Internet Appendix to "Equity Premium Forecasts Tend to Perform Worse Against a Buy-and-Hold Benchmark"

Appendix A: Implementation details for Rapach, Strauss, and Zhou (2010)

Rapach, Strauss, and Zhou (2010) predict logarithmic excess returns for the computation of forecast errors, but they do not explicitly describe which forecasts they use for the utility analysis. In their formulas for the forecast error metrics and the utility gains, the authors use the same symbol for predicted returns ($\hat{r}_t$).

Nevertheless, because the argument of the utility function is simple rather than logarithmic returns, I determine optimal portfolio weights with a combination forecast and means that are based on simple rather than logarithmic excess returns. This brings the utility gains in my replication closer to the ones reported in the paper.

Two coauthors of Rapach, Strauss, and Zhou (2010) have published code[1] that implements combination forecasts for another paper (Rapach and Zhou (2013))[2]. In this code, the trailing variance used in the computation of portfolio weights is estimated with the population variance (i.e., with division by T):[3]

```
FC_VOL(t)=mean(Y(R+P_0+(t-1)-window_VOL+1:R+P_0+(t-1)).^2)-...
    (mean(Y(R+P_0+(t-1)-window_VOL+1:R+P_0+(t-1))))^2;
```

whereas the variance of realized portfolio returns is determined with division by $T-1$:[4]

```
avg_utility=mean(return_portfolio)-0.5*gamma_MV*(std(return_portfolio))^2;
```

I follow this choice because it brings the utility gains in the replication closer to the ones reported in the paper that I replicate. The conclusions of the present paper do not change if division by $T-1$ is used for both variances.

---

[1] http://apps.olin.wustl.edu/faculty/zhou/HEF_2013_data_programs.zip

[2] Rapach, D. and G. Zhou, 2013, Forecasting stock returns. In Handbook of Economic Forecasting, Elsevier, Vol. 2, 328-383.

[3] From file: *Forecasts_quarterly.m*

[4] From file: *Perform_asset_allocation.m*

Appendix B

Table B1: Revisiting the performance of three- to twelve-month ahead forecasts of the US equity premium from Rapach, Ringgenberg, and Zhou (2016)

The table extends the analysis from Table 2 to the three-month, six-month, and 12-month horizons. The results are based on a replication of the analysis performed in Rapach, Ringgenberg, and Zhou (2016). The forecast is a regression-based prediction with a short interest index. Utility gains Δu indicate the per annum gain of a mean-variance investor who switches from a benchmark strategy to one that optimizes the S&P 500 weight based on the combination forecast; Sharpe ratio differences are determined for the same set of strategies and benchmarks. Benchmarks are based on (i) the trailing mean computed from the model sample beginning in 1947 (as in the original paper) and (ii) a buy-and-hold strategy. P-values are two-sided. The out-of-sample period in the paper and the replication is 1990:01 to 2014:12.

Interpretation: With the alternative benchmark, none of the performance metrics are statistically significant. Testing the statistical significance in utility gains or Sharpe ratios can cast a different light on results. Performance differences that may appear large can lack statistical significance.

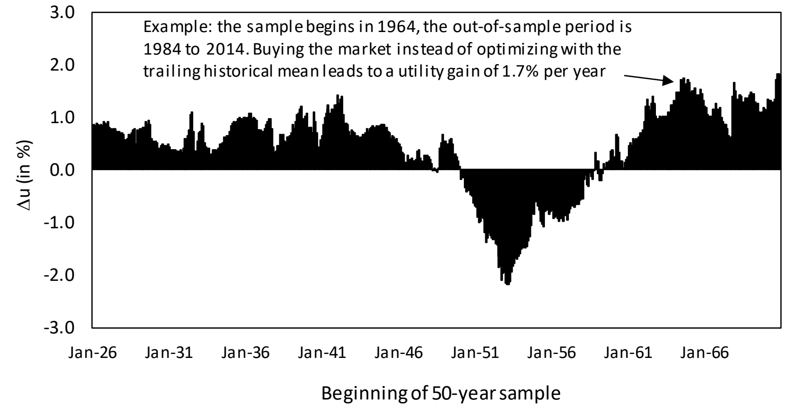
Benchmark:	Optimized with Trailing Mean		Buy-and-Hold	
	Original paper	Replication	Original Paper	Replication
	Panel A: Three-month forecast horizon			
Utility Gain Δu (in %)	4.65	4.65	2.06	2.06
p from delta method test	-	(0.023)	-	(0.380)
p from t-test on realized Δu	-	(0.046)	-	(0.389)
ΔSharpe Ratio	0.281	0.281	0.118	0.118
p from Jobson–Korkie test	-	(0.023)	-	(0.404)
	Panel B: Six-month forecast horizon			
Utility Gain Δu (in %)	5.44	5.44	3.18	3.18
p from delta method	-	(0.015)	-	(0.168)
p from t-test on realized Δu	-	(0.038)	-	(0.294)
ΔSharpe Ratio	0.353	0.353	0.199	0.199
p from Jobson–Korkie test	-	(0.027)		(0.217)
	Panel C: Twelve-month forecast horizon			
Utility Gain Δu (in %)	3.43	3.43	1.39	1.39
p from delta method	-	(0.139)	-	(0.635)
p from t-test on realized Δu	-	(0.115)	-	(0.491)
ΔSharpe Ratio	0.179	0.179	0.077	0.077
p from Jobson–Korkie test	-	(0.121)	-	(0.600)

