Do Return Prediction Models Add Economic Value?*

Tolga Cenesizoglu
HEC Montréal and CIRPÉE

Allan Timmermann
UCSD and CREATES

May 10, 2012

Abstract

We compare statistical and economic measures of forecasting performance across a large set of stock return prediction models with time-varying mean and volatility. We find that it is very common for models to produce higher out-of-sample mean squared forecast errors than a model assuming a constant equity premium, yet simultaneously add economic value when their forecasts are used to guide portfolio decisions. While there is generally a positive correlation between a return prediction model's out-of-sample statistical performance and its ability to add economic value, the relation tends to be weak and only explains a small part of the cross-sectional variation in different models' economic value.

Key words: predictability of stock returns, mean squared forecast error, economic and statistical measures of forecasting performance

*We thank an anonymous referee for many helpful suggestions. Seminar participants at HEC, University of Montreal, Imperial College, London, University of Toronto, Goldman Sachs and CREATES, University of Aarhus, provided valuable comments on a previous version of the paper. Timmermann acknowledges support from CREATES, funded by the Danish National Research Foundation.

1 Introduction

Can return forecasts deemed inaccurate by conventional statistical criteria such as (out-of-sample) mean squared forecast error or R^2 be of economic value to investors? This is a key question given the common finding in the stock return prediction literature that popular prediction models produce less accurate point forecasts of returns than a simple constant equity premium model, see, e.g., Campbell and Thompson (2008), Goyal and Welch (2008), and Rapach, Strauss and Zhou (2010). In fact, the relation between statistical and economic measures of forecasting performance is far from obvious given that the statistical criteria conventionally focus on the predictability of mean returns, while the economic value of a return forecast reflects the accuracy of predicted movements in the entire return distribution with weights that depend on the shape of the utility function.

To examine this issue, we analyze a variety of return prediction models that differ in whether they assume predictability in the mean and/or variance of returns. As a simple no-predictability benchmark in the spirit of Goyal and Welch (2008), we use a prevailing (constant) mean and variance model. We compare the performance of this model to that of time-varying mean, time-varying volatility and time-varying mean and volatility specifications in the EGARCH class studied by Glosten, Jagannathan and Runkle (1993). All in all, we consider four different mean-variance specifications and 15 different predictor variables, for a total of 60 return models.

To measure the statistical performance of the return forecasts, we use two measures of predictive accuracy, namely the (out-of-sample) root mean squared forecast errors (RMSE) and the weighted likelihood ratio test which measures the accuracy of probability distribution forecasts. Consistent with the evidence in Goyal and Welch (2008), we find that the simple constant mean model generates lower out-of-sample MSE-values than the majority of more complicated models that require the estimation of additional parameters. In contrast, the majority of models that allow for time-varying volatility generate more accurate density forecasts than the prevailing mean and volatility model when measured through the weighted likelihood ratio criterion.

We evaluate the economic significance of return predictability in two ways. First, we use our forecasts of the return distribution in an out-of-sample asset allocation exercise that combines stocks and T-bills. We consider two utility specifications−mean-variance utility and power utility−and three levels of risk aversion for a total of six different utility specifications. Portfolio performance is measured through the certainty equivalent return and the Sharpe ratio. Second, we use the non-parametric stochastic discount factor approach advocated by Almeida and Garcia (2008) to compute model-free estimates of the risk-adjusted returns associated with the individual prediction models relative to those associated with the constant mean and variance benchmark. Our empirical estimates of the economic performance measures suggest that allowing for a time-varying mean and volatility leads to improved portfolio performance. Specifically, we find that EGARCH models with a time-varying mean and volatility generate significant risk-adjusted returns of two to three percent per annum for the most successful individual predictor variables such as inflation, the term spread and the T-bill rate.

We next analyze the relationship between statistical and economic measures of forecasting performance and find that they can yield very different results. Specifically, across the 60 different sets of forecasts we find that it is very common for a model to produce higher RMSE-values than the prevailing mean model—and thus underperform in statistical terms—yet dominate this model in terms of economic performance measures such as the certainty equivalent return, risk-adjusted return, or the Sharpe ratio. Such disagreement between a model's (relative) statistical and economic performance arises for more than half of our model comparisons. It is much rarer for a model to produce worse predictive density forecasts than the prevailing mean model, yet outperform along economic criteria; this happens for less than 10% of the models.

A model's statistical performance is not, however, uninformative for its likely economic value. Specifically, for most utility specifications we find that there is a positive and significant correlation between the forecasting models' RMSE performance and their economic value. Hence the amount by which a model's RMSE performance exceeds or reduces that of the prevailing mean and variance benchmark carries information about the model's economic value. While the correlation between economic and statistical performance is positive, the predictive power of a model's RMSE-performance over its economic value tends to be low and is generally less than 10%.

Our paper is closely related to the large literature on predictability of the mean and volatility of stock returns. Empirical results have been somewhat mixed, indicating that while full-sample predictability of mean returns is quite strong, out-of-sample predictability of mean returns is considerably weaker and can be difficult to exploit. Ang and Bekaert (2007), Bali and Cakici (2010), Campbell (1987), Campbell and Shiller (1988), Campbell and Thompson (2008), Fama and French (1988, 1989), Ferson (1990), Ferson and Harvey (1993), Lettau and Ludvigsson (2001), Pesaran and Timmermann (1995), and Rapach, Strauss and Zhou (2010) find some evidence of predictability of the mean. Conversely, Bossaerts and Hillion (1999) and Goyal and Welch (2008) argue that the parameters of return prediction models are estimated with insufficient precision to render *ex-ante* return forecasts of much value.[1]

[1] Dangl and Halling (2008), Johannes, Korteweg and Polson (2009) and Paye and Timmermann (2006) attribute poor out-of-sample forecasting performance to parameter instability, while Lettau and van Nieuwerburgh (2008) propose shifts in the mean of predictor variables such as the dividend yield as a source of predictive failure.

Similarly, while the volatility of stock returns is known to follow a pronounced counter-cyclical pattern (Schwert (1989)), there is relatively weak evidence that macroeconomic state variables contain information useful for predicting stock market volatility. Engle, Ghysels and Sohn (2007) find some evidence that inflation volatility helps predict the volatility of stock returns. However, the volatility of interest rate spreads and growth in industrial production, GDP or the monetary base fail to consistently predict future volatility, with evidence being particularly weak in the post-WWII sample. This is consistent with findings in Paye (2010) and Ghysels, Santa-Clara and Valkanov (2006).

The outline of the paper is as follows. Section 2 introduces the models used to capture predictability of the return distribution while Section 3 presents the data and evaluates the statistical forecasting performance of the various models. Section 4 considers the economic value of return forecasts in portfolio selection experiments and Section 5 addresses the relationship between the statistical and economic measures of forecasting performance. Finally, Section 6 concludes.

2 Return Predictability Models

An extensive literature in finance has examined predictability of the mean and variance of stock returns; see Rapach and Zhou (2012) for a recent review. Our aim is to explore the statistical and economic performance of commonly used prediction models and so we follow the literature and focus on popular models such as linear regressions and time-varying volatility models.

As a simple no-predictability benchmark, we use a constant mean and volatility model:

$$r_{t+1} = \beta_0 + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma^2). \tag{1}$$

Here r_{t+1} is the stock return computed in excess of a safe T-bill rate. This model is similar in spirit to the prevailing mean model of Goyal and Welch (2008) with the addition of a constant volatility and the assumption of normally distributed returns. We refer to this as the prevailing mean and variance (PMV) model.

2.1 Time-variations in the Mean and Volatility of Returns

To explore if time-variations in the mean, volatility, or both, lead to better return forecasts, predictability of the mean and volatility of returns is analyzed in separate steps.

First, we allow for a time-varying mean (TVM), but hold volatility constant through the following model:

$$r_{t+1} = \beta_0 + \beta_1 x_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma^2), \tag{2}$$

where x_t is a state variable known at time t.

Conversely, a model that allows for time-varying volatility, but holds the mean constant can be studied through an EGARCH specification of the type considered by Glosten, Jagannathan, and Runkle (1993):

$$\begin{aligned} r_{t+1} &= \beta_0 + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma_{t+1}^2), \\ \log(\sigma_{t+1}^2) &= \delta_0 + \delta_2 \log(\sigma_t^2) + \delta_3 \left| \frac{\varepsilon_t}{\sigma_t} \right| + \delta_4 \frac{\varepsilon_t}{\sigma_t}. \end{aligned} \tag{3}$$

In this model the conditional volatility, σ_{t+1}, depends on the lagged volatility, σ_t, as well as the return innovation, ε_t, whose effect can depend on whether it is positive or negative. We refer to this as the PM-EGARCH model.

This model is naturally extended to a prevailing mean EGARCHX (PM-EGARCHX) model

$$\begin{aligned} r_{t+1} &= \beta_0 + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma_{t+1}^2), \\ \log(\sigma_{t+1}^2) &= \delta_0 + \delta_1 x_t + \delta_2 \log(\sigma_t^2) + \delta_3 \left| \frac{\varepsilon_t}{\sigma_t} \right| + \delta_4 \frac{\varepsilon_t}{\sigma_t}, \end{aligned} \tag{4}$$

which allows the economic state variables to affect the conditional variance. Conversely, the time-varying mean and volatility EGARCH (TVM-EGARCH) specification

$$\begin{aligned} r_{t+1} &= \beta_0 + \beta_1 x_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma_{t+1}^2), \\ \log(\sigma_{t+1}^2) &= \delta_0 + \delta_2 \log(\sigma_t^2) + \delta_3 \left| \frac{\varepsilon_t}{\sigma_t} \right| + \delta_4 \frac{\varepsilon_t}{\sigma_t}, \end{aligned} \tag{5}$$

allows the state variables to affect the conditional mean but not the volatility.

Finally, the most general model, labeled the time-varying mean EGARCHX (TVM-EGARCHX) model, allows the economic state variables to affect both the mean and volatility equations:

$$\begin{aligned} r_{t+1} &= \beta_0 + \beta_1 x_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma_{t+1}^2), \\ \log(\sigma_{t+1}^2) &= \delta_0 + \delta_1 x_t + \delta_2 \log(\sigma_t^2) + \delta_3 \left| \frac{\varepsilon_t}{\sigma_t} \right| + \delta_4 \frac{\varepsilon_t}{\sigma_t}. \end{aligned} \tag{6}$$

The models in equations (2)-(5) are all nested by Eq. (6). However, for purposes of forecasting returns out-of-sample it is possible that the simpler models dominate the most general specification since they have fewer parameters to estimate. The parameters of all the EGARCH models considered here can be estimated using maximum likelihood methods.

3 Statistical Measures of Return Predictability

We next describe the statistical measures used to measure return predictability and present initial empirical results. First, however, we introduce the data used in the analysis.

3.1 Data

Our empirical analysis uses a data set comprising monthly stock returns along with a set of fifteen predictor variables previously analyzed in Goyal and Welch (2008), and extended to the end of 2010.[2] Stock returns are measured by the S&P500 index and include dividends. A short T-bill rate is subtracted from stock returns to obtain excess returns.

The fifteen predictor variables fall into four broad categories:

- Valuation ratios capturing some measure of 'fundamental' value to market value such as
  - dividend-price ratio (d/p);
  - dividend yield (d/y);
  - earnings-price ratio (e/p);
  - 10-year earnings-price ratio ($e10/p$);
  - book-to-market ratio (b/m);
- Bond yield measures capturing level or slope effects from the term structure or measures of default risk, including
  - three-month T-bill rate (tbl);
  - yield on long term government bonds (lty);
  - term spread as measured by the difference between the yield on long-term government bonds and the three-month T-bill rate (tms);
  - default yield spread as measured by the yield spread between BAA and AAA rated corporate bonds (dfy);
  - default return spread as measured by the difference between the yield on long-term corporate bonds and government bonds (dfr);
- Estimates of equity risk and return such as
  - long term return (ltr);
  - stock variance, i.e., a volatility estimate based on daily squared returns ($svar$);
- Corporate finance variables, including

[2]We are grateful to Amit Goyal for providing this data.

