Empirical Finance: Methods & Applications

Asset Return Predictability

Pasquale Della Corte
Imperial College London

Week 3

Introduction

Asset return predictability is an exciting exercise

- Practitioners need asset return forecasts for their asset allocation, and better forecasts can boost investment performance,
- Academics need asset return forecasts to evaluate market efficiency tests and and produce better asset pricing models.

But predicting future asset returns is often challenging

- Asset returns contains a large unpredictable component, so even the best forecasting model can explain just a small portion of them,
- Also, competition among traders implies that successful forecasting models may be adopted by others, which can cause stock prices to move in a way that eliminates their superior predictive power.

Asset Return Predictability

Are asset return predictable in theory?

- Rational asset pricing models suggest that asset returns compensate investors for their exposure to aggregate (or undiversifiable) risk,
- Predictability can then result from exposure to time-varying aggregate risk, and forecasting models may capture this premium that varies over time.

But is predictability consistent with market efficiency?

- A market where prices 'fully reflect' available information is efficient (Fama, 1970),
- It is commonly believed that market efficiency means that asset returns are unpredictable,
- This is incorrect and asset return predictability is consistent with market efficiency as long as it reflects compensation for bearing risk.

Asset Return Predictability

There are at least three potential explanations for the return predictability.

Rational risk-based explanations

- Fluctuations in aggregate risk can produce time-varying discount rates and return predictability that is consistent with market efficiency,
- During recessions, for example, investors are more risk averse and demand higher expected returns to hold risk, and a variable that predicts the economy can then help predict returns,
- It is when risk-adjusted expected returns are non zero, we can then make a statement about market inefficiency.

Asset Return Predictability

Limits to arbitrage explanations

- Existence of trading frictions like lack of liquidity, borrowing constraints, transaction costs may give rise to security mispricing,

Behavioural biases explanations

- Investors may under-react or over-react to certain news because of psychological biases affecting their trading decision (e.g., limited attention)

Present Value Model

Introduction

What is a present value model?

- A popular framework to analyze asset returns,
- The key ingredients are expected future fundamentals and discount rates.

What does a present value model state?

- An asset price reflects its expected future fundamentals discounted to the present,
- The discount rate can be constant or time-varying.

Present Value Model

The traditional present value model defines today's asset price as

$$s_t = (1 - b)\, z_t + bE_t\left(s_{t+1}\right)$$

- s_t is the today's asset price,
- b_t is the subjective discount factor,
- z_t is the current fundamental,
- $E_t s_{t+1}$ is the expected future price.

Today's asset price reflect both current fundamentals and expected future fundamentals.

Present Value Model

The (subjective) discount factor b captures investors' impatience relative to fundamentals

$$s_t = (1 - b)\, z_t + bE_t\,(s_{t+1})$$

- A patient agent has higher $b \Longrightarrow$ higher weight on future fundamentals $E_t s_{t+1}$,
- An impatient agent has lower $b \Longrightarrow$ higher weight on the current fundamentals z_t.

Recall the law of iterated expectations

$$E_t\,(s_{t+2}) = E_t\,[E_{t+1}\,(s_{t+2})]$$

- At time t, an investor forecasts the future price s_{t+2} using her available information set,
- Our investor, however, will not generate a more precise forecast of s_{t+2} by conditioning on more specific information $\Longrightarrow E_t\,[E_{t+1}\,(s_{t+2})]$.

Present Value Model

Solve the model forward and impose the law of iterated expectations

$$s_t = (1-b)\sum_{j=0}^{k} b^j E_t\left(z_{t+j}\right) + b^k E_t\left(s_{t+k}\right)$$

Impose the no-bubble condition or no-Ponzi game, meaning that $b < 1$

$$\lim_{k\to\infty} b^k E_t\left(s_{t+k}\right) \to 0$$

The equilibrium price is

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t\left(z_{t+j}\right)$$

the expected present value of the future discounted fundamentals.

Engel and West (2005)

Consider both observable and unobservable fundamentals

$$s_t = (1-b)(z_t + u_t) + b\left[(z'_t + u'_t) + E_t(s_{t+1})\right]$$

Iterate forward, impose the no-bubble condition, and obtain

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t(z_{t+j} + u_{t+j}) + b\sum_{j=0}^{\infty} b^j E_t(z'_{t+j} + u'_{t+j})$$

- s_t is the (log) asset price at time t,
- b is the discount factor $(0 < b < 1)$,
- z_t and z'_t are the observed fundamentals at time t,
- u_t and u'_t are the unobserved fundamentals at time t,
- $E_t(\cdot)$ is the expectations operator as of time t.

A Present Value Model

The stock price model of Campbell and Shiller (1987) and West (1988)

$$s_t = b \sum_{j=0}^{\infty} b^j E_t \left(z'_{t+j} \right)$$

- s_t is the (log) stock price at time t,
- b is the discount factor $(0 < b < 1)$,
- z'_t is the dividend at time t,

Nested into Engel and West (2005) by setting everything else equal to zero.

A Present Value Model

The term structure model of Campbell and Shiller (1988)

$$s_t = (1 - b) \sum_{j=0}^{\infty} b^j E_t \left(z_{t+j} \right)$$

- s_t is the yield on a bond at time t,
- b is the discount factor $(0 < b < 1)$,
- z_t is the short-term rate at time t.

Nested into Engel and West (2005) by setting everything else equal to zero.

A Present Value Model

The **exchange rate model** of Frankel (1976) and Mussa (1976)

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t\left(z_{t+j}\right)$$

- s_t is the (log) nominal exchange rate at time t,
- b is the discount factor $(0 < b < 1)$,
- $b = \lambda/(1-\lambda)$ where λ is the semi-elasticity of money demand,
- z_t is the monetary fundamentals at time t,
- $z_t = (m_t - m_t^*) - (y_t - y_t^*)$.

Nested into Engel and West (2005) by setting everything else equal to zero.

Predictive Regressions

What is a Predictive Regression?

A predictive regression is a linear specification designed to assess the predictive power of past values of some economic or financial variable for the future values of another variable.

A univariate predictive regression is typically formulated as

$$y_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

- y_t is the asset return (in excess of the risk-free rate) between times t and $t-1$,
- x_{t-1} is the lagged predictor (e.g., dividend-price ratio) observed at time t,
- ε_t is an unpredictable shock with zero-mean and constant variance at time t,
- β measures the significance of x_t in predicting r_{t+1}.