Figure B1: Performance of the buy-and-hold strategy relative to an optimizing strategy based on the historical mean—for different 50-year sample periods and a risk aversion of five

The graphs show economic gains when switching from a strategy that optimizes the S&P 500 weight based on the historical mean to the buy-and-hold strategy. Utility gains $\boldsymbol{\Delta u}$ indicate the per annum gain of a mean-variance investor who switches from the optimizing strategy to the buy-and-hold strategy; Sharpe ratio differences are determined for the same switch. The analysis is conducted for different sample periods, each having a length of 50 years. The out-of-sample estimation period starts 20 years after the respective sample begins and lasts 30 years. The return frequency is monthly. Risk aversion is set to five, the variance of excess returns needed for the optimization is estimated with a five-year rolling window, and optimal equity weights are constrained to between 0% and 150%.

Interpretation: In the majority of the considered samples, the buy-and-hold strategy would have led to a better risk-adjusted performance than the optimizing strategy. Compared to the results in Figure 1, utility gains are lower because the buy-and-hold portfolio is too aggressive for a risk aversion of five; Sharpe ratio differences are slightly different from the ones in Figure 1 because the probability that the equity weight constraint becomes binding varies with the risk aversion.

Panel A: Δu = Utility from buy and hold minus utility from optimizing strategy



Panel B: ΔSharpe ratio = Sharpe ratio of buy and hold minus Sharpe ratio of optimizing strategy

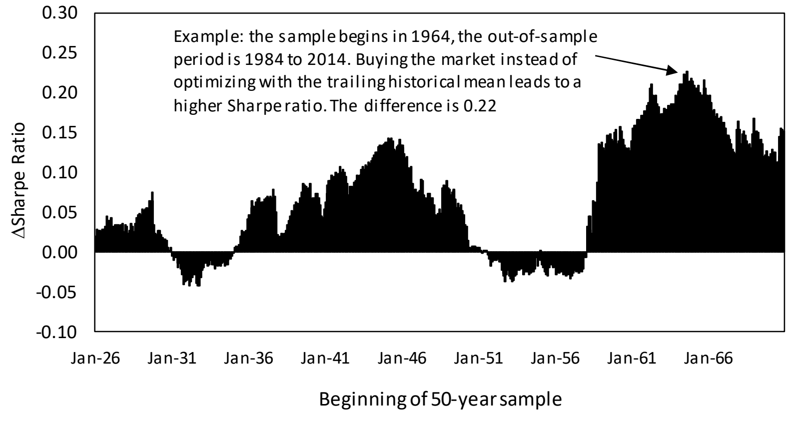
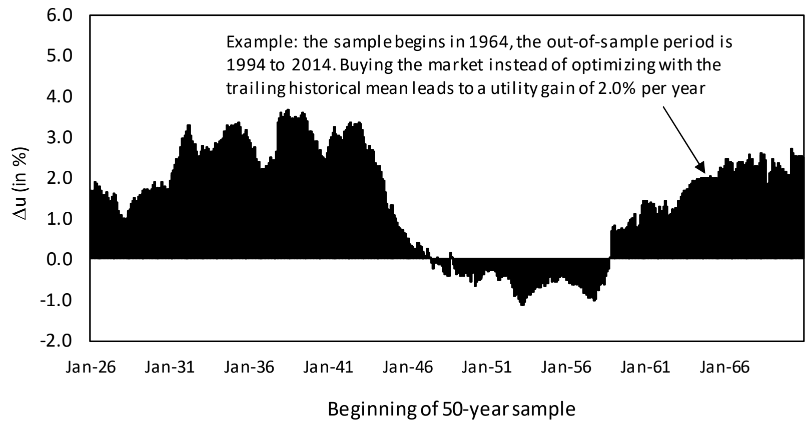