– dividend payout ratio measured by the log of the dividend-earnings ratio (d/e);

– net equity expansion measured by the ratio of 12-month net issues by NYSE-listed stocks over their end-year market capitalization ($ntis$);

Finally, we also consider the inflation rate ($infl$) measured by the rate of change in the consumer price index. Additional details on data sources and the construction of these variables are provided by Goyal and Welch (2008).

3.2 Statistical measures of predictive accuracy

To measure the accuracy of the forecasts from the return prediction models, we report the dominant statistical measure of forecasting performance used in the finance literature, namely the root mean squared error (RMSE). The RMSE-criterion only considers the precision of point forecasts and so focuses on predictability of the mean return. It is possible, however, that the mean is hard to predict, but that other parts of the return distribution can be predicted. To examine if this holds, we use the weighted likelihood ratio test statistic of Amisano and Giacomini (2007) to analyze the performance of individual density forecasts relative to the forecast implied by the PMV model. The test statistic takes the form

$$t = \frac{\overline{WLR}_P}{\widehat{\sigma}_P/\sqrt{P}}, \tag{7}$$

where $\overline{WLR}_P = P^{-1}\sum_{t=0}^{P-1} WLR_{t+1}$ is the average weighted likelihood ratio using P out-of-sample observations and $\widehat{\sigma}_P$ is an estimator of its variance. Each period the weighted likelihood ratio (WLR_{t+1}) is computed as the weighted average difference between the log scores of an individual model and the PMV model, evaluated at the actual return, r_{t+1}:

$$WLR_{t+1} = w(\bar{r}_{t+1})(\log f_{t+1|t}(r_{t+1}) - \log f_{t+1|t}^{PMV}(r_{t+1})), \tag{8}$$

where $w(\bar{r}_{t+1})$ is a weight function evaluated at the standardized return at time $t+1$ ($\bar{r}_{t+1}$). $f_{t+1|t}$ and $f_{t+1|t}^{PMV}$ are the predictive densities of the candidate model and the PMV model, respectively. We set $w(.) = 1$ throughout the analysis, thus focusing on the full distribution.

3.3 Empirical results

Throughout the analysis we consider the five forecasting models described in Section 2:

1. Prevailing mean and variance (PMV, Eq. (1)).

2. Time-varying mean (TVM, Eq. (2)).

3. Constant mean and time-varying volatility (PM-EGARCH and PM-EGARCHX, Eqs. (3, 4)),
4. Time-varying mean and volatility (TVM-EGARCH, Eq. (5)).
5. Time-varying mean and volatility with economic predictors (TVM-EGARCHX, Eq. (6)).

Estimates of the slope coefficients on the economic state variables for each of these models are presented in Table 1. Using conventional t-tests, the results indicate only sparse evidence of predictability of the dynamics in the first and second moments of returns. For the mean equation, the coefficient on the long-term equity return is significant at the 1% critical level in all three model specifications that allow for a time-varying mean, while the stock variance and the term spread are significant at the 10% critical level for two of the three models. Turning to the variance specifications, the coefficients on the stock variance, default return spread and the T-bill rate are significant at the 5% level in both the PM-EGARCHX and TVM-EGARCHX models.

We next turn to the out-of-sample prediction performance. To generate out-of-sample forecasts, we use a 30-year rolling window of the data. This guards against changes in the data generating process while also reducing the effect of parameter estimation error which would be greater under a shorter estimation window. Specifically, the first set of parameter estimates are produced for the 30-year sample, 1926:01-1955:12. One-step-ahead forecasts are then generated for returns in 1956:01. The following month our estimates are updated by rolling the estimation window one observation ahead and adding data for 1956:01. The updated model is used to produce return forecasts for 1956:02. This recursive forecasting procedure is repeated up to the end of the sample, generating 660 out-of-sample forecasts for the period 1956:01-2010:12.

Table 2 reports RMSE-values for the four forecasting models that allow for either a time-varying mean, time-varying variance, or both. In all cases we report the RMSE-value relative to that obtained by the benchmark PMV model, i.e. $(RMSE(PMV) - RMSE(Model))/RMSE(PMV)$. Positive values therefore indicate that the candidate model produces a lower RMSE than the PMV model, and so generates more accurate point forecasts, while negative values indicate inferior performance relative to the benchmark.

Consistent with the evidence in Goyal and Welch (2008), very few of the univariate models produce lower RMSE-values than the PMV model. In fact, only 11 out of 60 models produce lower RMSE-values than the PMV model. Only the models based on the long term return, inflation, the dividend yield or the dividend price ratio generate quite good point forecasts: For these specifications we find that two of three models with time-varying mean produce forecasts with a lower RMSE than the PMV model. The worst forecasting performance is found for the PM-EGARCHX models where not a single model dominates the PMV model along the RMSE-criterion, illustrating how estimation of additional parameters can adversely affect forecasting performance by this criterion.[3]

Table 3 reports values of the weighted likelihood ratio test applied to our data. Once again, positive values indicate that the model produces better density forecasts than the PMV model, while negative values suggest that the PMV model is (relatively) more accurate. For the TVM model the predictive density and RMSE results are largely in line, both suggesting that only the models based on the long term return, inflation, dividend yield or dividend-price ratio lead to better forecasts than the PMV model, whereas the other models are worse. This is as we would expect, since the TVM models only allow for predictability of the mean and so the two statistical criteria should yield similar results.

For the three remaining models that allow for a time-varying volatility, the predictive density results are very different from the RMSE results, however. In every single case, the specifications that allow for time-varying volatility lead to better out-of-sample density forecasts than the PMV model, with more than half of these comparisons being significant at the 10% critical level. In contrast, only six of these 45 models lead to lower RMSE-performance than the PMV model.[4]

We summarize the empirical analysis up to this point as follows. Consistent with the existing literature (e.g., Goyal and Welch (2008)) we find that very few models produce more accurate point forecasts than the simple prevailing mean and variance model, when measured by their RMSE-performance. In contrast, however, models that allow for a time-varying variance of returns produce much better density forecasts than this benchmark.

We next turn to a discussion of the economic performance measures.

4 Economic Measures of Return Predictability

The previous section indicated strong evidence of out-of-sample predictability of the density of stock returns, though not of the mean of this distribution. This section asks if such predictability is sufficiently large to be of economic value to investors by considering return predictability from the perspective of a small, risk averse investor with no market impact who chooses portfolio weights based on the return forecasts. We first explain how the forecasts are used in the portfolio selection procedure, then describe the economic performance measures and finally report empirical findings.

[3] The relative RMSE-value for the PM-EGARCH model is -0.15954% and thus is worse for all but one of the PM-EGARCHX models.

[4] For the PM-EGARCH model the weighted likelihood ratio test is 0.0323 which is smaller than that of all but one of the PM-EGARCHX models.

4.1 Portfolio Selection

Consider an investor who at time t allocates $w_t W_t$ of total wealth to stocks and the remainder, $(1-\omega_t)W_t$ to a risk-free asset, where $W_t = 1$ is the initial wealth, while ω_t is the share allocated to stocks. The investor's wealth at time $t+1$ is

$$W_{t+1} = (1-\omega_t)\exp(r^f_{t+1}) + \omega_t \exp(r^f_{t+1} + r_{t+1}), \tag{9}$$

where r_{t+1} is the stock return in excess of the risk-free rate, r^f_{t+1}, both continuously compounded.

Portfolio weights for period t can be obtained as the solution to the following optimization problem:

$$\omega^*_t = \arg\max_{\omega_t} E_t[U(W_{t+1})], \tag{10}$$

where $E_t[\cdot]$ denotes the conditional expectation based on the investor's information at time t. The investment horizon is set to one period and any intertemporal hedging component in the investor's portfolio choice is ignored. Hence we assume that the investor solves Eq. (10), holds the optimal portfolio for one period and then reoptimizes the portfolio weights the following period based on any new information.

We consider two specifications of investor preferences. First, under mean-variance utility,

$$U(W_{t+1}) = E_t[W_{t+1}] - \frac{\gamma}{2} Var_t(W_{t+1}),$$

where $E_t[.]$ and $Var_t(.)$ denote the conditional mean and variance of returns, respectively, and γ reflects the investor's absolute risk aversion. For this investor, the optimal portfolio holding, conditional on the predicted mean and variance, $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}^2_{t+1}$, is

$$\omega^*_t = \frac{\exp(\widehat{\mu}_{t+1} + \widehat{\sigma}^2_{t+1}/2) - 1}{\gamma \exp(r^f_{t+1})(\exp(\widehat{\sigma}^2_{t+1}) - 1)\exp(2\widehat{\mu}_{t+1} + \widehat{\sigma}^2_{t+1})}.$$

Second, we consider an investor with constant relative risk aversion (CRRA) preferences,

$$U(W_{t+1}) = \frac{W^{1-\gamma}_{t+1}}{1-\gamma}, \tag{11}$$

where γ is the investor's coefficient of relative risk aversion.

Suppose that the conditional distribution of log excess returns is normal with predicted mean and variance $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}^2_{t+1}$. Campbell and Viceira (2001) derive the following log-linearized approximation to the investor's wealth at time $t+1$:

$$\ln(W_{t+1}) \approx \ln(1 + r^f_{t+1}) + \omega_t r_{t+1} + \frac{1}{2}\omega_t(1-\omega_t)\widehat{\sigma}^2_{t+1}, \tag{12}$$

from which the approximate optimal portfolio weight for the CRRA investor is obtained:

$$\omega_t^* \approx \frac{\widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}^2/2}{\gamma \widehat{\sigma}_{t+1}^2}. \tag{13}$$

We use this equation to derive the optimal portfolio allocation of the CRRA investor under the models for returns in equations (1) - (6). In each case the return forecasts $\widehat{\mu}_{t+1}$, $\widehat{\sigma}_{t+1}^2$ are converted into portfolio weights, ω_t, and the resulting next-period return and utility are recorded. Following Kandel and Stambaugh (1996) and Geweke (2001), we restrict $\omega_t \in [0, 0.99]$ to ensure that the expected utility is finite under CRRA preferences. For comparability, we also initially impose this constraint under mean-variance utility, but subsequently relax the constraint to $\omega_t \in [0, 1.50]$ along the lines proposed by Campbell and Thompson (2008).

4.2 Economic measures of performance

Three different measures are used to evaluate the economic performance of the return forecasts. First, the selected portfolios are compared through their certainty equivalent return (CER) values. Under mean-variance utility, the CER is computed simply as the average realized utility over the out-of-sample period:

$$CER = P^{-1} \sum_{t=R+1}^{R+P} U(W_t^*), \tag{14}$$

where R is the length of the estimation sample (here, 30 years), while P is the length of the out-of-sample period that runs from 1956:1 through 2010:12.

Under CRRA utility, the CER is computed as

$$CER = \left((1-\gamma) P^{-1} \sum_{t=R+1}^{R+P} U(W_t^*) \right)^{1/(1-\gamma)} - 1, \tag{15}$$

where $P^{-1} \sum_{t=R+1}^{R+P} U(W_t^*)$ is again the mean realized utility.

