Methods for Optimization, Sparsification, and Timing of US and Asian Equity Portfolios

Kelvin Sun, Peter Emanuel

Abstract

We develop a process for constructing and timing entry into US and Asian equity portfolios with optimization, regression, and time series forecasting methods. We first fine-tune the traditional Markowitz mean-variance optimization framework with a semidefinite programming modification to the stocks' covariance matrices and an l_2-norm measure of risk in the optimization problem. Next, we form a cross-sectional regression model that decomposes the Asian portfolio's return by country. We then sparsify the Asian portfolio by first removing stocks from the less influential countries, as per the cross-sectional regression model output, and then further sparsify what remains with a lasso regression model. We also sparsify the US portfolio once with a lasso regression model. Finally, we use an AR(1) model to forecast the sparse Asian portfolio's equity risk premium over an example carry trade. With this AR(1) model, we backtest a timing strategy for entering into the sparse Asian portfolio when the premium over the carry trade is forecast to be large enough to justify the excess transaction costs and cost of shorting. We compare cumulative performance of a portfolio switching strategy as suggested by the AR(1) model backtest and as suggested by another round of Markowitz mean-variance optimization at the two-portfolio level. In conclusion, we find that there is significant improvement in performance of the sparse US portfolio, given the same level of risk and realistic levels of turnover, if we incorporate the sparse Asian portfolio with the AR(1) timing strategy.

1 Introduction

1.1 Motivation

A domestic investor would naturally seek to find investment opportunities in foreign assets, especially those of emerging markets, due to their potential for high growth. A well-diversified portfolio of foreign stocks can, in theory, achieve high returns and eliminate idiosyncratic risk. We take particular interest in Asian markets, as they have relatively robust stock exchanges and have seen spectacular economic growth over the past half-century[1].

One popular way to gain exposure to emerging market growth is through the carry trade. For example, an investment bank may borrow money in USD, exchange that USD into SGD, lend out the SGD, and lock in a forward rate to exchange SGD back into USD at some time in the future. A risk-free profit can be made if the interest rate differential between Singapore and the US is high enough to overcome the difference paid between USD/SGD spot and forward rates. If the forward rate is priced fairly, an investor may still profit off carry without locking in a forward rate, so long as the interest rate differential earned exceeds the potential FX loss at the future spot rate.

However, returns on carry trades, especially in emerging markets, are known to suffer catastrophic losses; carry trades exhibit high tail risk[2]. In fact, it may be said that the carry trade return is compensation for tail risk, and not for volatility, as the return tends to not be as volatile as the return on a typical portfolio of stocks. However, exposure to Asia through a portfolio of stocks alone may be preferred for an investor who does not like tail risk or does not want to venture into the currency markets. Said portfolio may be more volatile than a carry trade, but a Markowitz-style mean-variance

[1] International Monetary Fund. "Growth in East Asia: What We Can and What We Cannot Infer." Economic Issues, 1996.

[2] Dupuy, Philippe. "The Tail Risk Premia of the Carry Trades." Journal of International Money and Finance, vol. 59, 2015, pp. 123–145.

optimization[3] to that portfolio can mitigate this issue. Markowitz's framework gives an investor the ability to find his or her personal optimal portfolio, which maximizes return subject to his or her own unique maximum volatility tolerance.

1.2 Description of Dataset and Preprocessing

We use a dataset containing 5,865 daily prices on the 505 stocks in the S&P 500 Index from January 4th, 2000 to March 11th, 2022. We look at three countries to select stocks for our Asian portfolio. We use prices of the 67 stocks in Hong Kong's Hang Seng Index, the 30 stocks in Singapore's Straits Times Index, and the 225 stocks in Japan's Nikkei 225 Index over the same time period. We take the index constituent snapshots on March 11th, 2022. We use the exchange rate of USD/HKD, USD/SGD, and USD/JPY over the same time period to construct carry trades, along with a "risk-free" measure of borrowing and lending for each country. These are the 3-month SOFR for the US, the 3-month HIBOR for Hong Kong, the 3-month SIBOR for Singapore, and the 3-month TIBOR for Japan. We use Bloomberg-estimated values of these financing rates during the times before they existed and actual values thereafter.