Figure B2: Performance of the buy-and-hold strategy relative to an optimizing strategy based on the historical mean—for different 50-year sample periods and an initial estimation period of 30 years

The graphs show economic gains when switching from a strategy that optimizes the S&P 500 weight based on the historical mean to the buy-and-hold strategy. Utility gains $\boldsymbol{\Delta u}$ indicate the per annum gain of a mean-variance investor who switches from the optimizing strategy to the buy-and-hold strategy; Sharpe ratio differences are determined for the same switch. The analysis is conducted for different sample periods, each having a length of 50 years. The out-of-sample estimation period starts 30 years after the respective sample begins and lasts 20 years. The return frequency is monthly. Risk aversion is set to three, the variance of excess returns needed for the optimization is estimated with a five-year rolling window, and optimal equity weights are constrained to between 0% and 150%.

Interpretation: In the majority of the considered samples, the buy-and-hold strategy would have led to a better risk-adjusted performance than the optimizing strategy.

Panel A: Δu = Utility from buy and hold minus utility from optimizing strategy



Panel B: ΔSharpe ratio = Sharpe ratio of buy and hold minus Sharpe ratio of optimizing strategy

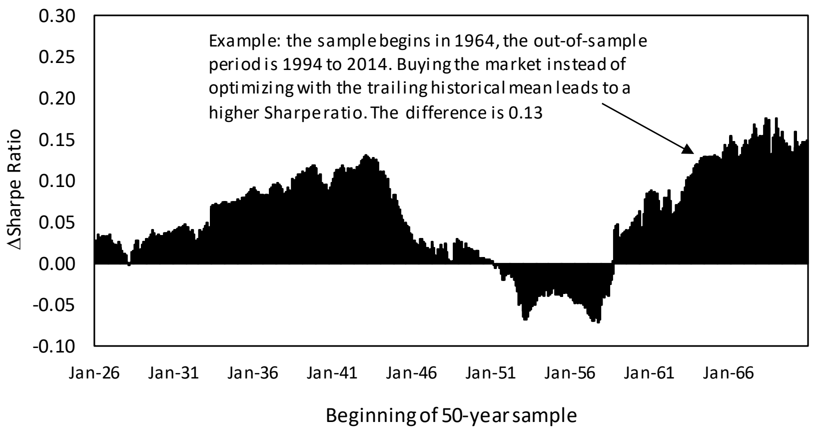
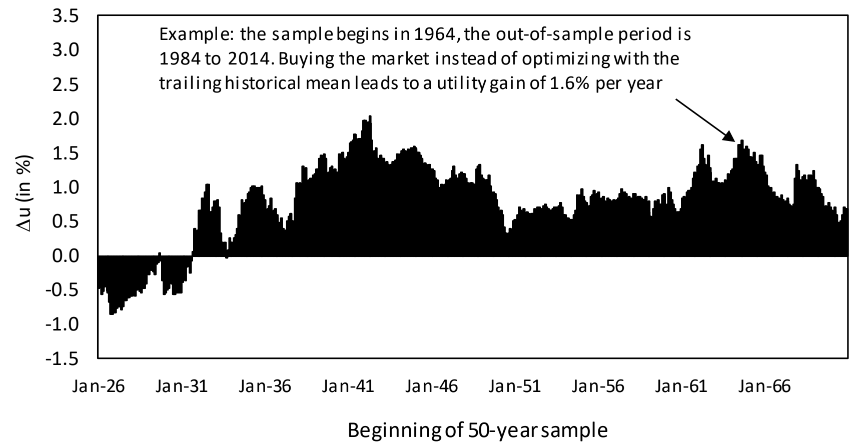


Figure B3: Performance of the buy-and-hold strategy relative to an optimizing strategy based on the historical mean—for different 50-year sample periods and with variance set to a constant