Our second measure of performance is the Sharpe ratio, calculated as the mean portfolio excess return divided by the portfolio return volatility, both computed over the out-of-sample period.

To see if the risk-adjusted returns associated with a specific forecasting model are significantly higher than those associated with the benchmark PMV model, we use the non-parametric stochastic discount factor approach of Almeida and Garcia (2008). In a first stage, the stochastic discount factor is backed out from data on a set of underlying risk factors. In a second stage, returns on a test portfolio are compared to a set of competing (benchmark) returns using the sample analog of an Euler equation. If the sample estimate of the product of the differential portfolio returns and the stochastic discount factor is positive and significant, we conclude that the test portfolio provides higher risk-adjusted returns than its competitor.

Specifically, we first estimate the stochastic discount factor from the conventional asset pricing model, $E(m_t F_t - 1) = 0$, where 1 and 0 are vectors of ones and zeros, respectively, and F_t is a vector of gross returns on a set of underlying risk factors that are priced by the stochastic discount factor, m_t. The sample analog of this equation is

$$\frac{1}{T}\sum_{t=1}^{T} m_t(F_t - \frac{1}{\overline{m}}) = 0, \tag{16}$$

where $\overline{m} = E(m_t)$ is the unconditional mean of the stochastic discount factor.

The sample estimate of the stochastic discount factor implied by the Almeida-Garcia (2008) approach is

$$\widehat{m}_t = T \times \overline{m} \times \frac{(1 - \psi\widehat{\varphi}'(F_t - \frac{1}{\overline{m}}))^{\frac{1}{\psi}}}{\sum_{t=1}^{T}(1 - \psi\widehat{\varphi}'(F_t - \frac{1}{\overline{m}}))^{\frac{1}{\psi}}}. \tag{17}$$

Here ψ can be interpreted as the parameter of an investor with HARA preferences of the form $U(W) = -\frac{1}{1+\psi}(1 - \psi W)^{\frac{1+\psi}{\psi}}$. Moreover, $\widehat{\varphi}$ can be obtained from the solution to an empirical likelihood problem.[5]

The performance of the returns on a specific portfolio, r_{jt}, measured relative to those of the PMV benchmark, r_t^{PMV}, can now be computed using the sample analog of the Euler equation for the estimated stochastic discount factor, $\widehat{m}_t$, and the differential portfolio returns, $r_{jt} - r_t^{PMV}$:

$$\widehat{\alpha}_j = \frac{1}{P}\sum_{t=R}^{R+P} \widehat{m}_t(r_{jt} - r_t^{PMV}), \tag{18}$$

where $\widehat{\alpha}_j$ measures the risk-adjusted return of portfolio j in excess of the PMV benchmark return.

In the empirical analysis monthly returns on the S&P500 are used to estimate the implied stochastic discount factor. The hyperparameter ψ is set to one-half, although the results are not sensitive to this choice. The unconditional mean of the stochastic discount factor, $\overline{m}$, is set to the reciprocal of the mean gross return on the risk free asset during the sample period 1956:01-2010:12. Performance is measured relative to the PMV model and p-values for the risk-adjusted returns associated with the sample Euler equation in (18) are calculated using the stationary bootstrap (Politis and Romano (1994)) on the alpha estimates.[6]

[5] Almeida and Garcia (2008) establish conditions under which

$$\widehat{\varphi} = \sup_{\varphi \in \Upsilon} \frac{\overline{m}^{1+\psi}}{T}\sum_{t=1}^{T} -\frac{1}{1+\psi}\left(1 - \psi\varphi'\left(F_t - \frac{1}{\overline{m}}\right)\right)^{\frac{1+\psi}{\psi}},$$

where φ is constrained to the set $\Upsilon = \{\varphi \in \mathbb{R}^k$, s.t. $t = 1, 2, \ldots, T : (1 - \psi\varphi'(F_t - \frac{1}{\overline{m}})) > 0\}$ and ψ is a hyperparameter of the so-called Cressie Read discrepancy function $\frac{x^{1+\psi}-1}{(1+\psi)\psi}$.

[6] Specifically, we generate 999 random samples of 660 observations from the estimated relative performance measures using the centered stationary bootstrap approach and a block size determined by a geometric distribution with a parameter of 0.05.

4.3 Empirical Results

Tables 4 and 5 present certainty equivalent return estimates, the Sharpe ratio, and risk-adjusted returns based on the out-of-sample return forecasts. To explore the sensitivity of the results to different preferences, we consider mean-variance utility as well as CRRA preferences, in both cases entertaining a coefficient of risk aversion, $\gamma = 3, 5$, or 10.

First consider the results based on mean-variance preferences (Table 4). For the TVM model we find mild evidence of improved economic performance relative to the PMV model, with a little under half of the 15 models bettering the CER or Sharpe ratio of the benchmark. The evidence is a bit stronger when judged by the risk-adjusted returns for which between 9 and 11 models produce better results, the models based on the long-term return, inflation or the term spread significantly so.

Turning to the three models that allow for time-varying volatility, the evidence is typically much stronger with the clear majority of models producing better economic performance than the PMV model. The strongest results are obtained for the PM-EGARCHX model which keeps the mean constant but allows the conditional variance to fluctuate through time. This model produces better out-of-sample performance for almost all three economic measures, irrespective of the level of risk aversion. Recalling from Table 2 that this specification produced worse RMSE performance than the PMV model across all 15 state variables, this is a clear illustration of the difference, indeed sharp contrast, that can exist between economic and statistical measures of predictive accuracy.

Under power utility (Table 5), the results are very comparable to those obtained for mean-variance utility, assuming a coefficient of risk aversion of $\gamma = 3$ or $\gamma = 5$. However, when γ rises to ten, the results differ in the case of the CER-values measured relative to the PMV model, which are higher under mean-variance utility but lower under power utility for most of the models that allow for time-varying volatility. For the two other measures of economic performance the results under power utility are in line with those obtained under mean-variance utility.

The results in tables 4 and 5 assume that investors cannot go short in the market portfolio or use leverage, i.e., $\omega_t \in [0, 0.99]$. They also do not impose economic constraints on the forecasts like the non-negative ex-ante equity risk premium proposed by Campbell and Thompson (2008). To see how these choices affect our results, Table 6 presents risk-adjusted relative return performance results for a mean-variance investor when we allow $\omega_t \in [0, 1.50]$, so the investor can use up to 50% leverage as assumed by Campbell and Thompson. Although the average risk-adjusted return performance improves somewhat, particularly for the less conservative investor ($\gamma = 3$), the basic results concerning economic and statistical significance of return predictability are very similar to the baseline results in Table 4. Table 7 shows results under mean-variance preferences and $\omega_t \in [0, 1.50]$ when we impose the non-negative equity risk premium constraint of Campbell and Thompson (2008). The results are very close to the comparable ones in Table 6, suggesting that imposing such constraints has little effect on our conclusions.

5 Relating the Statistical and Economic Measures of Forecasting Performance

Our empirical results suggest that a return prediction model with constant mean performs better than models with time-varying mean and/or volatility in terms of out-of-sample RMSE performance. However, this model generally performs poorly when measured by the economic value of its forecasts as used to guide portfolio selection. The opposite finding holds for models that allow for time-varying mean and volatility which typically produce poor forecasts of mean returns but often add economic value when used for portfolio decisions. This raises an important question, namely how closely related are different statistical and economic measures of forecasting performance?

It turns out that the message conveyed from a model's RMSE performance can be very different from that emerging from an analysis of the model's economic value. To see this, Table 8 shows the percentage of cases, for which we find that a model has a worse statistical but a better economic performance than the PMV benchmark. For most utility specifications between 50 and 70 percent of the prediction models generate higher RMSE-values, yet at the same time better economic performance than the PMV model. This strongly demonstrates that their inability to predict time-variations in the (mean) equity premium does not preclude the prediction models from adding economic value. Using instead the statistical measure of predictive accuracy based on the full density, we find that only between seven and 10 percent of the models generate worse statistical performance, yet better economic performance than the PMV benchmark.

To further examine these findings, we plot the individual models' certainty equivalent return and Sharpe ratio performance against their RMSE (Figure 1) or weighted likelihood ratio performance (Figure 2). In these graphs, the zero point on the axes indicates performance identical to that of the PMV model. Observations in the second quadrant therefore show better economic performance, but worse statistical performance, relative to the PMV benchmark, while observations in the fourth quadrant show worse economic performance, but better statistical performance. Such points indicate disagreement or 'dissonance' between the two criteria. The fact that most points fall in the second quadrant illustrates our earlier point that economic and statistical performance measures commonly disagree.

Figure 1 shows that the PM-EGARCHX models generate a tight distribution of (negative) RMSE-values, measured relative to the benchmark. In contrast, the TVM-EGARCHX forecasts generate the largest spread in RMSE- and CER-values. Interestingly forecasts from these models only have limited downside when measured by the Sharpe ratio.

Figure 2 confirms that only the TVM models produce worse density forecasts than the PMV model, although for these models we see a wide dispersion in the (relative) CER values and Sharpe ratio estimates. The three specifications with time-varying volatility produce very similar spreads in CER performance across the individual models.

Table 9 reports simple (linear) and rank correlations between, on the one hand, the two statistical measures of forecasting performance (RMSE or predictive density accuracy), and, on the other hand, the three economic measures of performance (CER, Sharpe ratio, and risk-adjusted returns). The rank correlation between RMSE and economic performance is generally high and falls in the range between 0.24 and 0.53. The corresponding linear correlation measure is much lower and lies in the range between 0.06 and 0.37. Interestingly, linear correlations between RMSE and economic performance tend to be higher under power utility than under mean variance utility. In most cases the R^2 from a cross-sectional regression of the individual models' economic performance on their statistical performance is modest with typical values below 10%.

Turning to the linear or rank correlation between the predictive density accuracy and the economic performance measure, with a single exception, again we find positive values of up to 0.50. The R^2-values from a cross-sectional regression of the individual models' economic performance on their statistical performance exceed 0.20 only under mean-variance utility with risk aversion of $\gamma = 3$, and otherwise are smaller than 0.10.

Comparing the strength of the correlation between different economic and statistical performance measures reveals an interesting picture. In general we find a stronger correlation between RMSE statistical performance and economic performance than between the probability density statistical measure and economic performance. The exception is under mean-variance utility where we see a much stronger positive correlation between CER and the probability density measure than between CER and the RMSE-values.

When we compare the statistical and economic performance measures separately within the class of TVM models or separately among the PM-EGARCHX models across specifications that include different economic state variables, somewhat higher correlations are obtained. This is to be expected since these form more homogenous subsets of models. However, the power of the statistical measures of predictive accuracy over economic performance remains muted as the proportion of the cross-sectional variation in economic performance explained by a model's statistical performance is less than one-third for the majority of cases.

In addition to presenting results for the baseline scenario with $\omega_t \in [0, 0.99]$, Tables 8 and 9 also show results under mean-variance preferences when we allow investors to apply up to 50% leverage and let $\omega_t \in [0, 1.50]$ as in the analysis of Campbell and Thompson (2008). This change has the biggest effect for $\gamma = 3$, for which we see an increase in the disagreement between the RMSE and economic measures of performance from 59% to 66% of all cases (Sharpe ratio) and from 61% to 72% (risk-adjusted returns). Conversely, for the more conservative investors we see slightly lower levels of disagreement, e.g., from 62% to 56% (Sharpe ratio) and from 72% to 69% (risk-adjusted returns) when $\gamma = 5$. Interestingly, despite the increased disagreement, Table 9 shows a slightly higher correlation between the statistical and economic performance measures under the wider band for the portfolio weights, particularly when $\gamma = 3$. Imposing the Campbell-Thompson non-negativity constraint on the equity risk premium has very little effect on the results in the two tables.

