DRCAPM Means									
1934-1999	0.90	0	9	0.88	0	9	1.00	0	10
1966-1999	1.28	0	10	1.16	0	9	1.33	0	10
1966-1981	1.07	0	3	0.89	0	2	0.85	0	0

Table 4 contains the results of the unconditional and conditional Henriksson-Merton and Treynor-Mazuy market-timing tests. There is little or no evidence of market-timing ability according to both unconditional tests. In the 1934-99 period the majority of the historic, Bayes-Stein and CAPM portfolios show negative timing ability. In the 1966-99 period there is more evidence of market-timing ability especially according to the conditional tests. But the tests indicate more market-timing ability for the historic, Bayes-Stein and CAPM portfolios than for the DR and DRCAPM portfolios. In the 1966-81 period the results are anomalous. The conditional tests, especially the Treynor-Mazuy conditional test, show strong evidence of market-timing ability for the historic, Bayes-Stein and CAPM estimators, and no evidence of market-timing ability for any of the dividend-yield riskfree-rate estimators, which seems to be at variance with the results reported in Figure 3 and Table 3.

The reported gammas (e.g. the timing coefficients) are averages calculated over ten power portfolios. The null hypothesis is that there is no timing ability. The alternative hypothesis is a one-tailed test that there is positive timing ability. The regressions are corrected for heteroskedasticity using White's (1980) correction. Number ≤ 0.05 refers to the number of portfolios out of ten whose timing coefficients are statistically significant at the 5% level. See Table 1 for label definitions.

Table 4: *Unconditional and Conditional Henriksson-Merton and Treynor-Mazuy Timing Coefficients for Ten Power Portfolios Estimated from Five Sets of Means*

	Henrikson-Merton						Treynor-Mazuy					
	Unconditional			Conditional			Unconditional			Conditional		
	Gamma	Number Negative	Number ≤ 0.05	Gamma	Number Negative	Number ≤ 0.05	Gamma	Number Negative	Number ≤ 0.05	Gamma	Number Negative	Number ≤ 0.05
Panel A: Historic Means												
1934-1999	0.04	6	0	0.09	1	0	-0.001	9	0	0.003	1	0
1966-1999	0.18	1	0	0.29	1	8	0.005	1	0	0.008	0	7
1966-1981	-0.10	3	0	0.17	1	6	-0.002	2	0	0.007	0	9
Panel B: Bayes-Stein Means												
1934-1999	0.01	8	0	0.08	1	0	-0.002	9	0	0.003	0	0
1966-1999	0.19	1	0	0.63	0	7	0.005	1	0	0.007	0	2
1966-1981	-0.06	2	0	0.14	1	3	0.000	1	0	0.005	1	9
Panel C: CAPM Means												
1934-1999	0.01	8	0	0.05	0	0	-0.002	9	0	0.001	1	0
1966-1999	0.08	1	0	0.15	0	0	0.002	1	0	0.004	0	0
1966-1981	-0.02	3	0	0.27	0	9	0.000	2	0	0.010	0	10
Panel D: DR Means												
1934-1999	0.33	0	2	0.20	0	1	0.009	0	1	0.005	1	6
1966-1999	0.43	0	5	0.19	0	3	0.008	0	4	0.003	3	1
1966-1981	0.38	0	0	0.13	0	0	0.004	2	0	-0.011	10	0
Panel E: DRCAPM Means												
1934-1999	0.15	0	1	0.18	0	1	0.001	8	1	0.003	0	0
1966-1999	0.28	0	4	0.24	0	6	0.006	0	0	0.005	0	6
1966-1981	0.08	2	0	0.20	0	1	-0.001	9	0	-0.005	10	0

Robustness of the Results

Four additional estimators of the means were examined. First, because of the variability in the dividend-yield riskfree-rate means, these forecasts were shrunk to historic means. The portfolios based on the dividend-yield riskfree-rate – historic (DRH) means performed better than those based solely on dividend-yield riskfree-rate means, but not as well as those based on dividend-yield riskfree-rate means shrunk to CAPM (e.g., DRCAPM) means. Second, in order to more fully investigate the Henriksson-Merton and Treynor-Mazuy market-timing test results, which indicate that there is little or no market timing ability in portfolios generated from different mean forecasts, two timing means were developed. The mean on the market (rather than the means on industries) was forecast using the dividend-yield riskfree-rate – historic mean method. These forecasts were called μ_{DRMKT} means. Then, a second set of CAPM means was developed using the μ_{DRMKT} means to set the slope of the SML, i.e., $\mu_{DRMCAPMt} = r_{Lt}\iota + (\mu_{DRMKTt} - r_{Lt})\hat{\beta}$. The DRMKT portfolios (that consisted of the market and either borrowing or lending only) and the industry-rotation DRMCAPM portfolios timed the market according to the Henriksson-Merton and Treynor-Mazuy market-timing tests. The industry-rotation portfolios based on the DRMCAPM means accumulated more wealth than the pure market timing DRM portfolios, but not as much as the industry rotation strategies based on the DRH and DRCAPM means. Finally, the robustness of the 32-quarter moving window approach was examined by comparing it with an expanding window approach that employed an all-of-history dividend-yield riskfree-rate – historic mean forecast of industry means. The portfolios based on this expanding-window method of forecasting the means performed much worse than portfolios based on the other dividend-yield riskfree-rate mean forecasts.