We front-fill stock price missing values. A stock is not considered for use in a portfolio before it existed or its pricing data is available; we do not back-fill missing values at the beginning of the dataset. We do not standardize data for the lasso regressions, as the stock returns and variances are similar enough as to not adversely affect the sparsification procedure when not standardized.

2 Construction of US and Asian Optimal Portfolios

2.1 Returns and Covariance Matrices

We determine optimal portfolio weights monthly and separately for two portfolios, the US and the three-country Asian portfolio. At month t, we take simple monthly returns, from the end of month *t-1* to the end of month t, for stocks that existed in month *t-2*. We take rolling 3-month covariance matrices from the start of month *t-2* to the end of month t, excluding the first 2 months of data. We invest in stocks in month *t+1* with the portfolio weights found at month t. Therefore, we have 264 months worth of returns and covariance matrices, 264 months of weighting schema for both US and Asia, and 263 months of returns to backtest the portfolios.

We find that the vast majority of covariance matrices we form in this first round of optimization are not positive semidefinite; the size of the dataset and small return values cause numerical instability. We require a covariance matrix V_t to be positive semidefinite so that the backtested portfolio variance ${x_t}^T V_{t+1} x_t$ is nonnegative given a stock weight vector x_t. To alleviate this problem, we transform each observed covariance matrix V_t into a positive definite covariance matrix W_t with the following program:

$$\begin{aligned} \min \quad & \|V_t - W_t\|_2 \\ \text{s.t.} \quad & W_t \succeq 0.01 \end{aligned} \tag{1}$$

This program produces sparse, positive definite covariance matrices W_t with no eigenvalue less than or equal to 0.01.

2.2 Portfolio Optimization

We generate optimal portfolio weight vector x_t to invest in stocks in month *t+1* for both the US and Asian stock universes separately. We define e to be a vector of ones with length equal to the number of stocks used in each monthly optimization round. We use l_2-norm as a measure of risk for computational efficiency. We attempt to produce weights that maximize portfolio return across i stocks, using individual stock monthly returns r_t, with the following program:

[3] Markowitz, Harry. "Portfolio Selection." The Journal of Finance, vol. 7, no. 1, 1952, pp. 77–91.

$$
\begin{aligned}
\max_{x} \quad & {r_t}^{\mathrm{T}} x_t \\
\text{s.t.} \quad & {e_{i,t}}^{\mathrm{T}} x_t = 1 \\
& \|W_t x_t\|_2 \leq 0.01
\end{aligned} \tag{2}
$$

Note that we allow short positions in stocks, so long as the weights sum to 1. The l_2-norm ceiling of 0.01 is not to be confused with the minimum eigenvalue floor of 0.01 in (1). The ceiling in (2) may be chosen arbitrarily to reflect volatility tolerance; we tend to produce portfolios that have low volatility for the sake of comparison to carry trades.

3 Sparse Replication of the US Portfolio

3.1 Lasso Regression Model for the Sparse US Portfolio

Once we formulate the full version of the optimal portfolio for the US S&P 500 Index, which has 505 stocks by the end of the dataset, we find that weights generated are not sparse, highly variable, and the monthly returns are consistently slightly positive. We cannot realistically construct a portfolio that alternates the weights every month too dramatically. Trading costs are simply too high.

We propose two solutions to overcome this. The first solution is to constrain how much each stock is allowed to change their weight in the portfolio every month. We decide to not proceed with this solution because it will be too computationally intensive and will too tightly constrain (2). The second solution is to find a sparse set of stocks within the S&P 500 Index that replicates the performance of the optimal portfolio with lasso regression. We seek to produce a sparse portfolio of 40 stocks or fewer, as we believe significant monthly rebalancing of a portfolio of this size or smaller is feasible.