We conclude the following from this analysis. Our results show that a model underperforms the constant benchmark in terms of RMSE performance does not carry much information about its ability to add value when used for portfolio selection. However, how much the model's RMSE performance exceeds or reduces that of the PMV benchmark *does* carry information about the model's economic value: in most cases there is a positive and significant correlation between the model's RMSE performance and its economic value. While this correlation is positive, it is tempered by the low predictive power of a model's RMSE-performance over its economic value.

6 Conclusion

This paper examines a large number of predictor variables, different preference specifications, and different forecasting models that allow for predictability of the conditional mean and/or the conditional variance of stock returns. We find evidence that allowing for predictability in both the mean and variance of returns leads to more accurate forecasts of the probability distribution of stock returns although it does not lead to better forecasts of mean returns. Return prediction models that use economic covariates to predict the second moment of the return distribution lead to particular improvements in economic performance.

We find that underperformance along conventional measures of forecasting performance such as root mean squared forecast errors contain little information when it comes to assessing whether return prediction models that allow for a time-varying mean or variance help or hurt investors. Many models underperform the constant mean and variance benchmark according to the RMSE criterion, but nevertheless produce superior economic performance when their forecasts are used as the basis for investment decisions. It is less common to find that a model produces less accurate probability distribution forecasts, while simultaneously outperforming in economic terms.

Our analysis suggests that the debate on return predictability has focused too narrowly on statistical measures of forecast precision such as root mean squared forecast errors or R^2 and that important insights can be learned by focusing instead on economic measures of forecasting performance.

References

[1] Almeida, C. and R. Garcia, 2008, Empirical Likelihood Estimators for Stochastic Discount Factors. Working paper, Getulio Vargas Foundation and Edhec, Nice.

[2] Amisano, G., and R. Giacomini, 2007, Comparing Density Forecasts via Weighted Likelihood Ratio Tests. Journal of Business and Economic Statistics 25, 177-190.

[3] Ang, A. and G. Bekaert, 2007, Stock Return Predictability: Is it There? Review of Financial Studies 20, 651-707.

[4] Bali, T. and N. Cakici, 2010, World Market Risk, Country-Specific Risk and Expected Returns in International Stock Markets. Journal of Banking and Finance 34(6), 1152-1165.

[5] Bossaerts, P. and P. Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What do we Learn? Review of Financial Studies 12(2), 405-428.

[6] Campbell, J.Y., 1987, Stock Returns and the Term Structure. Journal of Financial Economics 18(2), 373-399.

[7] Campbell, J.Y., and R., Shiller, 1988, The Dividend-Price Ratio, Expectations of Future Dividends and Discount Factors. Review of Financial Studies 1, 195-227.

[8] Campbell, J.Y. and S.B. Thompson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21, 1509-1532.

[9] Campbell, J.Y., and L.M. Viceira, 2001, Who should buy Long-Term Bonds? American Economic Review 91, 99-127.

[10] Dangl, T. and M. Halling, 2008, Predictive Regressions with Time-Varying Coefficients. forthcoming Journal of Financial Economics.

[11] Engle, R.F., E. Ghysels and B. Sohn, 2007, On the Economic Sources of Stock Market Volatility. Manuscript, New York University and University of North Carolina.

[12] Fama, E.F. and K.R. French, 1988, Dividend Yields and Expected Stock Returns. Journal of Financial Economics 22(1), 3-25.

[13] Fama, E.F. and K.R. French, 1989, Business Conditions and Expected Returns on Stocks and Bonds. Journal of Financial Economics 25(1), 23-49.

[14] Ferson, W., 1990, Are the Latent Variables in Time-Varying Expected Returns Compensation for Consumption Risk?, Journal of Finance, 45, 397-429.

[15] Ferson, W., and C., Harvey, 1993, The Risk and Predictability of International Equity Returns, Review of Financial Studies, 6, 527-566.

[16] Geweke, J., 2001, A Note on Some Limitations of CRRA Utility. Economics Letters 71, 341-345.

[17] Ghysels, E., P. Santa-Clara and R. Valkanov, 2006, Predicting Volatility: How to get the Most out of Return Data Sampled at Different Frequencies. Journal of Econometrics 131, 59-95.

[18] Glosten, L.R., R. Jagannathan, and D.E. Runkle, 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance 48, 1779-2801.

[19] Goyal, A., and I. Welch, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. Review of Financial Studies 21, 1455-1508.

[20] Johannes, M., A. Korteweg and N. Polson, 2009, Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns. Mimeo, Columbia University.

[21] Kandel, S. and R. Stambaugh, 1996, On the Predictability of Stock Returns: An Asset Allocation Perspective. Journal of Finance 51, 385-424.

[22] Lettau, M. and S. Ludvigsson, 2001, Consumption, aggregate wealth, and expected stock returns. Journal of Finance 56, 815-850.

[23] Lettau, M. and S. van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence. Review of Financial Studies 21, 1607-1652.

[24] Paye, B., 2010, Which Variables Forecast Aggregate Stock Market Volatility? Working Paper, Rice University.

[25] Paye, B. and A. Timmermann, 2006, Instability of Return Prediction Models. Journal of Empirical Finance 13, 274-315.

[26] Pesaran, M.H. and A. Timmermann, 1995, Predictability of Stock Returns: Robustness and Economic Significance. Journal of Finance 50(4), 1201-1228.

[27] Rapach, D.E., J.K. Strauss, and G. Zhou, 2010, Out-of-sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. Review of Financial Studies 23, 821-862.

[28] Rapach, D.E., and G. Zhou, 2012, Forecasting Stock Returns. Unpublished manuscript, University of Washington at St. Louis.

[29] Politis, D.N., and J.P. Romano, 1994, The Stationary Bootstrap. Journal of the American Statistical Association 89, 1303-1313.

[30] Schwert, G.W., 1989, Why does Stock Market Volatility Change over Time? Journal of Finance 44, 1115-1153.

Table 1: Coefficient Estimates from Linear Prediction and EGARCH Models

	TVM Mean Eq.	TVM-EGARCH Mean Eq.	TVM-EGARCHX Mean Eq.	PM-EGARCHX Variance Eq.	TVM-EGARCHX Variance Eq.
d/e	0.0052	0.0023	0.0054	-0.0349	-0.0531
svar	-1.0817**	-1.0241*	1.5149	102.1899***	103.0472***
dfr	0.1631	0.0800	0.0464	-6.9164**	-6.8686**
lty	-0.0225	0.0221	-0.0548	1.0632*	1.2325*
ltr	0.1477**	0.1658***	0.1671***	-1.4194	-1.5991
infl	-0.8295	-1.1243***	-1.3190***	10.9216	13.1027**
tms	0.2285*	0.1585**	0.1712	-0.5374	-0.7758
tbl	-0.0774	-0.0297	-0.0923	0.9181**	1.1260***
dfy	0.4389	0.6129*	0.4479	5.8670	4.9594
d/p	0.0056	0.0033	0.0044	-0.0054	-0.0164
d/y	0.0063	0.0037	0.0044	-0.0052	-0.0195
e/p	0.0020	0.0025	0.0022	0.0074	0.0016
b/m	0.0028	0.0008	0.0013	-0.0040	-0.0075
e10/p	0.1507	0.1090	0.1237	0.1993	-0.1639
ntis	-0.0553	-0.2240***	-0.1237	-1.6445**	-1.4026

Note: The coefficient estimates are based on estimation of linear prediction and EGARCH models over the sample 1956:01-2010:12. The time-varying mean (TVM) model allows the mean to vary over time as a function of the variable listed in the rows but assumes constant variance. The estimates for the TVM model are based on univariate regressions of monthly S&P 500 excess returns on lagged values of the predictor variables listed in the rows. The TVM-EGARCH models allow for both time-varying mean and variance but includes the predictor variables only in the mean equation but not in the variance equation. The PM-EGARCHX and TVM-EGARCHX models both allow for time-varying variances with predictor variables in the variance equation, with only the latter model including the predictor variables also in the mean equation. All EGARCH models assume conditionally normally distributed return innovations and are based on maximum likelihood estimation with the lagged predictor variables listed in the rows included as additional variables in the mean and/or variance equations. All standard errors are corrected for heteroskedasticity and autocorrelation. * significant at the 10% level. ** significant at the 5% level. *** significant at the 1% level.

Table 2: Relative Root Mean Squared Forecast Errors

	TVM	TVM-EGARCH	PM-EGARCHX	TVM-EGARCHX
d/e	-0.4484%	-0.2473%	-0.1148%	-0.2604%
svar	-0.1472%	-1.6086%	-0.0791%	-3.0958%
dfr	-0.4719%	-0.8370%	-0.0947%	-0.5124%
lty	-0.4411%	-0.3307%	-0.0459%	-0.9312%
ltr	0.1805%	0.1446%	-0.1237%	0.1321%
infl	0.3491%	0.0432%	-0.1057%	0.2083%
tms	0.1154%	-0.0083%	-0.1203%	-0.0619%
tbl	-0.2102%	-0.4040%	-0.0874%	-1.1427%
dfy	-0.4108%	-0.7731%	-0.1243%	-0.5221%
d/p	0.0724%	-0.3086%	-0.1103%	0.1643%
d/y	0.0184%	-0.0787%	-0.1080%	0.1184%
e/p	-0.3690%	-0.5471%	-0.1194%	-0.4322%
b/m	-0.5332%	-0.8363%	-0.0887%	-0.4509%
e10/p	-0.3802%	-0.3675%	-0.1036%	-0.1305%
ntis	-0.0964%	-0.1341%	-0.1946%	-0.1344%

Note: This table presents relative root mean square errors of models with the lagged predictor variables listed in the rows with respect to the prevailing mean and variance (PMV) model that assumes constant mean and variance. All other models are described in the caption to Table 1. A positive number suggests that the model with the predictor variables listed in the row has a lower RMSE relative to the PMV model, while a negative number suggests the opposite. All results are out-of-sample and cover the period 1956:01-2010:12. Forecasts are obtained by estimating the models recursively using a rolling window with the most recent 30 years of observations.