On a different dimension, a one-quarter portfolio change measure test was conducted. Like Grinblatt-Titman's (1993) results for mutual funds there was no statistically significant abnormal performance according to this measure. Moreover, the DR and DRCAPM portfolios performed abysmally in all three periods according to the one-quarter lag portfolio change measure, which is at complete odds with the other evidence presented in the paper.

Summary and Concluding Comments

This paper evaluates historic, Bayes-Stein, CAPM and two dividend-yield riskfree-rate estimators of asset means employing statistical and economic criteria in an industry-rotation setting using quarterly data. None of the estimators exhibit much in the way of out-of-sample predictive ability when judged by statistical criteria. The average out-of-sample R-squares are below 0.01 for all but the CAPM estimator. Moreover, there is little difference in the mean square errors of the five industry-rotation mean estimators. Yet, when the mean estimators are combined with a discrete-time power-utility portfolio selection model, the resulting portfolios earn economically significant returns. Even so, judged in terms of the compound return, standard deviation plots in Figures 1-3, or in terms of accumulated wealth, some of the resulting portfolios perform appreciably better than others. Specifically, the two dividend-yield riskfree-rate portfolios perform by far the best and the traditional CAPM portfolios perform the worst. For the most part, unconditional and conditional Jensen and Grinblatt-Titman tests support the compound return - standard deviation and wealth rankings, especially in the 1934-99 and 1966-99 periods. The unconditional and conditional Treynor-Mazuy and Henriksson-Merton market-timing tests indicate that the results did not arise from any market-timing ability.

So, are returns predictable? Clearly, the answer depends on the length of the decision horizon examined and metric chosen. With quarterly returns, the out-of-sample statistical answer is a clear "no"—and the economic answer is a resounding "yes." In this case, one can argue that the economic answer is compelling. You can't spend a slope coefficient, a t-statistic, an R-square, or a mean square error. But, four of the more risk-tolerant power-utility portfolios generated from DRCAPM means grew from one dollar to \$28,000, \$301,000, \$1,213,000 or \$228,000 over the 1934-99 period. These portfolios provided investors with real spending opportunities—especially when compared to an investment in the market that grew to only \$3,600, or an investment of 200% in the market financed by 100% borrowing that grew to \$9,700 over the same period. In this case, Leitch and Tanner (1991: p. 580) were correct in suggesting that: "… least-squares regression analysis may not be appropriate for many studies of economic behavior."

References

Ang, A. & Bekaert, G. (2003). Stock Return Predictability: Is it There? Manuscript, Columbia University.

Doron, A. (2002). Stock Return Predictability and Model Uncertainty. *Journal of Financial Economics, 64*: 423-458.

Doron, A. (2004). Stock Return Predictability and Asset Pricing Models. *Review of Financial Studies, 17*: 699-738.

Doron, A. & Tarun, C. (2006). Predicting Stock Returns. *Journal of Financial Economics, 82*: 387-415.

Balduzzi P. & Lynch, A.W. (1999). Transaction Costs and Predictability: Some Utility Cost Calculations. *Journal of Financial Economics, 52*: 47-78.

Barberis, N. (2000). Investing for the Long Run when Returns are Predictable. *Journal of Finance,* 55: 225-264.

Beller, K.R., Kling, J.L. & Levinson, M. (1998). Are Industry Stock Returns Predictable? *Financial Analysts Journal, 54*: 42-57.

Best, M.J. (1975). A Feasible Conjugate Direction Method to Solve Linearly Constrained Optimization Problems. *Journal of Optimization Theory and Applications, 16*: 25-38.

Best, M.J. & Grauer, R.R. (1985). Capital Asset Pricing Compatible with Observed Market Value Weights. *Journal of Finance, 40*: 85-103.

Best, M.J. & Grauer, R.R. (1991). On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies, 4*: 315-342.