What is a Predictive Regression?

The estimation of predictive regression is often challenging

- Predictors are highly persistent behaving like nearly nonstationary processes,
- Returns are generally noisier with fast mean-reverting dynamics,
- Shocks driving r_t and x_t often display a contemporaneous negatively correlation.

The joint presence of persistence and contemporaneous correlations

- Creates distortions to the least-squares estimates of β,
- Leads to spurious evidence of predictability.

Conventional in-sample analysis should be complemented with an out-of-sample assessment.

Out-of-Sample Predictability

How to assess the out-of-sample (OOS) predictive ability of a model?

Real-time information

- We should only use information available to an investor when the forecast is produced,
- This is not always possible, especially when we use macroeconomic predictors.

Benchmark model

- We need a competing model playing the role of reference model,
- We often use the random walk with drift as a benchmark model.

Out-of-Sample Predictability

How to assess the out-of-sample (OOS) predictive ability of a model?

Horse race analysis

- Estimate the model only using available information,
- Measure its forecasting ability in the subsequent period

Performance Evaluation

- Criteria to evaluate the performance of the competing model against the benchmark.

Out-of-Sample Predictability

Your benchmark model is defined as

$$y_t = \alpha + \varepsilon_t$$

Your competing model is defined as

$$y_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

We can recover the benchmark by simply setting $\beta = 0$ in the competing model.

Out-of-Sample Predictability

Suppose you have a dataset of T observations (e.g., monthly from 1976/01 to 2023/12)

- First, you estimate the model using an *in-sample* training period,
- Second, you generate a one-step ahead forecast for the next *out-of-sample* observation,
- Hence, you repeat these steps multiple times until you reach the end of the sample.

You must compare your competing model to a benchmark model using various criteria,

- The random walk with drift will act as a benchmark,
- This benchmark is nested into the competing model,
- You will run out-of-sample regressions for both models.

Out-of-Sample Predictability

We can generate out-of-sample forecasts using two methods

- Expanding window
- Rolling window

In both cases, we start with an initial in-sample period

- For example, 10 years of monthly observations (e.g., 1976/01 through 1985/12),
- In the expanding window, the number of in-sample data increases as you move forward,
- In the rolling window, the number of in-sample data is fixed as you move forward.

Expanding Window: Step 1

Out-of-Sample Predictability: Expanding Window

At time t (e.g., 1985/12)

- Run the following in-sample regression using only available information

$$y_{2\to t} = \alpha + \beta x_{1\to t-1} + \varepsilon_{2\to t},$$

  - ✓ $y_{2\to t}$ denotes data up to t (e.g., 1976/02 to 1985/12),
  - ✓ $x_{1\to t-1}$ indicates data up to $t-1$ (e.g., 1976/01 to 1985/11).

- Estimate the parameters α and β via least squares and denote them as

$$\widehat{\alpha} \quad \text{and} \quad \widehat{\beta}$$

Out-of-Sample Predictability: Expanding Window

At time t (e.g., 1985/12)

- Gather the estimates $\widehat{\alpha}$ and $\widehat{\beta}$ and construct the out-of-sample forecast as

$$\text{Or simply } \widehat{y}_{t+1} \longleftarrow E_t\left(y_{t+1}\right) = \widehat{\alpha} + \widehat{\beta} x_t$$

✓ $\widehat{y}_{t+1}$ is the forecast for the next period return (time $t+1$) using current information (time t),

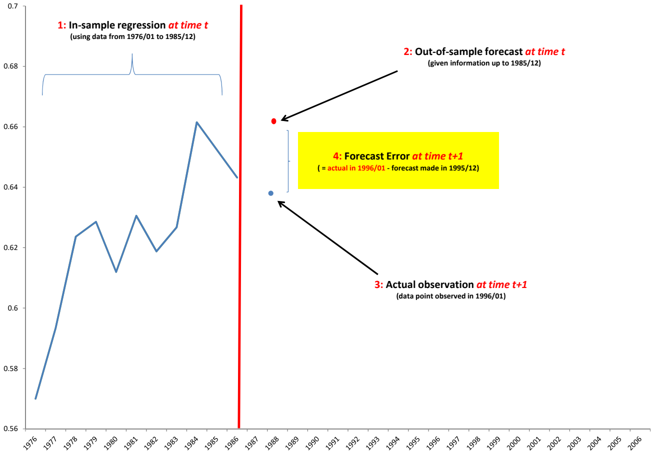
✓ x_t is the current information about your predictor (e.g., 1985/12)

At time $t+1$ (e.g., 1986/01)

- Observe the actual excess return y_{t+1} and construct the forecast error as

$$\widehat{\varepsilon}_{t+1} = y_{t+1} - \widehat{y}_{t+1}$$

Out-of-Sample Predictability: Expanding Window



Expanding Window: Step 2

Out-of-Sample Predictability: Expanding Window

At time $t+1$ (e.g., 1986/01)

- Add an extra observation to y and x and then re-run another in-sample regression as

$$y_{2\to t+1} = \alpha + \beta x_{1\to t} + \varepsilon_{2\to t+1},$$

  ✓ $y_{2\to t+1}$ denotes data up to $t+1$ (e.g., 1976/02 to 1986/01),
  ✓ $x_{1\to t}$ indicates data up to t (e.g., 1976/01 to 1985/12).

- Estimate again the parameters α and β via least squares and denote them as

$$\widehat{\alpha} \quad \text{and} \quad \widehat{\beta}$$

Warning: I use the same notation but these estimates may differ from the previous ones.