Table 3: Predictive Density Comparisons

	TVM	TVM-EGARCH	PM-EGARCHX	TVM-EGARCHX
d/e	-0.0016	0.0476***	0.0474**	0.0503**
svar	0.0015	0.0472**	0.0734***	0.0697***
dfr	-0.0030	0.0330	0.0466	0.0426
lty	-0.0050	0.0284	0.0347*	0.0355*
ltr	0.0024	0.0542**	0.0424*	0.0307
infl	0.0037	0.0540***	0.0595***	0.0666***
tms	-0.0008	0.0464**	0.0446**	0.0278
tbl	-0.0041	0.0317	0.0486**	0.0426*
dfy	-0.0036	0.0480**	0.0538**	0.0443*
d/p	0.0004	0.0365	0.0391	0.0389
d/y	0.0005	0.0571***	0.0361	0.0356
e/p	-0.0027	0.0433**	0.0317	0.0311
b/m	-0.0033	0.0333	0.0372	0.0248
e10/p	-0.0016	0.0408	0.0441*	0.0367
ntis	-0.0006	0.0687***	0.0512***	0.0674***

Note: This table presents the Weighted Likelihood Ratio test statistic of Amisano and Giacomini (2007) for pairwise comparison of the performance of two density forecasts. The models with the lagged predictor variables listed in the rows are compared against the PMV model. A positive number suggests that the model with the predictor variable listed in the row provides a better density forecast than the PMV model, while a negative number suggests the reverse. All results are out-of-sample and cover the period 1956:01-2010:12. Forecasts are obtained by estimating the models recursively using a rolling window with the most recent 30 years of observations. The individual models are described in the caption to Table 1. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 4: Risk-Adjusted Relative Return Performance for an Investor with Mean Variance Utility ($w_t \in [0, 0.99]$)

(a) $\gamma = 3$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.173%	-0.023	0.088%	0.797%	0.011	0.399%	1.560%	0.012	0.198%	0.888%	0.023	0.713%
svar	-0.333%	-0.014	0.111%	1.748%	0.007	0.254%	0.781%	0.039	0.573%	2.023%	-0.015	-0.155%
dfr	0.365%	0.033	0.623%	1.850%	0.032	0.647%	1.417%	0.048	0.700%**	1.838%	0.063	1.041%**
lty	-2.572%	-0.017	0.406%	-1.914%	0.072	1.367%	1.061%	0.013	0.201%	-2.246%	-0.014	0.443%
ltr	1.558%	0.052	1.030%	2.446%	0.071	1.271%*	1.151%	0.006	0.101%	2.389%	0.075	1.339%*
infl	1.028%	0.108	1.679%**	1.588%	0.043	0.812%	1.691%	0.027	0.393%*	2.042%	0.103	1.649%**
tms	2.311%	0.103	1.604%**	2.819%	0.113	1.739%**	1.401%	0.010	0.155%	3.403%	0.144	2.121%**
tbl	-1.770%	0.024	0.842%	-1.525%	0.042	1.050%	1.529%	0.016	0.255%	-1.451%	0.041	1.075%
dfy	0.646%	-0.089	-0.898%	2.006%	-0.082	-0.842%	0.615%	0.017	0.261%	1.814%	-0.058	-0.608%
d/p	-1.651%	-0.086	-0.629%	-0.306%	-0.069	-0.530%	0.951%	0.005	0.089%	-0.491%	-0.018	-0.036%
d/y	-1.610%	-0.095	-0.627%	-0.323%	-0.045	-0.180%	0.951%	0.007	0.115%	-0.553%	-0.032	-0.172%
e/p	-0.846%	-0.035	-0.236%	0.541%	0.023	0.503%	1.183%	0.025	0.355%	0.127%	0.016	0.392%
b/m	-0.467%	-0.082	-0.812%**	0.872%	-0.070	-0.769%	1.185%	0.020	0.291%	0.786%	0.001	0.080%
e10/p	-1.051%	-0.129	-1.076%	0.110%	-0.069	-0.613%	0.958%	0.007	0.123%	-0.305%	-0.049	-0.469%
ntis	1.888%	0.003	0.192%	3.076%	0.033	0.617%*	0.923%	-0.020	-0.227%	3.101%	0.019	0.408%

(b) $\gamma = 5$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.068%	0.028	0.643%	1.124%	0.046	0.782%	1.623%	0.041	0.496%	1.330%	0.085	1.357%**
svar	-0.184%	-0.005	0.251%	1.842%	0.022	0.393%	1.031%	0.068	0.870%*	1.999%	0.028	0.364%
dfr	0.459%	0.034	0.591%	2.059%	0.042	0.699%*	1.540%	0.078	0.946%**	2.093%	0.066	0.953%**
lty	-1.789%	0.002	0.590%	-1.129%	0.076	1.284%	1.210%	0.022	0.283%	-1.388%	-0.001	0.650%
ltr	1.672%	0.118	1.680%**	2.785%	0.152	2.047%***	1.241%	0.032	0.410%	2.727%	0.151	2.059%***
infl	1.062%	0.183	2.307%**	1.846%	0.106	1.474%**	1.705%	0.053	0.616%*	2.373%	0.171	2.311%***
tms	2.447%	0.151	1.957%**	3.037%	0.186	2.411%**	1.559%	0.023	0.281%	3.720%	0.217	2.822%***
tbl	-1.179%	0.043	0.980%	-0.815%	0.069	1.266%	1.706%	0.039	0.473%	-0.660%	0.081	1.420%*
dfy	0.892%	-0.063	-0.399%	2.268%	-0.021	-0.009%	0.872%	0.030	0.430%	2.020%	0.009	0.282%
d/p	-0.841%	-0.051	0.010%	0.140%	-0.015	0.237%	1.153%	0.017	0.251%	-0.056%	0.039	0.659%
d/y	-0.803%	-0.053	0.074%	0.150%	0.039	0.810%	1.152%	0.019	0.273%	-0.071%	0.020	0.504%
e/p	-0.387%	-0.038	-0.093%	0.780%	0.026	0.529%	1.245%	0.045	0.532%	0.330%	0.018	0.389%
b/m	0.140%	-0.080	-0.512%	1.230%	-0.066	-0.518%	1.276%	0.035	0.442%	0.999%	-0.007	0.008%
e10/p	-0.203%	-0.073	-0.167%	0.659%	-0.018	0.162%	1.169%	0.020	0.288%	0.193%	-0.008	0.163%
ntis	1.906%	0.056	0.824%*	3.191%	0.116	1.536%***	1.127%	0.004	0.152%	3.222%	0.109	1.451%***

(c) $\gamma = 10$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.034%	0.038	0.401%	0.866%	0.037	0.502%	1.125%	0.038	0.349%	1.077%	0.057	0.759%
svar	-0.077%	-0.021	0.068%	1.255%	0.001	0.137%	0.746%	0.046	0.465%	1.285%	0.012	0.156%
dfr	0.381%	-0.037	-0.025%	1.579%	-0.021	0.084%	1.106%	0.042	0.387%	1.651%	0.034	0.467%
lty	-0.913%	0.013	0.363%	-0.507%	0.022	0.530%	0.889%	-0.031	-0.123%	-0.632%	-0.051	0.202%
ltr	1.328%	0.103	1.194%**	2.368%	0.127	1.352%***	0.825%	0.036	0.326%	2.304%	0.142	1.480%***
infl	0.643%	0.162	1.277%**	1.406%	0.122	1.155%**	1.133%	0.041	0.331%	1.860%	0.184	1.785%**
tms	1.916%	0.153	1.515%**	2.346%	0.168	1.677%**	1.145%	0.018	0.212%	3.014%	0.198	2.046%***
tbl	-0.606%	0.052	0.546%	-0.356%	0.051	0.694%	1.311%	0.026	0.276%	-0.221%	0.038	0.688%
dfy	0.688%	-0.029	0.049%	1.718%	0.029	0.452%	0.619%	0.018	0.230%	1.473%	0.008	0.230%
d/p	-0.324%	-0.044	0.153%	0.195%	-0.041	0.046%	0.898%	0.028	0.302%	0.070%	0.017	0.325%
d/y	-0.294%	-0.032	0.260%	0.239%	0.030	0.526%	0.895%	0.039	0.379%	0.088%	0.002	0.259%
e/p	-0.109%	-0.076	-0.202%	0.543%	-0.024	0.042%	0.838%	0.024	0.233%	0.227%	-0.062	-0.271%
b/m	0.311%	-0.106	-0.423%*	0.950%	-0.098	-0.493%*	0.870%	0.038	0.352%	0.694%	-0.048	-0.249%
e10/p	0.219%	-0.094	-0.149%	0.643%	-0.023	0.174%	0.910%	0.023	0.274%	0.270%	-0.032	0.027%
ntis	1.287%	0.049	0.579%	2.314%	0.101	0.997%*	0.809%	-0.003	0.120%	2.376%	0.097	0.987%*

Note: This table reports the relative return performance (in percent per annum) of portfolios based on the models with the lagged predictor variables listed in the rows versus those based on the PMV model. The investor has a mean variance utility with a risk aversion parameter, γ of 3, 5 and 10. CER and RAR denote certainty equivalent return and risk-adjusted return, respectively. The risk-adjusted annual return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from the models with the lagged predictor variable listed in the row and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). A positive number suggests that the model with the lagged predictor variable listed in the row produces higher risk-adjusted returns than the PMV model. All results cover the period 1956:01-2010:12. The out-of-sample results are obtained by estimating the models recursively using a rolling window of the most recent 30 years of observations. The individual models are described in the caption to Table 1. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 5: Risk-Adjusted Relative Return Performance for an Investor with CRRA Utility ($w_t \in [0, 0.99]$)

(a) $\gamma = 3$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.357%	-0.026	0.042%	0.079%	0.008	0.353%	0.139%	0.011	0.175%	0.204%	0.020	0.672%
svar	-0.188%	-0.015	0.095%	0.065%	0.005	0.226%	0.475%	0.035	0.513%	-0.278%	-0.019	-0.228%
dfr	0.433%	0.031	0.590%	0.387%	0.029	0.613%	0.618%	0.045	0.656%**	0.800%	0.060	1.000%*
lty	-0.370%	-0.019	0.372%	0.587%	0.070	1.333%	0.131%	0.010	0.162%	-0.342%	-0.017	0.407%
ltr	0.644%	0.049	0.989%	0.852%	0.066	1.205%*	0.045%	0.004	0.076%	0.904%	0.070	1.263%*
infl	1.305%	0.104	1.625%**	0.474%	0.039	0.761%	0.339%	0.025	0.360%*	1.219%	0.099	1.599%**
tms	1.258%	0.100	1.560%**	1.346%	0.110	1.690%**	0.102%	0.008	0.132%	1.729%	0.140	2.072%**
tbl	0.102%	0.022	0.808%	0.272%	0.040	1.022%	0.180%	0.014	0.220%	0.267%	0.038	1.034%
dfy	-1.155%	-0.091	-0.936%	-1.117%	-0.085	-0.901%	0.183%	0.014	0.223%	-0.801%	-0.062	-0.665%
d/p	-1.019%	-0.087	-0.661%	-0.863%	-0.072	-0.578%	0.014%	0.003	0.055%	-0.227%	-0.020	-0.076%
d/y	-1.134%	-0.097	-0.663%	-0.555%	-0.047	-0.225%	0.038%	0.004	0.079%	-0.396%	-0.034	-0.207%
e/p	-0.440%	-0.036	-0.254%	0.287%	0.020	0.466%	0.288%	0.022	0.311%	0.189%	0.013	0.351%
b/m	-1.029%	-0.083	-0.835%*	-0.949%	-0.073	-0.817%	0.214%	0.017	0.251%	-0.007%	-0.001	0.047%
e10/p	-1.528%	-0.131	-1.122%	-0.859%	-0.072	-0.660%	0.044%	0.005	0.085%	-0.644%	-0.051	-0.502%
ntis	0.022%	0.000	0.151%	0.421%	0.030	0.573%	-0.308%	-0.023	-0.266%	0.219%	0.015	0.348%