We apply lasso regression with the monthly returns of the US portfolio as the response variable and the monthly returns of 505 individual stocks as the features, with the following objective function:

$$
\min_{\beta} \quad \sum_{t=1}^{263} (r_{i,t} x_{i,t\text{-}1} - \beta_0 - \sum_{i=1}^{505} (\beta_i r_{i,t}))^2 + \lambda \sum_{i=1}^{505} |\beta_i| \tag{3}
$$

We assign zeroes to the monthly returns of the stocks at month t that have not joined the index at each month *t-2*. We choose a regularization parameter λ of 0.01, producing 40 stocks that have non-zero coefficients. We believe that this subset of stocks captures the structure of the index to a certain degree, but since the number of stocks is greater than the number of months in the dataset, (3) may not have a unique solution. We then take the sparse set of stocks and produce a new sparse US portfolio, re-optimizing through (2). We need not apply the procedure from (1) to the covariance matrices of the sparse US portfolio because these covariance matrices are already positive semidefinite. The stocks that have not joined the index at each month *t-2* would not be included in the weighting scheme of month t and their entries in the covariance matrix in month *t+1* are 0. By doing so, we arrive at a new, sparse, optimized portfolio with their weights summing to 1.

The newly optimized portfolio has correlation of 0.09 to the full optimized portfolio, which is not outstanding. However, we achieve a 74% same-sign accuracy on monthly returns and an old versus new portfolio return MSE of 0.000166. Note that this is not the model MSE, rather, it is the MSE of the two already-optimized portfolio returns series.

4 Sparse Replication of the Asian Portfolio

4.1 Cross-Sectional Regression Model for Return Attribution by Country

We seek to perform a similar sparsification procedure to the Asian portfolio. Once sparsified, we modify covariance matrices through (1) and re-optimize through (2). As a precursor step, we first