Out-of-Sample Predictability: Expanding Window

At time $t+1$ (e.g., 1986/01)

- Gather the new estimates $\widehat{\alpha}$ and $\widehat{\beta}$ and update your out-of-sample forecast as

$$\text{Or simply } \widehat{y}_{t+2} \longleftarrow E_{t+1}\left(y_{t+2}\right) = \widehat{\alpha} + \widehat{\beta} x_{t+1}$$

  ✓ $\widehat{y}_{t+2}$ is the forecast for the next period (time $t+2$) using current information (time $t+1$),

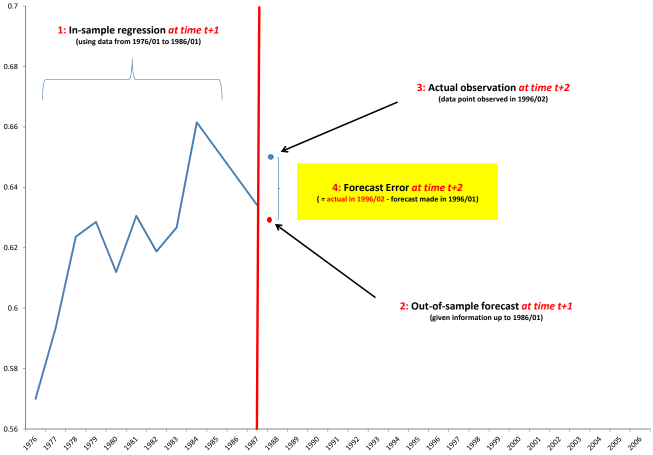
  ✓ x_{t+1} is the current information about your predictor (e.g., 1986/01)

At time $t+2$ (e.g., 1986/02)

- Observe the actual excess return y_{t+2} and construct the forecast error as

$$\widehat{\varepsilon}_{t+2} = y_{t+2} - \widehat{y}_{t+2}$$

Out-of-Sample Predictability: Expanding Window



Repeat These Steps until the End

Out-of-Sample Predictability: Expanding Window

At time $T-1$ (e.g., 2023/11)

- Add the final observation to y and x and then re-run another in-sample regression as

$$y_{2\rightarrow T-1} = \alpha + \beta x_{1\rightarrow T-2} + \varepsilon_{2\rightarrow T-1},$$

  ✓ $y_{2\rightarrow T-1}$ denotes data up to $T-1$ (e.g., 1976/02 to 2023/11),
  ✓ $x_{1\rightarrow T-2}$ indicates data up to $T-2$ (e.g., 1976/01 to 2023/10),

- Estimate again the parameters α and β via least squares and denote them as

$$\widehat{\alpha} \quad \text{and} \quad \widehat{\beta}$$

Warning: I use the same notation but these estimates may differ from the previous ones.

Out-of-Sample Predictability: Expanding Window

At time $T-1$ (e.g., 2023/11)

- Gather the new estimates $\widehat{\alpha}$ and $\widehat{\beta}$ and update your out-of-sample forecast as

$$\text{Or simply } \widehat{y}_T \longleftarrow E_{T-1}(y_T) = \widehat{\alpha} + \widehat{\beta} x_{T-1}$$

  - ✓ $\widehat{y}_T$ is the forecast for the next period (time T) using current information (time $T-1$),
  - ✓ x_{T-1} is the current information about your predictor (e.g., 2023/11).

At time T (e.g., 2023/12)

- Observe the actual excess return y_T and construct the forecast error as

$$\widehat{\varepsilon}_T = y_T - \widehat{y}_T$$

Rolling Window: Step 2

Out-of-Sample Predictability: Rolling Window

At time $t+1$ (e.g., 1986/01)

- Chop out the first observation but add a new observation to y and x so to have a fixed number of observations for your in-sample regression as

$$y_{3\to t+1} = \alpha + \beta x_{2\to t} + \varepsilon_{3\to t+1},$$

  - ✓ $y_{3\to t+1}$ denotes data up to $t+1$ (e.g., 1976/03 to 1986/01),
  - ✓ $x_{2\to t}$ indicates data up to t (e.g., 1976/02 to 1985/12).

- Estimate again the parameters α and β via least squares and denote them as

$$\widehat{\alpha} \quad \text{and} \quad \widehat{\beta}$$

Warning: I use the same notation but these estimates may differ from the previous ones.

Out-of-Sample Predictability: Rolling Window

At time $t+1$ (e.g., 1986/01)

- Gather the new estimates $\widehat{\alpha}$ and $\widehat{\beta}$ and update your out-of-sample forecast as

$$\text{Or simply } \widehat{y}_{t+2} \longleftarrow \boxed{E_{t+1}\left(y_{t+2}\right)} = \widehat{\alpha} + \widehat{\beta} x_{t+1}$$

✓ $\widehat{y}_{t+2}$ is the forecast for the next period (time $t+2$) using current information (time $t+1$),

✓ x_{t+1} is the current information about your predictor (e.g., 1986/01)

At time $t+2$ (e.g., 1986/02)

- Observe the actual excess return y_{t+2} and construct the forecast error as

$$\widehat{\varepsilon}_{t+2} = y_{t+2} - \widehat{y}_{t+2}$$

Repeat These Steps until the End

Out-of-Sample Predictability: Rolling Window

At time $T-1$ (e.g., 2023/11)

- Update y and x and then re-run another in-sample regression as

$$y_{T-w+1\rightarrow T-1} = \alpha + \beta x_{T-w\rightarrow T-2} + \varepsilon_{T-w+1\rightarrow T-1},$$

  ✓ $y_{T-w+1\rightarrow T-1}$ denotes data up to $T-1$ (e.g., 2014/01 to 2023/11),
  ✓ $x_{T-w\rightarrow T-2}$ indicates data up to $T-2$ (e.g., 2013/12 to 2023/10),

- Estimate again the parameters α and β via least squares and denote them as

$$\widehat{\alpha} \quad \text{and} \quad \widehat{\beta}$$

Warning: I use the same notation but these estimates may differ from the previous ones.

Out-of-Sample Predictability: Rolling Window

At time $T-1$ (e.g., 2023/11)

- Gather the new estimates $\widehat{\alpha}$ and $\widehat{\beta}$ and update your out-of-sample forecast as

$$\text{Or simply } \widehat{y}_T \longleftarrow \boxed{E_{T-1}(y_T)} = \widehat{\alpha} + \widehat{\beta} x_{T-1}$$

✓ $\widehat{y}_T$ is the forecast for the next period (time T) using current information (time $T-1$),

✓ x_{T-1} is the current information about your predictor (e.g., 2023/11).

At time T (e.g., 2023/12)

- Observe the actual excess return y_T and construct the forecast error as

$$\widehat{\varepsilon}_T = y_T - \widehat{y}_T$$

Economic Restrictions

Economic Restrictions

Campbell and Thompson (2008) recommend economic restrictions

- Often theory suggests the sign of the forecasts,
- One should then impose sign restrictions that are economically motivated.