(b) $\gamma = 5$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	0.390%	0.026	0.631%	0.399%	0.045	0.771%	0.176%	0.040	0.486%	0.851%	0.084	1.353%*
svar	0.148%	-0.005	0.247%	-0.085%	0.022	0.398%	0.725%	0.066	0.841%*	-0.025%	0.024	0.304%
dfr	0.429%	0.034	0.598%	0.343%	0.043	0.710%	0.605%	0.078	0.952%**	0.552%	0.068	0.980%*
lty	0.401%	0.001	0.588%	0.891%	0.076	1.299%	-0.009%	0.022	0.292%	0.344%	-0.002	0.650%
ltr	1.335%	0.116	1.675%**	1.682%	0.152	2.065%***	0.110%	0.031	0.399%	1.680%	0.152	2.078%***
infl	1.928%	0.181	2.309%***	1.045%	0.105	1.463%*	0.198%	0.051	0.598%*	1.764%	0.168	2.296%**
tms	1.571%	0.150	1.960%**	1.953%	0.186	2.425%**	-0.047%	0.021	0.265%	2.269%	0.215	2.820%***
tbl	0.642%	0.042	0.984%	0.852%	0.069	1.273%	0.051%	0.038	0.471%	0.948%	0.081	1.430%
dfy	-0.724%	-0.063	-0.398%	-0.517%	-0.022	-0.020%	0.186%	0.031	0.436%	-0.172%	0.007	0.265%
d/p	-0.201%	-0.052	-0.002%	0.010%	-0.016	0.221%	-0.040%	0.016	0.244%	0.494%	0.036	0.632%
d/y	-0.189%	-0.054	0.067%	0.589%	0.038	0.801%	-0.020%	0.018	0.265%	0.319%	0.017	0.475%
e/p	-0.183%	-0.038	-0.097%	0.233%	0.026	0.532%	0.250%	0.045	0.529%	0.049%	0.018	0.394%
b/m	-0.706%	-0.079	-0.513%	-0.775%	-0.066	-0.516%	0.147%	0.035	0.443%	-0.347%	-0.005	0.034%
e10/p	-0.398%	-0.074	-0.177%	-0.026%	-0.019	0.155%	0.009%	0.019	0.281%	0.020%	-0.009	0.161%
ntis	0.402%	0.055	0.819%*	1.010%	0.113	1.514%***	-0.024%	0.005	0.156%	0.965%	0.107	1.432%***

(c) $\gamma = 10$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	0.303%	0.039	0.413%	-0.169%	0.039	0.516%	-0.332%	0.040	0.359%	-0.253%	0.059	0.778%*
svar	-0.022%	-0.022	0.067%	-0.808%	-0.001	0.119%	-0.023%	0.045	0.458%	-0.605%	0.010	0.137%
dfr	-0.552%	-0.036	-0.024%	-1.042%	-0.019	0.090%	-0.472%	0.043	0.393%	-0.651%	0.035	0.471%
lty	0.356%	0.014	0.377%	0.174%	0.023	0.545%	-0.864%	-0.028	-0.114%	-0.043%	-0.049	0.216%
ltr	0.331%	0.105	1.213%**	0.299%	0.130	1.386%***	-0.270%	0.036	0.325%	0.423%	0.144	1.500%***
infl	0.723%	0.164	1.314%**	0.292%	0.124	1.184%**	-0.370%	0.042	0.341%	0.566%	0.185	1.815%***
tms	0.116%	0.157	1.556%**	0.058%	0.170	1.704%**	-0.548%	0.019	0.219%	0.089%	0.203	2.099%**
tbl	0.456%	0.053	0.565%	0.341%	0.052	0.714%	-0.667%	0.027	0.282%	0.174%	0.041	0.709%
dfy	-0.516%	-0.029	0.048%	-0.744%	0.028	0.447%	-0.304%	0.019	0.236%	-0.783%	0.009	0.235%
d/p	-0.415%	-0.045	0.152%	-0.445%	-0.041	0.049%	-0.319%	0.029	0.308%	-0.097%	0.017	0.336%
d/y	-0.262%	-0.031	0.272%	0.027%	0.029	0.528%	-0.252%	0.040	0.386%	-0.164%	0.003	0.271%
e/p	-0.747%	-0.077	-0.211%	-0.596%	-0.024	0.041%	-0.360%	0.026	0.248%	-0.827%	-0.061	-0.273%
b/m	-1.334%	-0.105	-0.427%	-1.397%	-0.097	-0.499%*	-0.278%	0.039	0.359%	-1.092%	-0.049	-0.259%
e10/p	-1.103%	-0.094	-0.151%	-0.500%	-0.022	0.182%	-0.349%	0.024	0.286%	-0.450%	-0.030	0.042%
ntis	-0.405%	0.050	0.599%	-0.435%	0.102	1.007%*	-0.445%	-0.002	0.125%	-0.522%	0.097	0.996%*

Note: This table reports the relative return performance (in percent per annum) of portfolios based on the models with the lagged predictor variables listed in the rows versus those based on the PMV model. The investor has a constant relative risk aversion utility with a risk aversion coefficient γ of 3, 5 and 10. CER and RAR denote certainty equivalent return and risk-adjusted return, respectively. The risk-adjusted annual return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from the models with the lagged predictor variable listed in the row and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). A positive number suggests that the model with the lagged predictor variable listed in the row produces higher risk-adjusted returns than the PMV model. All results cover the period 1956:01-2010:12. The out-of-sample results are obtained by estimating the models recursively using a rolling window of the most recent 30 years of observations. The individual models are described in the caption to Table 1. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 6: Risk-Adjusted Relative Return Performance for an Investor with Mean Variance Utility ($w_t \in [0, 1.50]$)

(a) $\gamma = 3$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.121%	0.016	0.821%	1.660%	0.037	1.075%	2.484%	0.035	0.683%	1.945%	0.072	1.885%*
svar	-0.311%	-0.006	0.357%	2.821%	0.018	0.571%	1.531%	0.062	1.256%*	3.093%	0.018	0.402%
dfr	0.687%	0.034	0.928%	3.116%	0.041	1.081%*	2.348%	0.075	1.453%**	3.155%	0.069	1.567%**
lty	-2.935%	-0.006	0.810%	-1.912%	0.072	1.938%	1.827%	0.019	0.406%	-2.330%	-0.005	0.932%
ltr	2.539%	0.104	2.402%**	4.178%	0.144	3.088%***	1.901%	0.022	0.467%	4.095%	0.144	3.104%**
infl	1.647%	0.172	3.444%***	2.780%	0.095	2.101%*	2.634%	0.042	0.785%	3.568%	0.155	3.327%***
tms	3.707%	0.140	2.867%**	4.603%	0.174	3.547%**	2.364%	0.013	0.276%	5.616%	0.202	4.115%***
tbl	-1.932%	0.036	1.453%	-1.404%	0.062	1.866%	2.586%	0.036	0.705%	-1.182%	0.075	2.133%
dfy	1.323%	-0.069	-0.747%	3.426%	-0.027	-0.111%	1.294%	0.028	0.645%	3.064%	-0.001	0.274%
d/p	-1.477%	-0.060	-0.176%	0.123%	-0.030	0.102%	1.727%	0.010	0.265%	-0.175%	0.021	0.725%
d/y	-1.419%	-0.060	-0.045%	0.128%	0.025	1.035%	1.725%	0.011	0.293%	-0.212%	0.005	0.520%
e/p	-0.713%	-0.041	-0.243%	1.164%	0.023	0.779%	1.907%	0.039	0.739%	0.478%	0.020	0.648%
b/m	0.071%	-0.082	-0.878%*	1.811%	-0.066	-0.832%	1.949%	0.031	0.619%	1.493%	-0.002	0.109%
e10/p	-0.518%	-0.083	-0.471%	0.884%	-0.027	0.076%	1.748%	0.016	0.377%	0.186%	-0.019	0.052%
ntis	2.938%	0.045	1.097%*	4.876%	0.100	2.122%**	1.695%	0.002	0.192%	4.922%	0.092	1.958%***

(b) $\gamma = 5$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.068%	0.039	0.801%	1.573%	0.052	1.065%	2.096%	0.039	0.617%	1.932%	0.064	1.439%*
svar	-0.164%	-0.013	0.213%	2.335%	0.003	0.239%	1.402%	0.049	0.872%	2.453%	0.023	0.411%
dfr	0.666%	-0.013	0.172%	2.803%	0.010	0.479%	2.015%	0.049	0.766%	2.897%	0.039	0.838%
lty	-1.825%	0.018	0.726%	-1.041%	0.039	1.148%	1.634%	-0.012	-0.061%	-1.311%	-0.032	0.510%
ltr	2.315%	0.114	2.098%**	4.018%	0.152	2.592%***	1.561%	0.030	0.497%	3.915%	0.153	2.625%***
infl	1.260%	0.177	2.618%**	2.528%	0.133	2.195%**	2.121%	0.050	0.712%	3.287%	0.182	3.092%***
tms	3.382%	0.169	2.777%**	4.160%	0.175	2.942%**	2.086%	0.023	0.400%	5.223%	0.224	3.806%***
tbl	-1.213%	0.055	1.093%	-0.725%	0.070	1.484%	2.331%	0.030	0.487%	-0.489%	0.056	1.454%
dfy	1.256%	-0.047	-0.213%	3.085%	0.009	0.452%	1.171%	0.021	0.448%	2.688%	0.015	0.466%
d/p	-0.669%	-0.055	0.087%	0.346%	-0.031	0.161%	1.605%	0.025	0.459%	0.105%	0.040	0.837%
d/y	-0.617%	-0.042	0.323%	0.400%	0.032	0.955%	1.601%	0.032	0.559%	0.126%	0.023	0.690%
e/p	-0.235%	-0.057	-0.217%	1.029%	-0.003	0.286%	1.580%	0.038	0.585%	0.439%	-0.037	-0.238%
b/m	0.523%	-0.094	-0.705%*	1.751%	-0.093	-0.919%*	1.633%	0.039	0.630%	1.321%	-0.041	-0.428%
e10/p	0.262%	-0.087	-0.271%	1.138%	-0.019	0.278%	1.632%	0.025	0.479%	0.473%	-0.016	0.180%
ntis	2.394%	0.059	1.126%	4.192%	0.105	1.814%**	1.501%	-0.003	0.160%	4.263%	0.101	1.764%**

(c) $\gamma = 10$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.034%	0.038	0.401%	0.915%	0.015	0.415%	1.172%	0.020	0.268%	1.149%	0.041	0.735%
svar	-0.073%	-0.034	0.011%	1.313%	-0.017	0.049%	0.756%	0.037	0.425%	1.304%	-0.006	0.048%
dfr	0.410%	-0.071	-0.197%	1.734%	-0.037	0.052%	1.184%	0.034	0.394%	1.844%	-0.004	0.311%
lty	-0.913%	0.013	0.363%	-0.504%	0.019	0.528%	0.917%	-0.052	-0.241%	-0.614%	-0.065	0.175%
ltr	1.474%	0.113	1.464%***	2.710%	0.095	1.417%**	0.846%	0.037	0.372%	2.648%	0.104	1.525%**
infl	0.643%	0.162	1.278%**	1.511%	0.109	1.180%**	1.181%	0.029	0.285%	2.053%	0.180	2.029%***
tms	2.063%	0.148	1.689%**	2.544%	0.163	1.865%**	1.208%	0.002	0.159%	3.376%	0.151	2.063%**
tbl	-0.606%	0.052	0.546%	-0.356%	0.048	0.682%	1.425%	0.002	0.193%	-0.214%	0.032	0.681%
dfy	0.702%	-0.030	0.066%	1.843%	0.031	0.578%	0.624%	0.017	0.235%	1.562%	0.006	0.282%
d/p	-0.324%	-0.044	0.153%	0.207%	-0.034	0.095%	0.970%	0.004	0.196%	0.074%	0.002	0.254%
d/y	-0.294%	-0.033	0.256%	0.263%	0.034	0.572%	0.966%	0.015	0.278%	0.097%	-0.017	0.168%
e/p	-0.109%	-0.076	-0.202%	0.547%	-0.028	0.023%	0.860%	0.003	0.121%	0.228%	-0.067	-0.303%
b/m	0.318%	-0.127	-0.577%	0.980%	-0.104	-0.526%*	0.897%	0.031	0.346%	0.707%	-0.058	-0.295%
e10/p	0.241%	-0.097	-0.095%	0.660%	-0.033	0.150%	0.978%	0.006	0.220%	0.277%	-0.040	-0.001%
ntis	1.312%	0.056	0.668%	2.485%	0.106	1.175%*	0.832%	-0.007	0.137%	2.584%	0.113	1.337%*