The graphs show economic gains when switching from a strategy that optimizes the S&P 500 weight based on the historical mean to the buy-and-hold strategy. Utility gains Δu indicate the per annum gain of a mean-variance investor who switches from the optimizing strategy to the buy-and-hold strategy; Sharpe ratio differences are determined for the same switch. The analysis is conducted for different sample periods, each having a length of 50 years. The out-of-sample estimation period starts 20 years after the respective sample begins and lasts 30 years. The return frequency is monthly. Risk aversion is set to three, optimal equity weights are constrained to between 0% and 150%, and the variance of excess returns needed for the optimization is set to 0.1875%, corresponding to an annual volatility of 15%.

Interpretation: In the majority of the considered samples, the buy-and-hold strategy would have led to a better risk-adjusted performance than the optimizing strategy.

Panel A: Δu = Utility from buy and hold minus utility from optimizing strategy



Panel B: ΔSharpe ratio = Sharpe ratio of buy and hold minus Sharpe ratio of optimizing strategy

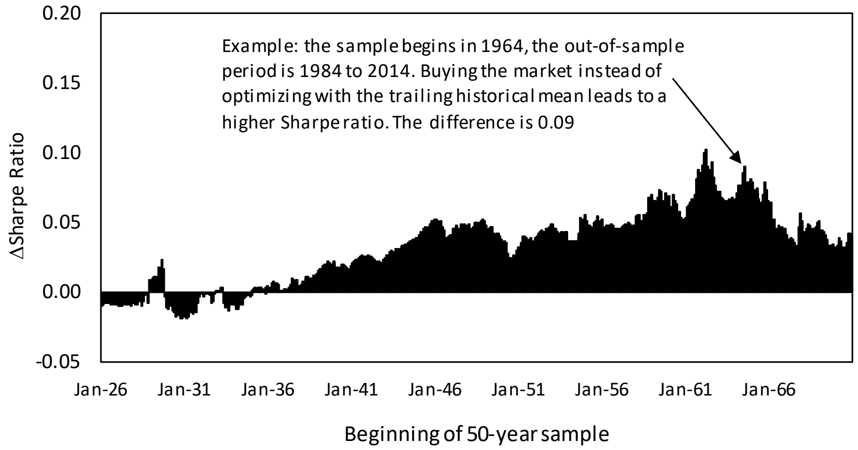
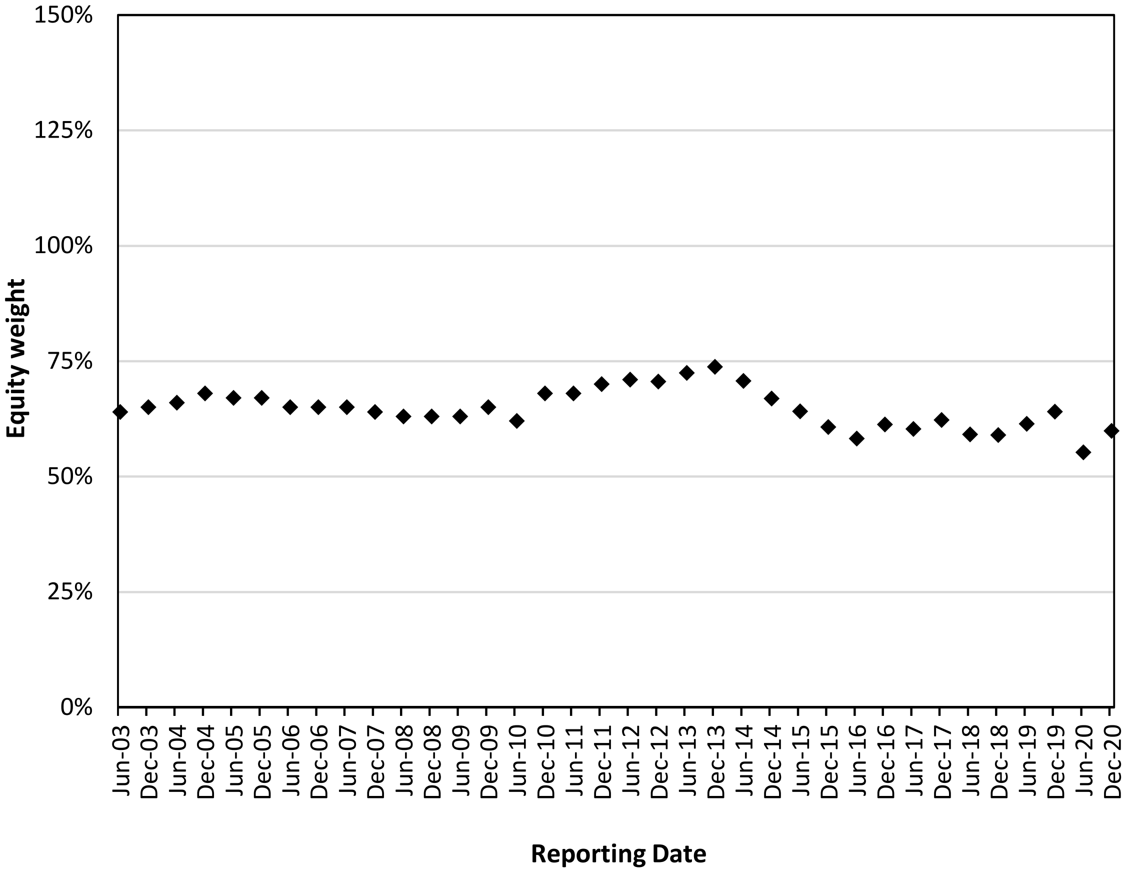