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.033
Model:                            OLS   Adj. R-squared:                  0.027
Method:                 Least Squares   F-statistic:                     5.433
Date:                Sun, 15 May 2022   Prob (F-statistic):            0.00479
Time:                        17:19:05   Log-Likelihood:                 1116.0
No. Observations:                 319   AIC:                            -2226.
Df Residuals:                     316   BIC:                            -2215.
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
HK             0.0064      0.001      6.978      0.000       0.005       0.008
SP             0.0024      0.001      1.785      0.075      -0.000       0.005
JT             0.0032      0.000      6.447      0.000       0.002       0.004
==============================================================================
Omnibus:                      210.804   Durbin-Watson:                   1.895
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2030.178
Skew:                           2.671   Prob(JB):                         0.00
Kurtosis:                      14.145   Cond. No.                         2.74
==============================================================================
```

Figure 1: An example output of (4)

decompose portfolio returns at each month by country. If one country's contribution to the portfolio monthly returns dominates the others, we may completely remove all stocks from the non-contributing countries prior to performing lasso regression. We use a cross-sectional regression model to perform the decomposition each month. The response variable is the (fixed) vector of stock weights in month *t-1* times returns in month *t*, given by $r_{i,t-1} x_{i,t}$. In other words, the response variable is a vector whose components are individual stock contributions to the overall strategy return of month *t*. The features are indicator variables on stock country of origin. We relate the features and response with the following equation:

$$r_{i,t1} x_{i,t-1} = \beta_{\mathrm{HK,i,t}} \mathbb{1}_{\mathrm{HK,i,t}} + \beta_{\mathrm{SP,i,t}} \mathbb{1}_{\mathrm{SP,i,t}} + \beta_{\mathrm{JT,i,t}} \mathbb{1}_{\mathrm{JT,i,t}} + \epsilon_{\mathrm{i,t}} \tag{4}$$

Figure 1 presents an example month of the output of (4). In this month, the Asian portfolio returned 1.2%, over half of which (0.64%) is attributed to the Hong Kong portion of the portfolio. In fact, looking at the evolution of model coefficients across time, it is clear that the Hong Kong portion of the portfolio dominates its returns, with the exception of a few months where Singapore dominates. The Japan portion of the portfolio is stable and appears to only serve to reduce volatility by construction. Figure 2 displays this fact graphically.

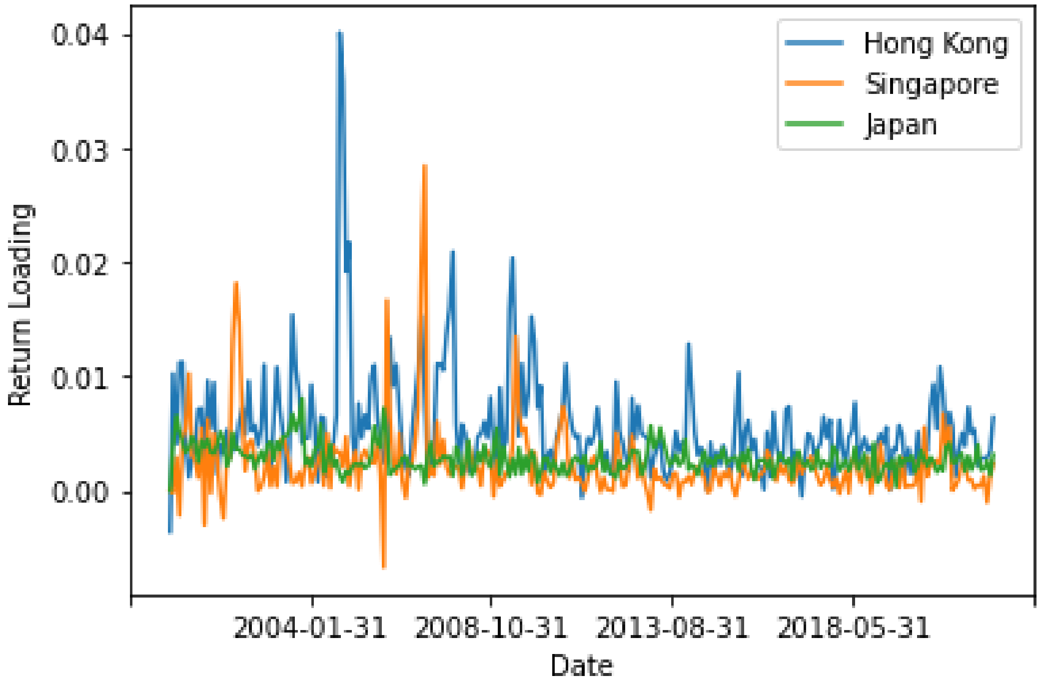

Figure 2: Time series of cross-sectional country return loadings

We conclude that the Hong Kong portion of each month's Asian portfolio, though only about 20% of this universe of Asian stocks by count, is the driving force behind the portfolio returns. Therefore, we drop all stocks from Singapore and Japan prior to performing lasso regression. The Asian portfolio is now the Hong Kong portfolio.

4.2 Lasso Regression Model for the Sparse Hong Kong Portfolio