Risk-based explanations predict a positive expected excess return when risk is high

1. Set $\widehat{\beta} = 0$ when $\widehat{\beta} < 0$,
2. Set $\widehat{y}_{t+1} = 0$ when $\widehat{y}_{t+1} < 0$,
3. Set $\widehat{\beta} = 0$ when $\widehat{\beta} < 0$ and $\widehat{y}_{t+1} = 0$ when $\widehat{y}_{t+1} < 0$.

Performance Evaluation

Performance Evaluation

What's the next?

- We need to evaluate the accuracy of the out-of-sample forecasts,
- This evaluation must be done for all models.

How do we do it?

- We must minimize a loss function

$$\text{Loss Function} = L(y_{t+1}, \widehat{y}_{t+1})$$

- y_{t+1} is the sequence of actual excess returns in our sample,
- $\widehat{y}_{t+1}$ is the sequence of out-of-sample forecasts,
- We focus on one-step-ahead forecasts.

Performance Evaluation

Properties of the loss function

1. The loss of any forecast is non-negative

$$L(y_{t+1}, \widehat{y}_{t+1}) \geq 0$$

2. There is an optimal forecast $y^{\star}_{t+1}$ where

$$L(y_{t+1}, y^{\star}_{t+1}) = 0$$

3. The loss is non-decreasing away from $y^{\star}_{t+h}$

$$y^{B}_{t+1} > y^{A}_{t+1} > y^{\star}_{t+1} \Longrightarrow L(y_{t+1}, y^{B}_{t+1}) > L(y_{t+1}, y^{A}_{t+1}) > L(y_{t+1}, y^{\star}_{t+1})$$

Popular Loss Function

Mean Squared Error (MSE)

- The mean squared error minimize the average variance of the forecast errors

$$MSE = \frac{1}{T-t} \sum_{i=t+1}^{T} \left(y_i - \widehat{y}_i\right)^2$$
$$= \frac{1}{T-t} \sum_{i=t+1}^{T} \widehat{\varepsilon}_i^2$$

Interpretation

- A model with a lower MSE produces more accurate out-of-sample forecasts.
- This loss function, however, can be sensitive to few large forecast errors.

Alternative Loss Functions

Mean Absolute Error (MAE)

- The mean absolute error minimize the average absolute forecast errors

$$MAE = \frac{1}{T-t} \sum_{i=t+1}^{T} |\widehat{\varepsilon}_i|$$

Root Mean Squared Error (RMSE)

- The root mean squared error minimize the average volatility of forecast errors

$$RMSE = \sqrt{\frac{1}{T-t} \sum_{i=t+1}^{T} \widehat{\varepsilon}_i^2}$$

Interpretation

- A model with a lower MAE or RMSE produces more accurate out-of-sample forecasts.

Statistical Evaluation

Out-of-sample R-squared suggested by Campbell and Thompson (2008)

$$R^2_{oos} = 1 - \frac{MSE_{MOD}}{MSE_{BEN}}$$

- MSE_{BEN} is the mean squared error of the benchmark model (BEN),
- MSE_{MOD} is the mean squared error of the competing model (MOD).

Interpretation

- $R^2_{oos} > 0$ means that *MOD* outperforms *BEN* with a lower MSE.

Statistical Evaluation

Out-of-sample RMSE difference by Welch and Goyal (2008)

$$\Delta RMSE = RMSE_{BEN} - RMSE_{MOD}$$

- $RMSE_{BEN}$ is the root mean squared error of the benchmark model (BEN),
- $RMSE_{MOD}$ is the root mean squared error of the competing model (MOD).

Interpretation

- $\Delta RMSE > 0$ means that *MOD* outperforms *BEN* being its RMSE lower.

Mincer-Zarnowitz Regression

The Mincer-Zarnowitz regression (MZ) tests for the optimal forecast

$$y_{t+1} = \alpha + \beta \widehat{y}_{t+1} + \eta_t$$

- If the forecast is optimal, we should expect $\alpha = 0$ and $\beta = 1$.

The augmented Mincer-Zarnowitz regression (AMZ)

$$y_{t+1} = \alpha + \beta \widehat{y}_{t+1} + \gamma_1 x_{1,t} + \ldots + \gamma_k x_{k,t} + \eta_t$$

- If the forecast is optimal, we should expect $\alpha = 0$ and $\beta = 1$, after controlling for measurable variables like liquidity, volatility, etc.

Statistical Evaluation

We could also test the null hypothesis H_0 of equal predictive ability

$$H_0 : MSE_{BEN} = MSE_{MOD}$$

- MSE_{BEN} is the mean squared error of the benchmark model (BEN),
- MSE_{MOD} is the mean squared error of the competing model (MOD).

Clark and West (2006, 2007) note that

- *MOD* employs additional parameters (perhaps useless for prediction) and these additional parameters may generate estimation error,
- If *MOD* and *BEN* were equally good, $MSE_{MOD} > MSE_{BEN}$ and this bias must be corrected.

Statistical Evaluation

Clark and West (2007) propose a simple test designed for nested models

$$f_i = (y_i - \overline{y}_i)^2 - [(y_i - \widehat{y}_i)^2 - (\overline{y}_i - \widehat{y}_i)^2]$$

- y_{i+1} is the actual asset return at time i,
- $\overline{y}_i$ is the forecast of the benchmark model,
- $\widehat{y}_i$ is the forecast of the competing model,
- $(\overline{y}_i - \widehat{y}_i)^2$ is the correction term.

You construct f_{i+1} for each out-of-sample observation in your sample $(i = t+1, \ldots, T)$

$$f_{t+1}, f_{t+2}, \ldots, f_T$$

Statistical Evaluation

Regress the f_i on a constant as

$$f_i = \rho + \eta_t$$

Hence, construct a simple t-test

$$CW = \frac{\widehat{\rho}}{se(\widehat{\rho})}$$

The asymptotic distribution of CW is non-standard

- Critical values from a normal distribution provide a good approximation,
- If $CW > 1.65$, we reject the null hypothesis at the 5% level (one-sided test),
- If $CW > 1.28$, we reject the null hypothesis at the 10% level (one-sided test),
- A bootstrap procedure is, however, common when the sample is small.

Market Timing Evaluation

Henriksson and Merton (1981) propose a simple market timing test

$$p = \frac{c}{n}$$

- $c =$ number of correct forecasts (same direction),
- $n =$ total number of forecasts,
- $p =$ probability of directional accuracy.