Note: This table reports the relative return performance (in percent per annum) of portfolios based on the models with the lagged predictor variables listed in the rows versus those based on the PMV model. The investor has a mean variance utility with a risk aversion parameter, γ of 3, 5 and 10, and is allowed to use leverage so that $w_t \in [0, 1.50]$. CER and RAR denote certainty equivalent return and risk-adjusted return, respectively. The risk-adjusted annual return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from the models with the lagged predictor variable listed in the row and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). A positive number suggests that the model with the lagged predictor variable listed in the row produces higher risk-adjusted returns than the PMV model. All results cover the period 1956:01-2010:12. The out-of-sample results are obtained by estimating the models recursively using a rolling window of the most recent 30 years of observations. The individual models are described in the caption to Table 1. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 7: Risk-Adjusted Relative Return Performance for an Investor with Mean Variance Utility ($w_t \in [0, 1.50]$ imposing a non-negative ex-ante equity risk premium)

(a) $\gamma = 3$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.112%	0.026	0.859%	1.664%	0.041	1.069%	2.484%	0.035	0.683%	1.962%	0.081	1.886%**
svar	-0.307%	-0.003	0.313%	2.825%	0.016	0.481%	1.531%	0.062	1.256%*	3.094%	0.020	0.435%
dfr	0.690%	0.036	0.915%	3.124%	0.041	1.023%*	2.348%	0.075	1.453%**	3.162%	0.066	1.453%**
lty	-2.905%	0.006	0.683%	-1.892%	0.075	1.717%	1.827%	0.019	0.406%	-2.304%	0.009	0.836%
ltr	2.548%	0.102	2.232%***	4.188%	0.135	2.818%**	1.901%	0.022	0.467%	4.105%	0.135	2.840%**
infl	1.654%	0.168	3.285%***	2.787%	0.094	2.000%*	2.634%	0.042	0.785%	3.579%	0.156	3.233%**
tms	3.713%	0.135	2.695%**	4.611%	0.165	3.294%**	2.364%	0.013	0.276%	5.624%	0.192	3.842%***
tbl	-1.909%	0.042	1.288%	-1.385%	0.064	1.631%	2.586%	0.035	0.702%	-1.155%	0.084	1.978%
dfy	1.331%	-0.056	-0.578%	3.434%	-0.013	0.093%	1.294%	0.028	0.645%	3.069%	0.006	0.375%
d/p	-1.459%	-0.051	-0.099%	0.131%	-0.015	0.277%	1.727%	0.010	0.265%	-0.171%	0.029	0.832%
d/y	-1.386%	-0.051	-0.036%	0.137%	0.040	1.177%	1.725%	0.011	0.293%	-0.207%	0.013	0.642%
e/p	-0.706%	-0.033	-0.164%	1.170%	0.028	0.821%	1.907%	0.039	0.739%	0.481%	0.023	0.676%
b/m	0.090%	-0.077	-0.828%	1.815%	-0.062	-0.772%	1.949%	0.031	0.619%	1.494%	-0.002	0.094%
e10/p	-0.479%	-0.066	-0.319%	0.893%	-0.013	0.248%	1.748%	0.016	0.377%	0.191%	-0.013	0.126%
ntis	2.943%	0.052	1.202%*	4.882%	0.106	2.199%***	1.695%	0.002	0.192%	4.926%	0.098	2.038%***

(b) $\gamma = 5$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.063%	0.049	0.824%	1.575%	0.055	1.061%	2.096%	0.039	0.617%	1.942%	0.071	1.439%
svar	-0.161%	-0.009	0.186%	2.338%	0.001	0.185%	1.402%	0.049	0.872%	2.453%	0.025	0.431%
dfr	0.668%	-0.011	0.163%	2.808%	0.011	0.445%	2.015%	0.049	0.766%	2.901%	0.037	0.770%
lty	-1.807%	0.030	0.650%	-1.029%	0.041	1.014%	1.634%	-0.012	-0.061%	-1.295%	-0.019	0.452%
ltr	2.320%	0.112	1.996%**	4.024%	0.145	2.430%**	1.561%	0.030	0.497%	3.921%	0.146	2.467%***
infl	1.264%	0.173	2.522%**	2.532%	0.132	2.134%**	2.121%	0.050	0.712%	3.294%	0.183	3.036%***
tms	3.386%	0.166	2.674%**	4.165%	0.169	2.790%**	2.086%	0.023	0.400%	5.228%	0.217	3.642%***
tbl	-1.199%	0.061	0.993%	-0.714%	0.072	1.343%	2.331%	0.030	0.485%	-0.473%	0.064	1.361%
dfy	1.261%	-0.036	-0.112%	3.089%	0.019	0.575%	1.171%	0.021	0.448%	2.691%	0.020	0.526%
d/p	-0.659%	-0.047	0.133%	0.351%	-0.019	0.266%	1.605%	0.025	0.459%	0.108%	0.047	0.901%
d/y	-0.597%	-0.035	0.329%	0.406%	0.044	1.041%	1.601%	0.032	0.559%	0.129%	0.030	0.764%
e/p	-0.231%	-0.051	-0.169%	1.033%	0.001	0.311%	1.580%	0.038	0.585%	0.441%	-0.035	-0.221%
b/m	0.534%	-0.090	-0.675%*	1.754%	-0.089	-0.883%*	1.633%	0.039	0.630%	1.322%	-0.042	-0.437%
e10/p	0.285%	-0.074	-0.181%	1.144%	-0.009	0.381%	1.632%	0.025	0.479%	0.476%	-0.011	0.225%
ntis	2.398%	0.064	1.189%	4.195%	0.110	1.860%**	1.501%	-0.003	0.160%	4.265%	0.105	1.811%*

(c) $\gamma = 10$

	TVM			TVM-EGARCH			PM-EGARCHX			TVM-EGARCHX		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
d/e	-0.031%	0.047	0.412%	0.916%	0.018	0.413%	1.172%	0.020	0.268%	1.155%	0.047	0.736%*
svar	-0.071%	-0.031	-0.002%	1.314%	-0.018	0.022%	0.756%	0.037	0.425%	1.304%	-0.004	0.058%
dfr	0.411%	-0.070	-0.201%	1.736%	-0.037	0.035%	1.184%	0.034	0.394%	1.846%	-0.006	0.277%
lty	-0.904%	0.025	0.325%	-0.498%	0.022	0.461%	0.917%	-0.052	-0.241%	-0.606%	-0.053	0.147%
ltr	1.477%	0.111	1.413%**	2.713%	0.090	1.336%**	0.846%	0.037	0.372%	2.651%	0.099	1.446%**
infl	0.645%	0.159	1.230%**	1.513%	0.109	1.150%**	1.181%	0.029	0.285%	2.056%	0.181	2.001%***
tms	2.065%	0.146	1.638%**	2.547%	0.158	1.789%**	1.208%	0.002	0.159%	3.379%	0.146	1.981%**
tbl	-0.599%	0.058	0.497%	-0.350%	0.050	0.612%	1.425%	0.002	0.192%	-0.206%	0.039	0.634%
dfy	0.704%	-0.021	0.117%	1.845%	0.040	0.639%	0.624%	0.017	0.235%	1.563%	0.011	0.313%
d/p	-0.318%	-0.038	0.176%	0.210%	-0.023	0.148%	0.970%	0.004	0.196%	0.075%	0.008	0.286%
d/y	-0.284%	-0.025	0.259%	0.266%	0.045	0.615%	0.966%	0.015	0.278%	0.098%	-0.010	0.205%
e/p	-0.107%	-0.071	-0.178%	0.549%	-0.025	0.035%	0.860%	0.003	0.121%	0.229%	-0.065	-0.295%
b/m	0.323%	-0.124	-0.562%	0.982%	-0.101	-0.508%*	0.897%	0.031	0.346%	0.708%	-0.058	-0.299%
e10/p	0.253%	-0.087	-0.050%	0.663%	-0.024	0.201%	0.978%	0.006	0.220%	0.278%	-0.036	0.021%
ntis	1.314%	0.061	0.699%	2.487%	0.110	1.199%*	0.832%	-0.007	0.137%	2.586%	0.116	1.361%*

Note: This table reports the relative return performance (in percent per annum) of portfolios based on the models with the lagged predictor variables listed in the rows versus those based on the PMV model. The investor has a mean variance utility with a risk aversion parameter, γ of 3, 5 and 10, and is allowed to use leverage so that $w_t \in [0, 1.50]$. The investor also imposes the economic constraint that ex-ante equity risk premium is non-negative as in Campbell and Thompson (2008). CER and RAR denote certainty equivalent return and risk-adjusted return, respectively. The risk-adjusted annual return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from the models with the lagged predictor variable listed in the row and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). A positive number suggests that the model with the lagged predictor variable listed in the row produces higher risk-adjusted returns than the PMV model. All results cover the period 1956:01-2010:12. The out-of-sample results are obtained by estimating the models recursively using a rolling window of the most recent 30 years of observations. The individual models are described in the caption to Table 1. * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 8: Disagreement between Statistical and Economic Measures of Performance

(a) Mean Variance Utility ($w_t \in [0, 0.99]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
RMSE	59%	54%	61%	66%	62%	72%	67%	54%	69%
Density Test	7%	7%	10%	8%	10%	10%	10%	8%	10%

(b) CRRA Utility ($w_t \in [0, 0.99]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
RMSE	52%	52%	61%	56%	62%	72%	15%	52%	69%
Density Test	7%	7%	10%	10%	10%	10%	7%	8%	10%

(c) Mean Variance Utility ($w_t \in [0, 1.50]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
RMSE	66%	59%	72%	67%	56%	69%	67%	48%	67%
Whole Dist	8%	8%	10%	10%	8%	10%	10%	8%	10%

(d) Mean Variance Utility ($w_t \in [0, 1.50]$ imposing a non-negative ex-ante equity risk premium)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
RMSE	66%	64%	74%	67%	57%	69%	67%	48%	67%
Whole Dist	8%	10%	10%	10%	8%	10%	10%	8%	10%

Note: This table presents the percentage of 61 models that have worse statistical but better portfolio performance than the PMV model. CER is certainty equivalent return. The risk-adjusted return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from different models and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). RMSE is the root mean square error and Density Test is the Weighted Likelihood Ratio test statistic of Amisano and Giacomini (2007) for pairwise comparison of the performance of two density forecasts.