As was also the case with the US portfolio, the Hong Kong portfolio exhibits consistent, slightly positive monthly returns and highly variable weights between months. Again, we cannot realistically construct a portfolio like this. Further sparsifying the set of 67 Hong Kong stocks while maintaining the return profile of the original Asian portfolio creates a more feasible portfolio rebalancing scenario.

To this end, we apply the same lasso regression procedure to the 67 Hong Kong stock monthly returns as was done to the US stock monthly returns in (3). We use a smaller regularization parameter λ of 0.00001 and produce a sparse set of 30 Hong Kong stocks. This sparse set of returns has covariance matrices modified in (1) and is re-optimized in (2). The resulting sparse Hong Kong portfolio monthly returns have a correlation of 0.57 to the the original Asian portfolio monthly returns with an astounding 99% same-sign accuracy and an old versus new portfolio return MSE of 0.0000332. Figure 3 compares the two lasso regression models in summary.

Sparse Portfolio	#Stocks Selected	% of Original	Correlation to Original	Same-Sign Accuracy	MSE
US	39	8%	0.09	74%	1.66E-04
Hong Kong	30	9%	0.57	99%	3.32E-05

Figure 3: Lasso regression model comparison

5 Sparse Hong Kong Portfolio Timing

5.1 AR(1) Model Forecast of Equity Risk Premium over the Carry Trade

Now that we have formed a sparse Hong Kong portfolio of stocks as our entryway into Asian equity exposure, we formulate a method for forecasting the long-run sparse Hong Kong portfolio monthly excess return to a risk-free asset. As was stated in the introduction, a popular method for gaining exposure to emerging market growth is through a carry trade. The USD/HKD carry trade is rather trivial; USD/HKD FX risk is minimal because the HKD is pegged to the USD within a band. Therefore, we discard the forward rate component and view the uncovered USD/HKD carry trade purely as an investment in the Hong Kong risk-free rate. In other words, the monthly excess return of the sparse Hong Kong portfolio over the carry trade *is* the monthly excess return of the sparse Hong Kong portfolio over the risk-free rate.

We train a series of ARIMA models on the entire 263-month return series of the difference between the sparse Hong Kong portfolio monthly returns and the USD/HKD carry trade, E_t, picking the model with the lowest BIC. The resulting model is of AR(1) with the following equation:

$$E_{t+1} = 0.0027 + 0.8111(E_t - 0.0027) + \epsilon_{\mathrm{t}} \tag{5}$$

The trend of the ACF of the residuals of E_t, the low standard error of AR(1) parameters μ (0.0027) and ϕ (0.8111), and the failure to reject the null hypothesis of the Box-Pierce test are all strong evidence of insignificant, white noise residuals, as shown in Figures 4, 5, and 6, respectively. An AR(1) model is clearly appropriate to use to forecast E_t.

We construct a 5-year-ahead forecast of E_t with (5), which converges to a long-run average of 0.68%. Figure 7 illustrates this forecast. The observed average risk premium of the full Asian portfolio over a one-third-evenly-split USD/HKD USD/SGD USD/JPY carry trade over the life of the dataset is 1.39%. Therefore, we estimate that about half of the Asian equity risk premium over the carry trade comes from just 30 Hong Kong stocks.

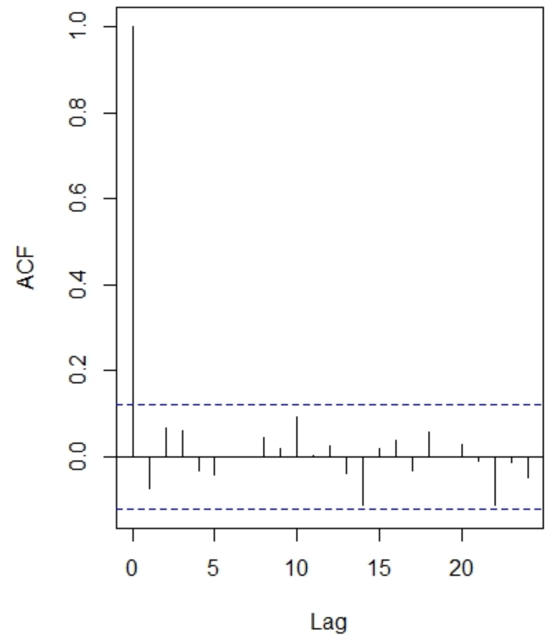

Figure 4: ACF of AR(1) model residuals