Figure B4: Equity weights of an exemplary mutual fund

For the time period from 2003 to 2020, for which the fund's reports are available on the Edgar system of the Securities and Exchange Commission (SEC), the graph shows the reported equity weights of the mutual fund American Balanced Fund. As of June 2018, American Balanced Fund was the largest US balanced fund.[5] The equity share for American Balanced Fund is collected from the fund's SEC filings N-CSR (annual) and N-CSR (semi-annual). During these years, American Balanced Fund did not report positions in equity futures or options.

Interpretation: The equity weight of a representative asset allocation fund fluctuates in a range that is considerably smaller than the 0% to 150% range that the predictability literature often chooses to make its benchmark weights realistic.



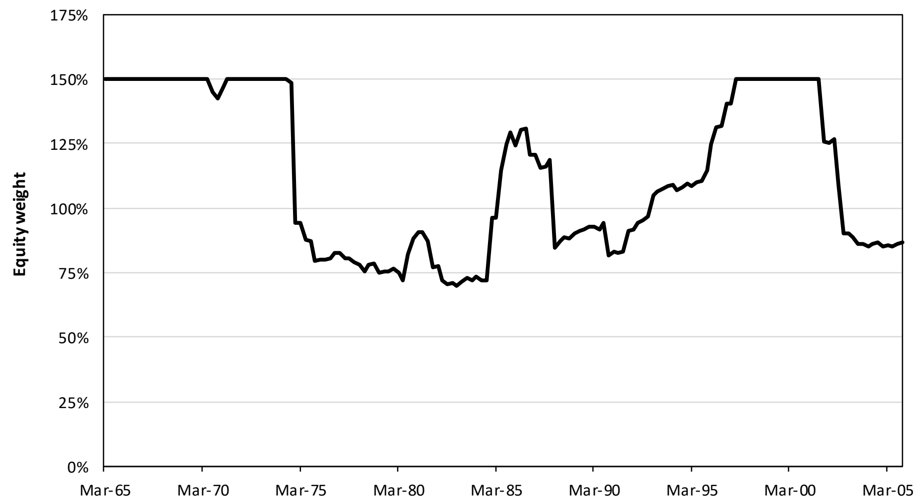
[5] Source: Pension&Investments (https://www.pionline.com/article/20181029/INTERACTIVE/181029977/the-largest-balanced-asset-allocation-mutual-funds-most-used-by-dc-plans)

Figure B5: Equity weights of the optimizing historical-mean benchmark in the out-of-sample periods of Rapach, Strauss, and Zhou (2010) and Rapach, Ringgenberg, and Zhou (2016).

Using the out-of-sample periods and assumptions made in the original papers, I determine the equity weight of the optimizing historical-mean benchmark. For Rapach, Strauss, and Zhou (2010), assumptions are: Trailing mean estimated since 1947:1; variance estimated using a rolling ten-year window; risk aversion $\gamma = 3$; weights limited to between 0 and 1.5. For Rapach, Ringgenberg, and Zhou (2010), assumptions are: Trailing mean estimated since 1973:12; variance estimated using a rolling ten-year window; risk aversion $\gamma = 3$; weights limited to between -0.5 and 1.5

Interpretation: The equity weights in historical-mean benchmarks of predictability studies can vary considerably.