Best, M.J. & Grauer, R.R. (1992). Positively Weighted Minimum-Variance Portfolios and the Structure of Asset Expected Returns. *Journal of Financial and Quantitative*

Analysis, 27: 513-537.

Black, F. & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal, 48*: 28-43.

Bossaerts, P. & Hillion, P. (1999). Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn? *Review of Financial Studies, 12*: 405-428.

Brandt, M.W. (1999). Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach. *Journal of Finance, 54*: 1609-1645.

Brennan, M.J., Schwartz, E.S. & Lagnado, R. (1997). Strategic Asset Allocation. *Journal of Economic Dynamics and Control, 21*: 1377-1403.

Brennan, M.J. & Xia, Y. (2002). Dynamic Asset Allocation under Inflation. *Journal of Finance, 57*: 1201-1238.

Campbell, J. & Viceira, L. (May 1999). Consumption and Portfolio Decisions When Expected Returns are Time Varying. *Quarterly Journal of Economics, 114*: 433-495.

Connor. G. (1997). Sensible Return Forecasting for Portfolio Management. *Financial Analysts Journal, 53*: 44-51.

Dybvig, P.H. & Ross, S.A. (1985). The Analytics of Performance Measurement Using a Security Market Line. *Journal of Finance, 40*: 401-416.

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. (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics, 25*: 23-49.

Ferson, W.E. & Schadt, R.W. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. *Journal of Finance, 51*: 25-462.

Ferson, W.E. & Warther, V.A. (1996). Evaluating Fund Performance in a Dynamic Market. *Financial Analysts Journal, 52*: 20-28.

Ferson, W.E. & Siegel, A.F. (2001). The Efficient Use of Conditioning Information in Portfolios. *Journal of Finance, 56*: 967-982.

Ferson, W.E., Sarkissian, S. & Simin, T. (2003). Spurious Regressions in Financial Economics? *Journal of Finance, 58*: 1393-1413.

Frost, P.A., & Savarino, J.E. (1988). For Better Performance: Constrain Portfolio Weights. *Journal of Portfolio Management,* 15: 29-34.

Goetzmann, W.N. & Jorion P. (1993). Testing the Predictive Power of Dividend Yields. *Journal of Finance, 48*: 663-680.

Grauer, R.R. (1991). Further Ambiguity When Performance Is Measured by the Security Market Line. *Financial Review, 26*: 569-585.

Grauer, R.R. (2007). Extreme Mean-Variance Solutions: Estimation Error versus Modeling Error? Manuscript, Simon Fraser University.

Grauer, R. R. & Hakansson, N.H. (1986). A Half-Century of Returns on Levered and Unlevered Portfolios of Stocks, Bonds, and Bills, with and without Small Stocks. *Journal of Business, 59*: 287-318.

Grauer, R.R. & Hakansson, N.H. (1987). Gains from International Diversification: 1968-85 Returns on Portfolios of Stocks and Bonds. *Journal of Finance, 42*: 721-739.

Grauer, R.R. & Hakansson, N.H. (1995). Stein and CAPM Estimators of the Means in Asset Allocation. *International Review of Financial Analysis, 4*: 35-66.

Grauer, R.R., Hakansson, N.H. & Shen, F.C. (1990). Industry Rotation in the U.S.

Stock Market: 1934-1986 Returns on Passive, Semi-passive, and Active Strategies. *Journal of Banking and Finance, 14:* 513-535.

Grauer, R.R. & Shen, F.C. (2000). Do Constraints Improve Portfolio Performance? *Journal of Banking and Finance, 24*: 1253-1274.

Green, R.C. (1986). Benchmark Portfolio Inefficiency and Deviations from the Security Market Line. *Journal of Finance, 41*: 295-312.

Green, R.C. & Hollifield, B. (1992). When Will Mean-Variance Efficient Portfolios Be Well-Diversified? *Journal of Finance, 47*: 1785-1810.

Grinblatt, M. & Titman, S. (1993). Performance Measurement Without Benchmarks: An Examination of Mutual Fund Returns. *Journal of Business, 66*: 47-68.

Hakansson, N.H. (1971). On Optimal Myopic Portfolio Policies, With and Without Serial Correlation of Yields. *Journal of Business, 44*: 324-34.

Hakansson, N.H. (1974). Convergence to Isoelastic Utility and Policy in Multiperiod Portfolio Choice. *Journal of Financial Economics, 1*: 201-24.

Hawawini, G. & Keim, D. (1995). On the Predictability of Common Stock Returns: World Wide Evidence. In Jarrow, R.A. Maksimovic, V. & Ziemba, W.T. editors: *Finance, Handbooks in Operations Research & Management Science,* Amsterdam: North Holland (9): 497-544.

