ECON4917/ECMT6003 Applied Project (DRAFT)

Christian Bousamra, Alessandro Petagna, Patrick Mannion

May 2025

Introduction

This report presents an empirical analysis of forecasting Australian stock returns, building upon the work of Rapach, Strauss, and Zhou (2013). We replicate key findings from their study regarding the predictability of international stock returns, particularly the role of U.S. predictors for Australia. We then develop a custom forecasting model using a rolling window estimation approach and compare its out-of-sample performance to benchmark models, including the historical average and a recursive estimation model.

Part (a): Replication of Table II Results

Objective

Replicate the coefficient estimates for the lagged Australian Treasury bill rate ($\hat{\beta}_{AUS,b}$) and log dividend yield ($\hat{\beta}_{AUS,d}$) from the predictive regression model in Table II of Rapach, Strauss, and Zhou (2013). Compute and report the Newey-West HAC t-statistics with lag order 4.

Model

$$r_{AUS,t+1} = \beta_{AUS,0} + \beta_{AUS,b} bill_{AUS,t} + \beta_{AUS,d} dy_{AUS,t} + \epsilon_{AUS,t+1}$$

Results

Coefficient Estimates (OLS):

-
  - $\hat{\beta}_{AUS,0}$ (Intercept): -0.218815
  - $\hat{\beta}_{AUS,b}$ (Lagged Bill): -0.049174
  - $\hat{\beta}_{AUS,d}$ (Lagged DY): 0.683115

- **Newey-West HAC t-statistics (Lag order 4):**
  t-stat for $\hat{\beta}_{AUS,0}$: -1.637195 t-stat for $\hat{\beta}_{AUS,b}$: -9.029476 t-stat for $\hat{\beta}_{AUS,d}$: 5.902007

Screenshot of Part (a) Results

Interpretation

The estimated coefficients for the lagged Australian Treasury bill rate (-0.0492) and log dividend yield (0.6831) are consistent with the values reported in Table II of Rapach, Strauss, and Zhou (2013). The Newey-West HAC t-statistics indicate the statistical significance of these predictors for forecasting Australian equity premium. The t-statistic of -9.029 for the lagged bill rate and 5.902 for the lagged dividend yield suggest that both variables are statistically significant predictors of the Australian equity premium at conventional significance levels.

```
% Results for Part (a) %
Coefficient Estimates (OLS) for Part a:
beta_AUS,0 (Intercept): -0.218815
beta_AUS,b (Lagged Bill): -0.049174
beta_AUS,d (Lagged DY): 0.683115

Newey-West HAC t-statistics (Lag order 4) for Part a:
t-stat for beta_AUS,0: -1.637195
t-stat for beta_AUS,b: -9.029476
t-stat for beta_AUS,d: 5.902007
```

Figure 1: MATLAB Output for Part (a)

Part (b): Replication of Table III Results

Objective

Replicate the coefficient estimate for the lagged U.S. equity premium ($\hat{\beta}_{AUS,USA}$) from the predictive regression model in Table III of Rapach, Strauss, and Zhou (2013). Report the Newey-West HAC t-statistic with lag order 4 and interpret the finding.

Model

$$r_{AUS,t+1} = \beta_{AUS,0} + \beta_{AUS,b} bill_{AUS,t} + \beta_{AUS,d} dy_{AUS,t} + \beta_{AUS,USA} r_{USA,t} + \epsilon_{AUS,t+1}$$

Results

- **Coefficient Estimates (OLS):**
  - $\hat{\beta}_{AUS,0}$ (Intercept): -0.540316
  - $\hat{\beta}_{AUS,b}$ (Lagged Bill): -0.060479
  - $\hat{\beta}_{AUS,d}$ (Lagged DY): 0.911855
  - $\hat{\beta}_{AUS,USA}$ (Lagged US Return): 0.203402

- **Newey-West HAC t-statistics (Lag order 4):**
  - t-stat for $\hat{\beta}_{AUS,0}$: -4.285467
  - t-stat for $\hat{\beta}_{AUS,b}$: -12.015831
  - t-stat for $\hat{\beta}_{AUS,d}$: 8.437367
  - t-stat for $\hat{\beta}_{AUS,USA}$: 49.258497

Screenshot of Part (b) Results

```
% Results for Part (b) - Replicating Table III (Australia vs. USA) %
  beta_AUS,USA estimate: 0.203402
  HAC t-statistic: 49.258497
```

Figure 2: MATLAB Output for Part (b)

Interpretation

The estimated coefficient for the lagged U.S. equity premium ($\hat{\beta}_{AUS,USA}$) is 0.2034, which is close to the 0.2 reported in Table III of the paper. The highly significant Newey-West HAC t-statistic of 49.26 indicates a strong positive statistical relationship between the lagged U.S. equity premium and the current Australian equity premium, even after controlling for local Australian predictors. This finding supports the authors' argument about the significant role of U.S. stock returns in predicting returns in other industrialized countries like Australia.

Part (c): Replication of Out-of-Sample R-squared (Table VII, Column 2)

Objective

Replicate the out-of-sample R-squared (R^2_{OS}) statistic reported in the second column of Table VII. This statistic compares the out-of-sample forecasting performance of a model with a constant and lagged U.S. returns against a historical average forecast.