Panel A: Equity weights of the historical-mean benchmark used in Rapach, Strauss, and Zhou (2010)



Panel B: Equity weights of the historical-mean benchmark used in Rapach, Ringgenberg, and Zhou (2016)

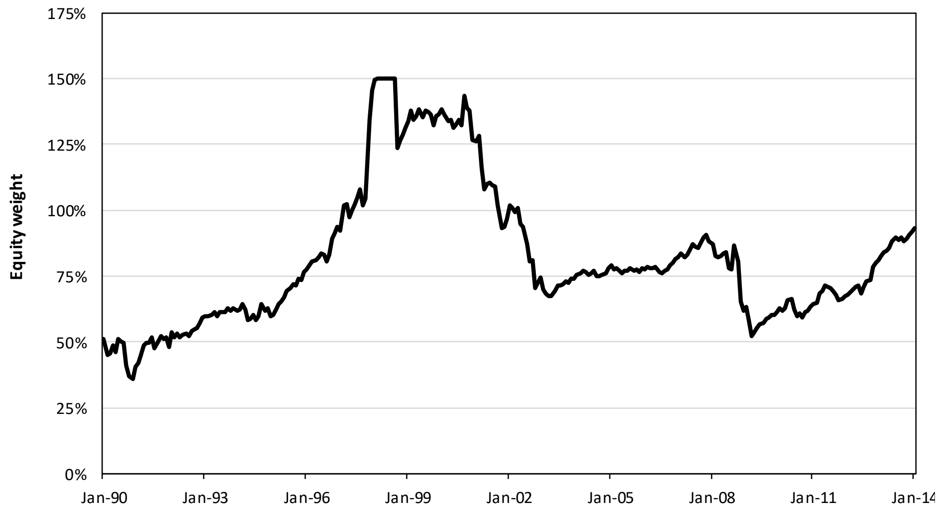
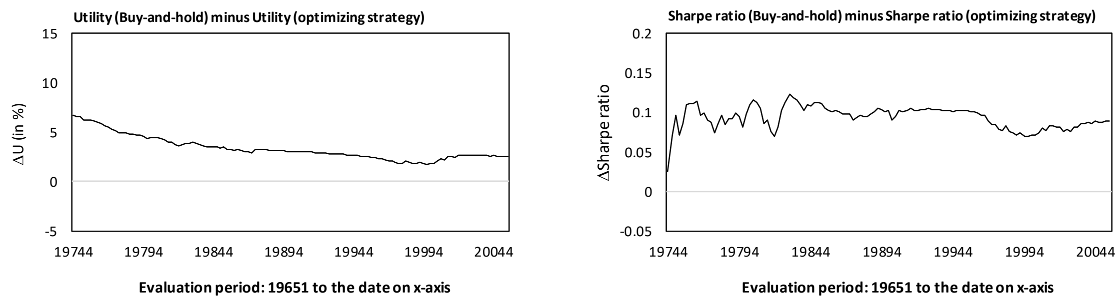


Figure B6: Performance of the buy-and-hold strategy relative to an optimizing strategy based on the historical mean—for expanding windows within the out-of-sample periods of the two replicated papers

The graphs show economic gains when switching from a strategy that optimizes the S&P 500 weight based on the historical mean to the buy-and-hold strategy. Utility gains $\boldsymbol{\Delta u}$ indicate the per annum gain of a mean-variance investor who switches from the optimizing strategy to the buy-and-hold strategy; Sharpe ratio differences are determined for the same switch. The analysis is conducted for expanding windows within the out-of-sample periods of two papers, using the return frequency and parameter assumptions made in these papers.

Interpretation: The superior performance of the buy-and-hold strategy is not only visible ex post; it would have also been visible to the imaginary investors in the out-of-sample periods.

Panel A: Economic gains from switching to the buy-and-hold strategy in the out-of-sample period of Rapach, Strauss, and Zhou (2010)



Panel B: Economic gains from switching to the buy-and-hold strategy in the out-of-sample period of Rapach, Ringgenberg, and Zhou (2016)