Hypothesis Testing

$$H_0 : p = 0.5 \quad \text{(no market timing)}$$

$$H_A : p > 0.5 \quad \text{(positive market timing)}$$

Henriksson and Merton (1981)

p is distributed as a Binomial distribution with

$$E[p] = r/n$$

$$Var[p] = \frac{p(1-p)}{n}$$

Since the Binomial is approximately Normal distribution for large samples

$$Z = \frac{E[p] - 0.5}{\sqrt{Var[p]}} \sim N(0,1)$$

If $Z > 1.65$, we reject the null hypothesis at the 5% level (one-sided test).

Economic Evaluation

"Economists are often puzzled as to why profit-maximizing firms buy economic forecasts. Summary statistics [...] **rarely reveal major differences** *between professional forecasting services and a simple naive approach of no change in the variable being forecast... Yet, millions of dollars are spent annually both producing and purchasing these apparently worthless forecasts."*

Leitch and Tanner, *American Economic Review*, 1992

"The standard criteria show no consistent relationship with profits. Indeed, MAE and RMSE criteria have perverse signs (better forecasts should have lower average errors and higher profits, and thus, the simple correlations should be negative)."

Leitch and Tanner, *American Economic Review*, 1992

- A purely statistical analysis may not be informative of possible tangible economic gains from using dynamic forecasts in active portfolio management.

Economic Evaluation

In addition to an extensive literature on statistical evaluation, there is recent research proposing a methodology for assessing the economic value of exchange rate predictability.

A purely statistical analysis of predictability is not particularly informative to an investor as it falls short of measuring whether there are tangible economic gains from using dynamic forecasts in active portfolio management.

We review this approach based on dynamic asset allocation that is used, among others, by West, Edison and Cho (1993), Fleming, Kirby and Ostdiek (2001), Han (2006), and Della Corte, Sarno and Thornton (2008) and Della Corte, Sarno and Tsiakas (2009, 2011), Della Corte, Jeanneret, and Patelli (2023).

Economic Evaluation

Consider an investor allocating her wealth between a riskless asset and N risky assets

$$\begin{bmatrix} \widehat{r}_{1,t+1} \\ \widehat{r}_{2,t+1} \\ \vdots \\ \widehat{r}_{N,t+1} \end{bmatrix} = \begin{bmatrix} \widehat{y}_{1,t+1} \\ \widehat{y}_{1,t+2} \\ \vdots \\ \widehat{y}_{N,t+1} \end{bmatrix} + \begin{bmatrix} r_{f,t} \\ r_{f,t} \\ \vdots \\ r_{f,t} \end{bmatrix}$$

$$\widehat{\boldsymbol{r}}_{t+1} = \widehat{\boldsymbol{y}}_{t+1} + \boldsymbol{r}_{f,t}$$

- $\widehat{\boldsymbol{r}}_{t+1} \rightarrow$ vector of expected returns on the risky assets,
- $\widehat{\boldsymbol{y}}_{t+1} \rightarrow$ vector of expected excess returns on the risky assets,
- $\boldsymbol{r}_{f,t} \rightarrow$ vector of riskless rate of returns.

Maximum Expected Return

Each period t, the investor solves the following problem

$$\begin{aligned} \max_{\boldsymbol{w}_t} \quad \widehat{r}_{p,t+1} &= \boldsymbol{w}_t'\widehat{\boldsymbol{r}}_{t+1} + \left(1 - \boldsymbol{w}_t'\boldsymbol{\iota}\right) r_{f,t} \\ \text{s.t.} \quad \sigma_p^{\star} &= \left(\boldsymbol{w}_t'\widehat{\boldsymbol{\Sigma}}_{t+1}\boldsymbol{w}_t\right)^{1/2} \end{aligned}$$

- $\widehat{r}_{p,t+1} \rightarrow$ expected portfolio return,
- $\boldsymbol{w}_t \rightarrow$ vector of weights on the risky assets,
- $\widehat{\boldsymbol{\Sigma}}_{t+1} \rightarrow$ is conditional covariance matrix of the risky returns,
- $\boldsymbol{\iota} \rightarrow$ vector of ones,
- $r_{f,t} \rightarrow$ riskless return,
- $\sigma_p^{\star} \rightarrow$ target portfolio volatility (e.g., 10% in annual terms).

Maximum Expected Return

The solution for $\boldsymbol{w_t}$ is in closed-form

$$\boldsymbol{w_t} = \frac{\sigma_p^*}{\sqrt{C_t}} \widehat{\boldsymbol{\Sigma}}_{t+1}^{-1} (\widehat{\boldsymbol{r}}_{t+1} - i_t)$$

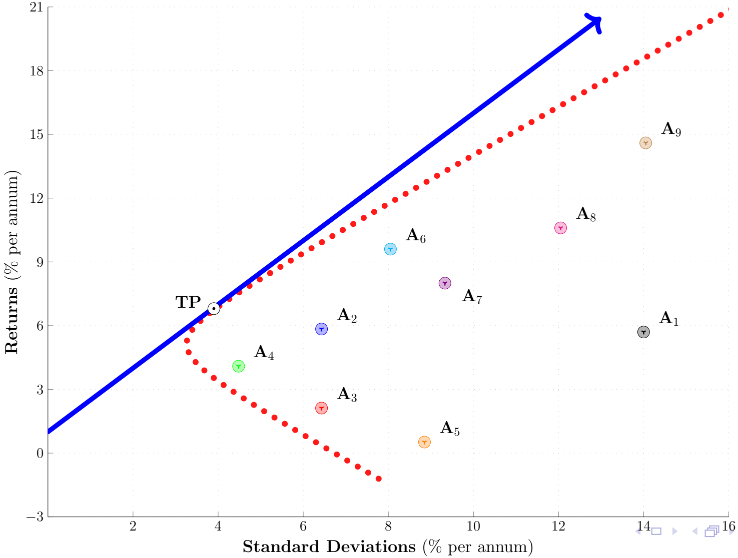
where

$$C_t = (\widehat{\boldsymbol{r}}_{t+1} - r_{ft})' \widehat{\boldsymbol{\Sigma}}_{t+1}^{-1} (\widehat{\boldsymbol{r}}_{t+1} - r_{f,t}).$$

$$\widehat{\boldsymbol{\Sigma}}_{t+1} = \begin{bmatrix} \sigma_{\varepsilon_1\varepsilon_1} & \sigma_{\varepsilon_1\varepsilon_2} & \cdots & \sigma_{\varepsilon_1\varepsilon_N} \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma_{\varepsilon_2\varepsilon_2} & \cdots & \sigma_{\varepsilon_2\varepsilon_N} \\ \vdots & \cdots & \ddots & \cdots \\ \sigma_{\varepsilon_N\varepsilon_1} & \sigma_{\varepsilon_N\varepsilon_2} & \cdots & \sigma_{\varepsilon_N\varepsilon_N} \end{bmatrix}_t$$

$\widehat{\boldsymbol{\Sigma}}_{t+1}$ is the sample covariance of the regression residuals at time (more sophisticated methods like EWMA, GARCH, SV, and DCC could also be used).