Table 9: Relation between Statistical and Economic Measures of Performance

(a) Mean Variance Utility ($w_t \in [0, 0.99]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
				RMSE					
Rank Corr	0.25	0.35	0.27	0.24	0.46	0.40	0.24	0.53	0.51
Linear Corr	0.06	0.23	0.24	0.07	0.30	0.30	0.10	0.36	0.36
R^2	0.00	0.00	0.01	0.06	0.09	0.13	0.06	0.09	0.13
				Density Comparison Test					
Rank Corr	0.47	0.21	0.13	0.50	0.33	0.23	0.49	0.34	0.27
Linear Corr	0.47	0.24	0.14	0.49	0.30	0.19	0.46	0.28	0.19
R^2	0.22	0.24	0.21	0.06	0.09	0.08	0.02	0.04	0.03

(b) CRRA Utility ($w_t \in [0, 0.99]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
				RMSE					
Rank Corr	0.38	0.34	0.27	0.43	0.45	0.39	0.52	0.53	0.52
Linear Corr	0.27	0.24	0.24	0.35	0.30	0.31	0.37	0.36	0.36
R^2	0.07	0.12	0.14	0.06	0.09	0.13	0.06	0.09	0.13
				Density Comparison Test					
Rank Corr	0.24	0.20	0.12	0.17	0.33	0.22	-0.03	0.34	0.26
Linear Corr	0.24	0.24	0.13	0.15	0.30	0.19	-0.04	0.27	0.18
R^2	0.06	0.02	0.00	0.06	0.09	0.08	0.02	0.03	0.03

(c) Mean Variance Utility ($w_t \in [0, 1.50]$)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
				RMSE					
Rank Corr	0.23	0.43	0.35	0.23	0.52	0.49	0.25	0.51	0.49
Linear Corr	0.06	0.29	0.29	0.08	0.34	0.34	0.11	0.37	0.36
R^2	0.00	0.01	0.01	0.08	0.12	0.13	0.08	0.12	0.13
				Density Comparison Test					
Rank Corr	0.50	0.32	0.20	0.49	0.34	0.26	0.49	0.32	0.27
Linear Corr	0.48	0.30	0.19	0.47	0.27	0.18	0.45	0.26	0.18
R^2	0.24	0.22	0.20	0.09	0.07	0.07	0.04	0.03	0.03

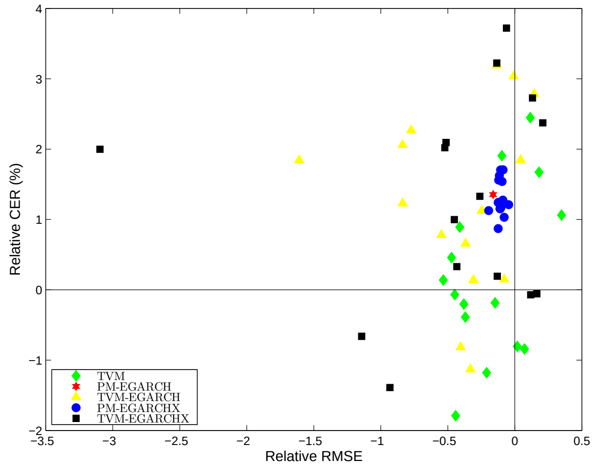
(d) Mean Variance Utility ($w_t \in [0, 1.50]$ imposing a non-negative ex-ante equity risk premium)

	$\gamma = 3$			$\gamma = 5$			$\gamma = 10$		
	CER	Sharpe	RAR	CER	Sharpe	RAR	CER	Sharpe	RAR
				RMSE					
Rank Corr	0.23	0.43	0.37	0.23	0.52	0.50	0.25	0.50	0.51
Linear Corr	0.06	0.29	0.30	0.08	0.34	0.35	0.11	0.36	0.37
R^2	0.00	0.01	0.01	0.08	0.12	0.13	0.09	0.12	0.14
				Density Comparison Test					
Rank Corr	0.50	0.31	0.23	0.49	0.32	0.28	0.48	0.31	0.28
Linear Corr	0.48	0.29	0.21	0.47	0.26	0.19	0.45	0.25	0.19
R^2	0.23	0.22	0.20	0.09	0.07	0.06	0.04	0.04	0.04

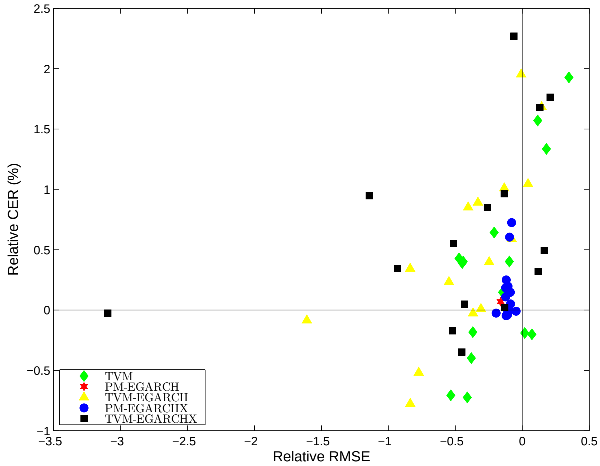
Note: This table presents linear and rank correlations between statistical and economic measures of performance as well as R^2s from regressions of economic measures on statistical measures. CER is certainty equivalent return. The risk-adjusted return (RAR) is calculated as the sample average of the stochastic discount factor times the difference between portfolio returns based on forecasts from different models and those from the portfolio based on the PMV model. The stochastic discount factor is based on the approach of Almeida and Garcia (2008). RMSE is the root mean square error and Density Test is the Weighted Likelihood Ratio test statistic of Amisano and Giacomini (2007) for pairwise comparison of the performance of two density forecasts.

Figure 1: Relation between RMSE and Measures of Economic Performance

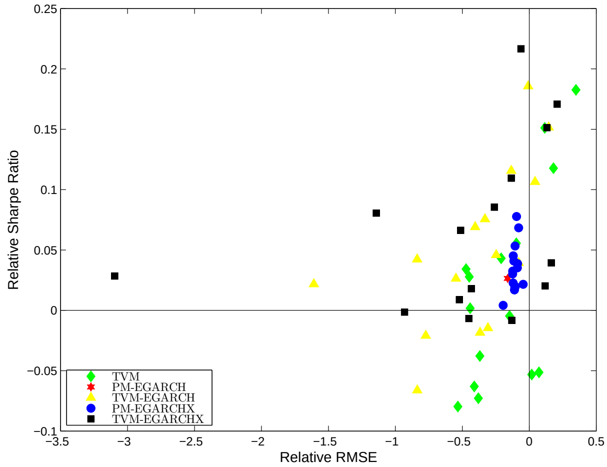
(a) CER (MV) vs. RMSE



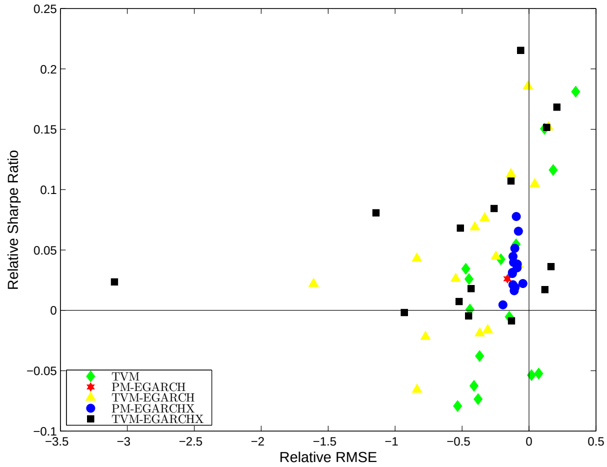
(b) CER (CRRA) vs. RMSE



(c) Sharpe Ratio (MV) vs. RMSE



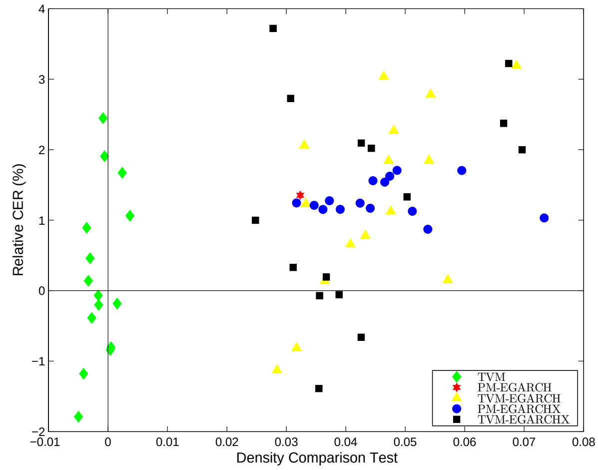
(d) Sharpe Ratio (CRRA) vs. RMSE



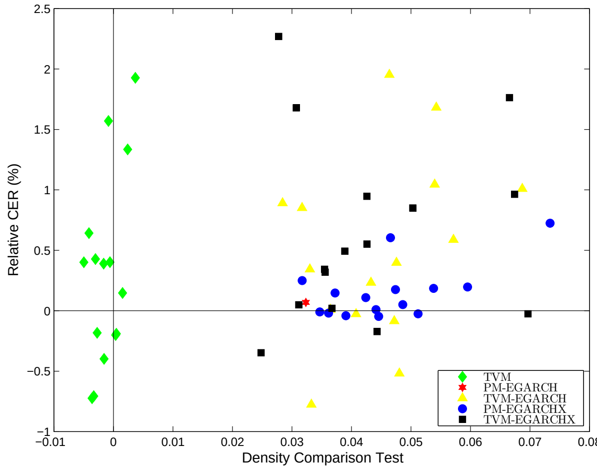
Note: This figure presents scatter plots of out-of-sample economic performance measures against RMSE (root mean square error). CER (certainty equivalent return) and Sharpe ratio are measured relative to those based on the PMV model so that the origin in the figure corresponds to the PMV model. The investor has a risk aversion parameter γ of 5 and either mean variance (MV) (panels a and c) or constant relative risk aversion (CRRA) utility (panels b and d). All results are computed out-of-sample and cover the period 1956:01-2010:12.

Figure 2: Relation between Predictive Density Test and Measures of Economic Performance

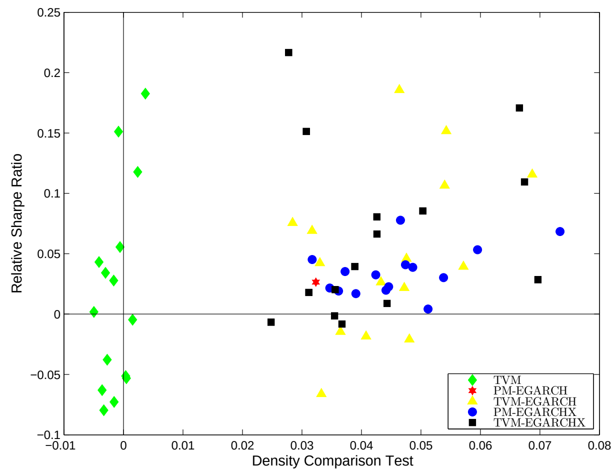
(a) CER (MV) vs. Pred. Dens. Test



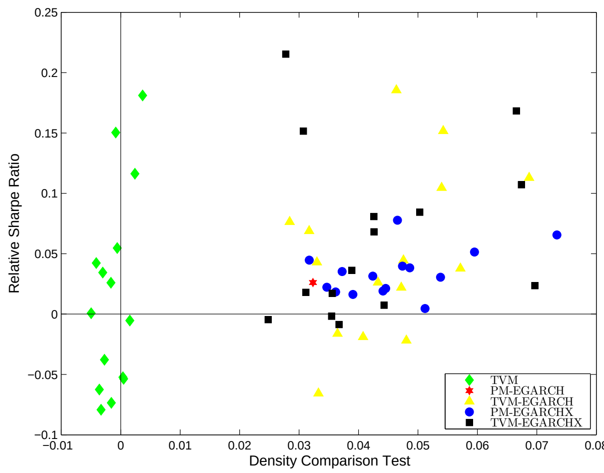
(b) CER (CRRA) vs. Pred. Dens. Test



(c) Sharpe (MV) vs. Pred. Dens. Test



(d) Sharpe (CRRA) vs. Pred. Dens. Test



Note: This figure presents scatter plots of out-of-sample economic performance measures against the predictive density test (the Weighted Likelihood Ratio test statistic of Amisano and Giacomini (2007) for the full return distribution). CER (certainty equivalent return) and Sharpe ratio are measured relative to those based on the PMV model so that the origin in the figure corresponds to the PMV model. The investor has a risk aversion parameter γ of 5 and either mean variance (MV) (panels a and c) or constant relative risk aversion (CRRA) utility (panels b and d). All results are computed out-of-sample and cover the period 1956:01-2010:12.