Henriksson, R.D. & Merton, R.C. (1981). On Market Timing and Investment Performance II. Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business, 54:* 513-533.

Hodrick R.J. (1992). Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. *Review of Financial Studies, 5*: 357-386.

Huberman, G. & Ross, S.A. (1983). Portfolio Turnpike Theorems, Risk Aversion and Regularly Varying Utility Functions. *Econometrica, 51*: 1104-19.

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

Jobson, J.D., Korkie, B. & Ratti, V. (1979). Improved Estimation for Markowitz Portfolios using James-Stein Type Estimators. *Proceedings of the American Statistical Association,* Business & Economics Statistics Section, *41*: 279-284.

Jobson, J.D. & Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association, 75*: 544-554.

Jobson, J.D. & Korkie, B. (1981). Putting Markowitz Theory to Work. *Journal of Portfolio Management, 7*: 70-74.

Jorion, P. (1985). International Portfolio Diversification with Estimation Risk. *Journal of Business, 58*: 259-278.

Jorion, P. (1986). Bayes-Stein Estimation for Portfolio Analysis. *Journal of Financial and Quantitative Analysis 21*: 279-292.

Jorion, P. (1991). Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection. *Journal of Banking and Finance, 15*: 717-727.

Klemkosky, R.C. & Bharati, R. (1991). Time-Varying Expected Returns and Asset Allocation. *Journal of Portfolio Management, 21:* 80-88.

Leitch, G. & Tanner, J.E. (1991). Economic Forecast Evaluation: Profits versus the Conventional Error Measures. *American Economic Review, 81* (June 1991), 580-599.

Leland, H. (1972). On Turnpike Portfolios. *In Mathematical Methods in Investment and*

Finance. Shell, K. & Szego, G.P., editors., Amsterdam: North-Holland.

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*: 13-47.

Lynch, A.W. (2001). Portfolio Choice and Equity Characteristics: Characterizing the Hedging Demands Induced by Return Predictability. *Journal of Financial Economics, 62*: 67-130.

Lynch, A.W. & Balduzzi. P. (2000). Predictability and Transaction Costs: The Impact on Rebalancing Rules and Behavior. *Journal of Finance, 55:* 2285-2309.

Marquering, W. & Verbeek, M. (2001). The Economic Value of Predicting Stock Index Returns and Volatility. Manuscript, Erasmus University

Merton, R.C. (1971). Optimum Consumption and Portfolio Rules in a Continuous Time Model. *Journal of Economic Theory, 3*: 373-413.

Merton, R.C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, *41*: 867-888.

Merton, R.C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics, 8*: 323-361.

Michaud, R.O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal, 45*: 31-42

Mossin, J. (1968). Optimal Multiperiod Portfolio Policies. *Journal of Business, 41*: 215-229.

Pelaez, R.F. (1998). Economically Significant Stock Market Forecasts. *Financial Review, 33*: 65-76.

Pastor, L. (2000). Portfolio Selection and Asset Pricing Models. *Journal of Finance, 55*: 179-223.

Pastor, L. & Stambaugh, R.F. (2000). Comparing Asset Pricing Models: An Investment Perspective. *Journal of Financial Economics, 56*: 335-381.

Roll, R. (1978). Ambiguity When Performance Is Measured by the Securities Market Line. *Journal of Finance, 33*: 1051-1069.

Ross, S.A. (1974). Portfolio Turnpike Theorems for Constant Policies. *Journal of Financial Economics, 1*: 171-98.

Schwert, G.W. (2003). Anomalies and Market Efficiency. In Constantinides, G.M., Harris, M. & Stultz, R. editors: *Handbook in the Economics of Finance*, Amsterdam: Elsevier Science B.V.: 939-972.

Shanken, J. (1990). Intertemporal Asset Pricing: an Empirical Investigation. *Journal of Econometrics, 45*: 99-120.

Sharpe, W.F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, 19*: 425-442.

Solnik, B. (1993). The Performance of International Asset Allocation Strategies Using Conditioning Information. *Journal of Empirical Finance, 1*: 33-55.

Stein, C. (1955). Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. *Proceedings of the 3rd Berkeley Symposium on Probability and Statistics I*, Berkeley: University of California Press.

Stambaugh, R.F. (1999). Predictive Regressions. *Journal of Financial Economics, 54*: 375-421.

Treynor, J.L. & Mazuy, K. (1966). Can mutual funds outguess the market? Harvard

Business Review, 44: 131-136.
Torous, W., Valkanov, R. & Yan, S. (2005) On Predicting Stock Returns with Nearly Integrated Explanatory Variables. *Journal of Business, 78* (in press).
White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica, 48*: 817-838.