Maximum Expected Return: An Illustration



Maximum Expected Return

At time t, for each exchange rate model

- Generate the one-step ahead forecasts: $\widehat{\boldsymbol{r}}_{t+1}$ and $\widehat{\boldsymbol{\Sigma}}_{t+1}$,
- Compute the optimal portfolio weights $\boldsymbol{w_t}$ via mean-variance,
- Rebalance the wealth given the new portfolio weights.

At time $t+1$, compute the realized portfolio return as

$$r_{p,t+1} = \boldsymbol{w_t}'\boldsymbol{r}_{t+1} + \left(1 - \boldsymbol{w_t}'\boldsymbol{\iota}\right) r_{f,t}$$

and

$$\boldsymbol{r}_{t+1} = \boldsymbol{y}_{t+1} + \boldsymbol{r}_{f,t}$$

- $r_{p,t+1} \rightarrow$ realized portfolio return,
- $\boldsymbol{r}_{t+1} \rightarrow$ vector of realized risky returns,
- $\boldsymbol{y}_{t+1} \rightarrow$ vector of realized excess returns.

Performance Evaluation

Quadratic utility

$$U\left(W_{t+1}\right) = W_{t+1} - \frac{\lambda}{2} W_{t+1}^2 = W_t \underbrace{R_{p,t+1}}_{1+r_{p,t+1}} - \frac{\lambda}{2} W_t^2 \underbrace{R_{p,t+1}^2}_{(1+r_{p,t+1})^2}$$

Constant degree of relative risk aversion

$$\frac{\lambda}{1-\lambda} = \delta$$

Average Realized Utility (West, Edison and Cho, 1993)

$$\overline{U} = W_t \sum_{i=t+1}^{T} \left\{ R_{p,i} - \frac{\delta}{2\left(1+\delta\right)} R_{p,i}^2 \right\}$$

an estimate of the expected utility generated by a given W_t.

Performance Evaluation

Fleming, Kirby & Ostdiek (2001)

$$\sum_{i=t+1}^{T} \left\{ \left(R^{\star}_{p,i} - \mathcal{F}\right) - \frac{\delta}{2\left(1+\delta\right)} \left(R^{\star}_{p,i} - \mathcal{F}\right)^2 \right\}$$

$$= \sum_{t=t+1}^{T} \left\{ R_{p,i} - \frac{\delta}{2\left(1+\delta\right)} R^2_{p,i} \right\}$$

- R_p $\rightarrow$ gross portfolio return implied by the benchmark model,
- $R^{\star}_p$ $\rightarrow$ gross portfolio return implied by the competing model,
- $\mathcal{F}$ $\rightarrow$ performance fee you are willing to pay to switch to *MOD*,
- $\mathcal{F}$ $\rightarrow$ in annualized basis points (i.e., $\mathcal{F} = 100$ means 1%),
- W_0 is set equal to \$1 every period.

Performance Evaluation

Goetzmann, Ingersoll, Spiegel & Welch (2007)

$$M = \frac{1}{(1-\delta)} \ln \left\{ \frac{1}{T-1} \sum_{t=t+1}^{T} \left(\frac{R_{p,i}}{R_{f,i-1}} \right)^{1-\delta} \right\}$$

- $\delta \rightarrow$ can be thought as the degree of relative risk aversion,
- $M \rightarrow$ the certainty equivalent generated by a given strategy,
- $R_{f,t} = 1 + r_{f,t} \rightarrow$ gross domestic riskless return.

Consider the difference of performance measures as

$$\mathcal{P} = M_{MOD} - M_{BEN}$$

- $\mathcal{P}$ is the performance fee that you are willing to pay to switch from *BEN* to *MOD* and is reported in basis points per annum.

Performance Evaluation

The **Sharpe ratio** ($\mathcal{SR}$) quantifies excess return per unit of vol

- Average excess return divided by its standard deviation

$$SR = \frac{E\left(r_{p,t+1} - r_{f,t}\right)}{\sqrt{Var\left(r_{p,t+1} - r_{f,t}\right)}}$$

The **Sortino ratio** ($\mathcal{SO}$) is excess return per unit of 'bad' vol

- Average excess return over the standard deviation of negative excess returns

$$SO = \frac{E\left(r_{p,t+1} - r_{f,t}\right)}{\sqrt{Var\left(r_{p,t+1} - r_{f,t} < 0\right)}}$$

- $\mathcal{SO}$ differentiates between 'ups and downs' movements in portfolio returns and a large $\mathcal{SO}$ means a low risk of a large losses.

Transaction Costs Impact

Transaction costs are critical to assess the profitability of a trading strategy and large investor may face lower costs than small investors.

We compute the break-even transaction cost that makes an investor indifferent between models

$$\tau_{t+1} = \frac{1}{N} \sum_{j=1}^{N} \left| w_{t+1}^{j} - \underbrace{w_t^j \left(\frac{1 + r_{t+1}^j}{1 + r_{p,t+1}} \right)}_{\tilde{w}_{t+1}^{j}} \right|$$

- w_t^i $\rightarrow$ weight on asset i at time t after rebalancing,
- $\tilde{w}_{t+1}^i$ $\rightarrow$ weight on asset i at time $t+1$ before rebalancing,
- w_{t+1}^i $\rightarrow$ weight on asset i at time $t+1$ after rebalancing.

Performance Evaluation

The break-even transaction cost is computed as

$$\tau^{be} = \frac{\overline{R}_p^{\star} - \overline{R}_p}{\overline{\tau}^{\star} - \overline{\tau}}$$

- $\overline{R}_p^{\star} \rightarrow$ average gross portfolio return for *MOD*,
- $\overline{R}_p \rightarrow$ average gross portfolio return for *BEN*,
- $\overline{\tau}^{\star} \rightarrow$ average transaction cost for *MOD*,
- $\overline{\tau} \rightarrow$ average transaction cost for *BEN*.