Results

Out-of-Sample R-squared (Constant + Lagged US vs Historical Average):

- – R^2_{OS}: 2.972892%

Screenshot of Part (c) Results

```
% Results for Part (c) %
Out-of-Sample R-squared (Constant + Lagged US vs Historical Average) for Australia (Part c - Table VII, Col 2):
R2OS: 2.972892%
```

Figure 3: MATLAB Output for Part (c)

Interpretation (PROBABLY WRONG)

The calculated out-of-sample R-squared of 2.97% indicates that the model using a constant and lagged U.S. returns achieved a 2.97% reduction in Mean Squared Forecast Error (MSFE) compared to the historical average forecast over the out-of-sample period (1985:01-2010:12). While this shows some predictive ability relative to the historical average, it differs from the -0.69% reported in Table VII of the paper.

Part (d): Custom Model - Rolling Window Estimation

Objective

Build a custom forecasting model by employing rolling window estimation for the predictive regression model from Part (b). Investigate the out-of-sample performance of this model relative to the historical average.

Model

$$r_{AUS,t+1} = \beta_{AUS,0} + \beta_{AUS,b} bill_{AUS,t} + \beta_{AUS,d} dy_{AUS,t} + \beta_{AUS,USA} r_{USA,t} + \epsilon_{AUS,t+1}$$

(Estimated using a 59-month rolling window)

Results

Out-of-Sample R-squared (Rolling Window Predictive Regression vs Historical Average):

- – Rolling Window Size: 59 months
  – R^2_{OS}: 13.739038%

Screenshot of Part (d) Results

```
% Results for Part (d) %
Out-of-Sample R-squared (Rolling Window Predictive Regression vs Historical Average) for Australia (Part d):
Rolling Window Size: 59 months
R2OS: 13.739038%
```

Figure 4: MATLAB Output for Part (d)

Interpretation

By using a 59-month rolling window for parameter estimation, the predictive regression model achieved an out-of-sample R-squared of 13.74% relative to the historical average. This significantly positive R2OS indicates that the rolling window model provided substantially more accurate out-of-sample forecasts for Australian equity premium than the historical average over the 1985:01-2010:12 period, resulting in a 13.74% reduction in MSFE. This suggests that allowing the model parameters to adapt over time using a rolling window is beneficial for forecasting in this context.

Part (e): Comparison of Rolling Window Model with Historical Average

Objective

Formally compare the out-of-sample performance of the rolling window model (from Part d) with the historical average forecast using the Diebold-Mariano test and interpret the findings.

Comparison Statistics

MSFE (Rolling Window Model): 20.204741 MSFE (Historical Average): 23.422809 Out-of-Sample R-squared (Rolling vs HA): 13.739038%

Diebold-Mariano Test (vs HA)

DM Statistic: 1.521861 P-Value: 0.128044

Screenshot of Part (e) Results

```
% Results for Part (e) %
Comparison of Rolling Window Model vs Historical Average (Part e):
MSFE (Rolling Window Model): 20.204741
MSFE (Historical Average): 23.422809
Out-of-Sample R-squared (Rolling vs HA): 13.739038%
Diebold-Mariano Test (vs HA):
  DM Statistic: 1.521861
  P-Value: 0.128044
```

Figure 5: MATLAB Output for Part (e)

Interpretation

The comparison of MSFE and the positive R2OS clearly show that the rolling window model outperforms the historical average in terms of forecast accuracy. The Diebold-Mariano test provides a statistical assessment. With a DM statistic of 1.52 and a p-value of 0.128, we fail to reject the null hypothesis of equal predictive accuracy at a standard 5% significance level. Although the rolling window model has a lower MSFE, this difference is not statistically significant at the 5% level according to the DM test. This suggests that while the rolling window model is better in terms of point forecasts (lower MSFE), the evidence for its statistical superiority over the historical average is not strong enough to reject the null hypothesis at the 5% level.

Part (f): Comparison of Rolling Window Model with Recursive Model

Objective

Compare the out-of-sample performance of the rolling window model (from Part d) with the recursive predictive regression model (from Part b) using the MSFE ratio and the Diebold-Mariano test, and interpret the findings.

Comparison Statistics

MSFE (Rolling Window Model): 20.204741 MSFE (Recursive Model from Part b): 21.943879
MSFE Ratio (Rolling / Recursive): 0.920746

Diebold-Mariano Test (vs Recursive)

DM Statistic: 1.245437 P-Value: 0.212971

Screenshot of Part (f) Results

```
% Results for Part (f) %
Comparison of Rolling Window Model vs Recursive Model (Part f):
MSFE (Rolling Window Model): 20.204741
MSFE (Recursive Model from Part b): 21.943879
MSFE Ratio (Rolling / Recursive): 0.920746
Diebold-Mariano Test (vs Recursive):
  DM Statistic: 1.245437
  P-Value: 0.212971
```

Figure 6: MATLAB Output for Part (f)

Interpretation

Comparing the rolling window model to the recursive model, the rolling window approach yields a lower MSFE (20.20 vs 21.94), resulting in an MSFE ratio of 0.92. This indicates that the rolling window model is more accurate in terms of MSFE. The Diebold-Mariano test for this comparison gives a DM statistic of 1.25 and a p-value of 0.213. At a 5% significance level, we fail to reject the null hypothesis of equal predictive accuracy. Similar to the comparison with the historical average, while the rolling window model has lower MSFE, the statistical evidence for its superiority over the recursive model is not statistically significant at the 5% level based on the DM test. This could suggest that, for this dataset and model, the benefits of parameter adaptation in the rolling window are not statistically distinguishable from the recursive approach over this specific out-of-sample period, or that the power of the DM test is limited in this context.