An investor prefers *MOD* to *BEN* is she pays lower transaction costs than τ^{be}.

Model Combination

Combined Forecasts

So far, we have evaluated the performance of individual models relative to the benchmark but ex-ante we ignore which model is true and this generates model uncertainty.

We account for model uncertainty using combined forecasts from the full set of predictive regressions. The superior performance of combined forecasts is known since the seminal work of Bates and Granger (1969).

We review three types of combined forecasts

- Simple model averaging,
- Statistical model averaging,
- Economic model averaging.

Combined Forecasts

For each risky asset,

- We estimate predictive regressions for $n = 1, \ldots, M$ models,
- Each regression provides an individual forecast $\widehat{y}^m_{t+1}$.

The combined forecast $\widetilde{y}_{t+1}$ is a weighted average of individual forecasts

$$\widetilde{y}_{t+1} = \sum_{m=1}^{M} \kappa_t^m \times \widehat{y}^m_{t+1}$$

- κ_t^m is the ex-ante weight on m at time t,
- How do we compute these ex-ante model weights?

Simple Model Averaging

The **mean rule**

- For each forecast, set $\kappa_t^m = 1/M$.

The **median rule**

- Select the median across κ_t^m.

The **trimmed mean rule**

- Remove the largest and lowest values from κ_t^m,
- Set $\kappa_t^m = 1/\left(M-2\right)$ for the remaining individual forecasts.

Statistical Model Averaging

At time t, construct the combining weights as

$$\kappa_t^m = \frac{1/\,MSE_t^m}{\sum_{m=1}^{M} 1/\,MSE_t^m}$$

- $\kappa_t^m \rightarrow$ a low MSE_t^m implies a high weight for the forecast n,
- $MSE_t^m \rightarrow$ mean-squared error based on data up to time t,
- See Bates and Granger (1969) and Stock and Watson (2004).

Economic Model Averaging

At time t, construct the combining weights as

$$\kappa_t^m = \frac{SR_t^m}{\sum_{n=1}^{M} SR_t^m}$$

- $\kappa_t^m \rightarrow$ a high SR_t^m implies a high weight for the forecast n,
- $SR_t^m \rightarrow$ Sharpe ratio based on data up to time t,
- See Della Corte and Tsiakas (2012).

A Simple Application

A Currency Fund Manager

Which model should an asset manager use for asset allocation?

- We answer this question by considering a set of models,
- We run out-of-sample tests using statistical and economic criteria.

Our benchmark model

- Random Walk with drift (RW)

Our competing model

- Purchasing Power Parity (PPP),
- Uncovered Interest Parity (UIP),
- Monetary Fundamentals (MF),
- Symmetric Taylor rule (TR_s),
- Asymmetric Taylor rule (TR_a).

Predictive Regression

Consider the following one-step ahead predictive regression

$$y_t = \alpha + \beta x_{t-1} + u_t$$

where $y_t = s_t - s_{t-1}$ is the nominal exchange rate return.

Random Walk with drift (RW)

$$x_{t-1} = 0$$

Uncovered Interest Parity (UIP)

$$x_{t-1} = i_{t-1} - i^*_{t-1}$$

Purchasing Power Parity (PPP)

$$x_{t-1} = p_{t-1} - p^*_{t-1} - s_{t-1}.$$

Predictive Regression

Monetary Fundamentals (MF)

$$x_{t-1} = \left(m_{t-1} - m_{t-1}^*\right) - \left(y_{t-1} - y_{t-1}^*\right) - s_{t-1}$$

Symmetric Taylor Rule (TR_s)

$$x_{t-1} = 0.1\left(g_{t-1} - g_{t-1}^*\right) + 1.5\left(\pi_{t-1} - \pi_{t-1}^*\right)$$

- the foreign country does not target the spot rate,

Asymmetric Taylor Rule (TR_a)

$$x_{t-1} = 0.1 q_{t-1} + 0.1\left(g_{t-1} - g_{t-1}^*\right) + 1.5\left(\pi_{t-1} - \pi_{t-1}^*\right)$$

- the foreign country targets the exchange rate in its Taylor rule,
- We set β_0, β_1 and β_2 as in Engel, Mark and West (2007).

Data Sample

The sample data consists of

- End-of-month data ranging from January 1976 to June 2010,
- nine currencies relative to the US dollar (USD): Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Deutsche mark-euro (EUR), British pound (GBP), Japanese yen (JPY), Norwegian kroner (NOK), New Zealand dollar (NZD), and Swedish kronor (SEK),
- A variety of sources such as Datastream, OECD, etc.

Our OOS period runs from January 1986 to December 2010

- First OOS forecast uses data from Jan-76 to Dec-85,
- Moving forward using an expanding window until May 2010.

The Bootstrap Algorithm

This appendix summarizes the bootstrap algorithm we use for generating critical values for the OOS test statistics under the null of no exchange rate predictability against a one-sided alternative of linear predictability. Following Mark (1995) and Kilian (1999), the algorithm consists of the following steps:

1. Define the IS period for $\left\{\Delta s_{t+1}, x_t\right\}_{t=1}^{M}$ and the OOS period for $\{\Delta s_{t+1}, x_t\}_{t=M+1}^{T-1}$. We generate $P = (T-1) - M$ OOS forecasts $\{\Delta\bar{s}_{t+1|t}, \Delta\widehat{s}_{t+1|t}\}_{t=M+1}^{T-1}$ by estimating the predictive regression

$$\Delta s_{t+1} = \alpha + \beta x_t + \varepsilon_{t+1}$$

and then computing the test statistic of interest, $\widehat{\tau}$.

2. Define the data generating process (DGP) as

$$\begin{aligned}\Delta s_{t+1} &= \alpha + \beta x_t + u_{1,t+1} \\ x_t &= \mu + \rho_1 x_{t-1} + \ldots + \rho_p x_{t-p} + u_{2,t},\end{aligned}$$

and estimate this model subject to the constraint that β in the first equation is zero, using the full sample of observations $\left\{\Delta s_{t+1}, x_t\right\}_{t=1}^{T-1}$. The lag order p in the second equation is determined by a suitable lag order selection criterion such as the Bayesian information criterion (BIC).

The Bootstrap Algorithm

3. Generate a sequence of pseudo-observations $\left\{\Delta s_t^*, x_{t-1}^*\right\}_{t=1}^{T-1}$ as follows:

$$
\begin{aligned}
\Delta s_{t+1}^* &= \widehat{\alpha} + u_{1,t+1}^* \\
x_t^* &= \widehat{\mu} + \widehat{\rho}_1 x_{t-1}^* + \ldots + \widehat{\rho}_p x_{t-p}^* + u_{2,t}^*.
\end{aligned}
$$

The pseudo-innovation term $u_t^* = (u_{1,t}^*, u_{2,t}^*)'$ is randomly drawn with replacement from the set of observed residuals $\widehat{u}_t = (\widehat{u}_{1,t}, \widehat{u}_{2,t})'$. The initial observations $\left(x_{t-1}^*, \ldots, x_{t-p}^*\right)'$ are randomly drawn from the actual data. Repeat this step $B = 10{,}000$ times.

4. For each of the B bootstrap replications, define an IS period for $\left\{\Delta s_{t+1}^*, x_t^*\right\}_{t=1}^{M}$, and an OOS period for $\left\{\Delta s_{t+1}^*, x_t^*\right\}_{t=M+1}^{T-1}$. Then, generate P OOS forecasts $\{\Delta \bar{s}_{t+1|t}^*, \Delta \widehat{s}_{t+1|t}^*\}_{t=M+1}^{T-1}$ by estimating the predictive regression

$$
\Delta s_{t+1}^* = \alpha^* + \beta^* x_t^* + u_{1,t+1}^*
$$

both under the null and the alternative for $t = M+1, \ldots, T-1$, and construct the test statistic of interest, $\widehat{\tau}^*$.

The Bootstrap Algorithm

5. Compute the one-sided p-value of $\widehat{\tau}$ as

$$p\text{-value} = \frac{1}{B}\sum_{j=1}^{B} I(\widehat{\tau}^* > \widehat{\tau}),$$

where $I(\cdot)$ denotes an indicator function, which is equal to 1 when its argument is true and 0 otherwise.

Statistical Evaluation

	UIP	PPP	MF	TR_s	TR_a	UIP	PPP	MF	TR_s	TR_a
			AUD					**CAD**		
R^2_{oos} (%)	-0.26	0.34^a	-1.21	-0.34	0.91^b	0.06	0.30^a	-1.32	-0.41	0.09^a
$\Delta RMSE$ (%)	0.00	0.01^a	-0.02	-0.01	0.02^b	0.00	0.01^a	-0.01	0.00	0.01^a
CW	-0.49	0.96	-0.53	-1.04	1.76^b	1.21^a	1.16^a	-0.22	-1.13	0.68
			CHF					**EUR**		
R^2_{oos} (%)	-1.68	1.11^c	-0.35	-0.81	-2.37	-1.55	0.74^b	-0.73	-1.06	-1.71
$\Delta RMSE$ (%)	-0.03	0.02^c	-0.01	-0.01	-0.04	-0.02	0.01^b	-0.01	-0.02	-0.03
CW	0.64	1.88^b	1.78^b	-0.17	0.36	0.10	1.66^b	1.10^a	-0.56	-0.05
			GBP					**JPY**		
R^2_{oos} (%)	-1.22	1.22^c	-0.33	-0.69	1.64^c	0.71^b	0.30^a	-1.67	-0.85	-1.64
$\Delta RMSE$ (%)	-0.02	0.02^c	0.00	-0.01	0.02^c	0.01^b	0.01^a	-0.03	-0.01	-0.03
CW	0.29	1.93^b	0.88	-1.14	2.87^c	1.78^b	0.91	1.07^a	-0.88	-0.39
			NOK					**NZD**		
R^2_{oos} (%)	-0.50	0.68^b	-1.42	-1.53	1.01^b	0.11	-0.37	-0.65	-0.06	1.15^c
$\Delta RMSE$ (%)	-0.01	0.01^b	-0.02	-0.02	0.02^b	0.00	-0.01	-0.01	0.00	0.02^c
CW	-0.04	1.46^b	-1.29	-0.88	2.00^c	0.85	-0.61	-0.62	0.61	2.10^c
			SEK							
R^2_{oos} (%)	-1.00	0.72^b	-0.28	-0.60	0.77^b					
$\Delta RMSE$ (%)	-0.02	0.01^b	0.00	-0.01	0.01^b					
CW	-0.85	1.59^b	0.71	-1.31	1.69^b					

Sources: Della Corte & Tsiakas (2012).

Economic Evaluation

	μ_p	σ_p	$\mathcal{SR}$	$\mathcal{SO}$	$\mathcal{F}$	$\mathcal{P}$	τ^{be}
RW	10.8	11.4	0.54	0.73			
UIP	11.9	11.1	0.65	0.99	131	143	173
PPP	13.3	11.3	0.76	0.97	252	247	70
MF	11.1	11.7	0.55	0.73	10	4	—
TRs	7.1	11.6	0.21	0.23	−384	−433	—
TRa	12.1	11.4	0.65	0.83	130	121	161

Sources: Della Corte & Tsiakas (2012).

Model Averaging

	μ	σ	$\mathcal{SR}$	$\mathcal{SO}$	$\mathcal{F}$	$\mathcal{P}$	τ^{be}
RW	10.8	11.4	0.54	0.73			
Simple Model Averaging							
Mean	13.3	11.6	0.74	0.89	234	206	81
Median	12.0	12.0	0.61	0.74	76	52	21
Trimmed Mean	12.0	12.1	0.60	0.72	65	37	23
Statistical Model Averaging							
MSE	13.3	11.6	0.74	0.89	232	204	81
$MSE(\kappa = 60m)$	13.2	11.7	0.73	0.88	222	195	76
$MSE(\kappa = 36m)$	13.2	11.7	0.73	0.88	218	191	74
$MSE(\kappa = 12m)$	13.1	11.7	0.72	0.87	207	179	70
Economic Model Averaging							
SR	12.9	11.4	0.72	0.89	207	187	103
$SR(\kappa = 60m)$	13.5	11.5	0.76	0.96	261	241	129
$SR(\kappa = 36m)$	13.4	11.5	0.76	0.95	254	234	126
$SR(\kappa = 12m)$	13.4	11.5	0.76	0.96	254	235	128

Sources: Della Corte & Tsiakas (2012).

Model Averaging