```
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1    mean
      0.8111  0.0027
s.e.  0.0356  0.0013

sigma^2 = 1.577e-05:  log likelihood = 1081.34
AIC=-2156.68   AICc=-2156.59   BIC=-2145.97
```

Figure 5: AR(1) model output

```
        Box-Pierce test

data:  E_residuals
X-squared = 12.889, df = 20, p-value = 0.8821
```

Figure 6: Box-Pierce test on AR(1) model residuals

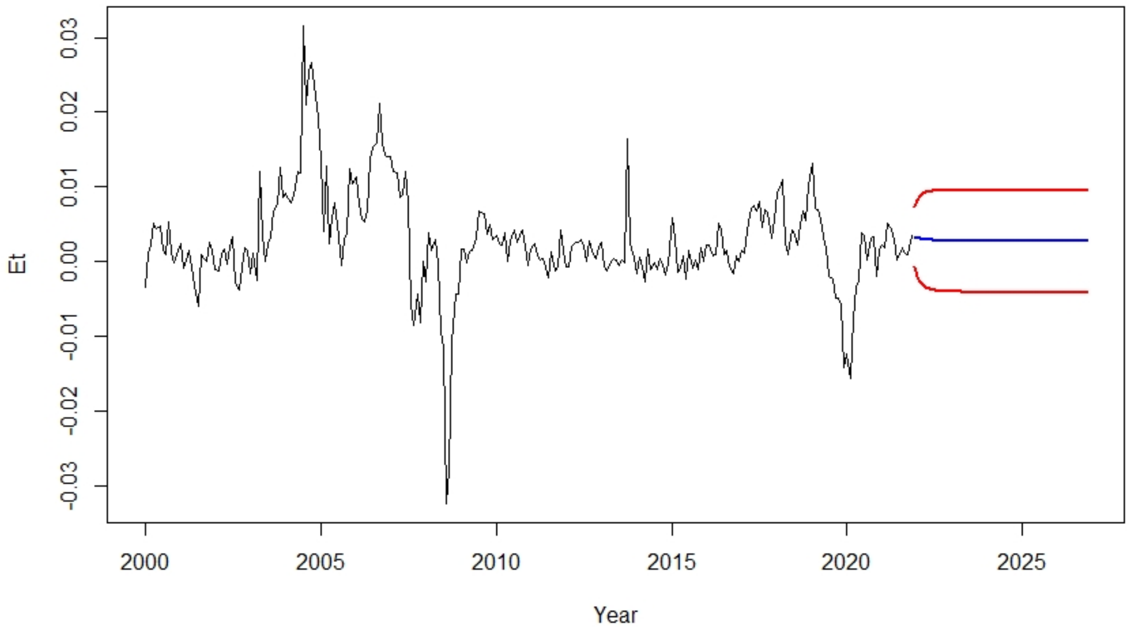

Figure 7: 5-year-ahead 95% confidence-bound AR(1) forecast of E_t

5.2 Backtest of AR(1) Model Timing Strategy

We infer, based on the fact that the best model chosen to fit $\mathrm{E_t}$ was of AR(1) and of I(0), that this time series is already stationary, and thus mean-reverting. Therefore, we may be able to time entry into the sparse Hong Kong portfolio. For example, we may purchase it at the end of month t when the AR(1) model forecasts a high enough $\mathrm{E_{t+1}}$ and sell it prior to future month $t+l$ when the AR(1) model forecasts E_{t+l+1} to decline enough. Only the observations of $\mathrm{E_t}$ known prior to the end of month t are used to fit an AR(1) model during the backtest to avoid lookahead bias. Therefore, we fit a new AR(1) model each month with more and more information about $\mathrm{E_t}$ until the AR(1) model converges to (5). We use AR(1) throughout the backtest, even if, at certain months, $\mathrm{E_t}$ would have had a lower BIC if fit with a different ARIMA model, to reduce overfitting and maintain simplicity.

We do not seek to enter into the sparse Hong Kong portfolio when its next-month-forecasted $\mathrm{E_{t+1}}$ is merely positive. Trading costs and the cost of shorting eat into this return. We assume these costs to total 0.3% each month and will only enter in to the sparse Hong Kong portfolio at the end of month t when its next-month-forecasted $\mathrm{E_{t+1}}$ is greater than 0.5%.

Figures 8 shows the monthly risk premium of the sparse Hong Kong portfolio over the USD/HKD carry trade throughout the life of the dataset. The region highlighted in green represents months where we enter into the sparse Hong Kong portfolio according to the AR(1) model forecast, which aligns well with periods of high realized $\mathrm{E_{t+1}}$. Figure 9 shows the monthly return of the USD/HKD carry trade (in this case, the risk-free asset) over the same period. The region highlighted in red also represents months where we enter into the sparse Hong Kong portfolio, which aligns well with poor USD/HKD carry trade performance. We would surely not enter in to the USD/HKD carry trade in the highlighted time periods as we would rather earn the Hong Kong equity risk premium instead.

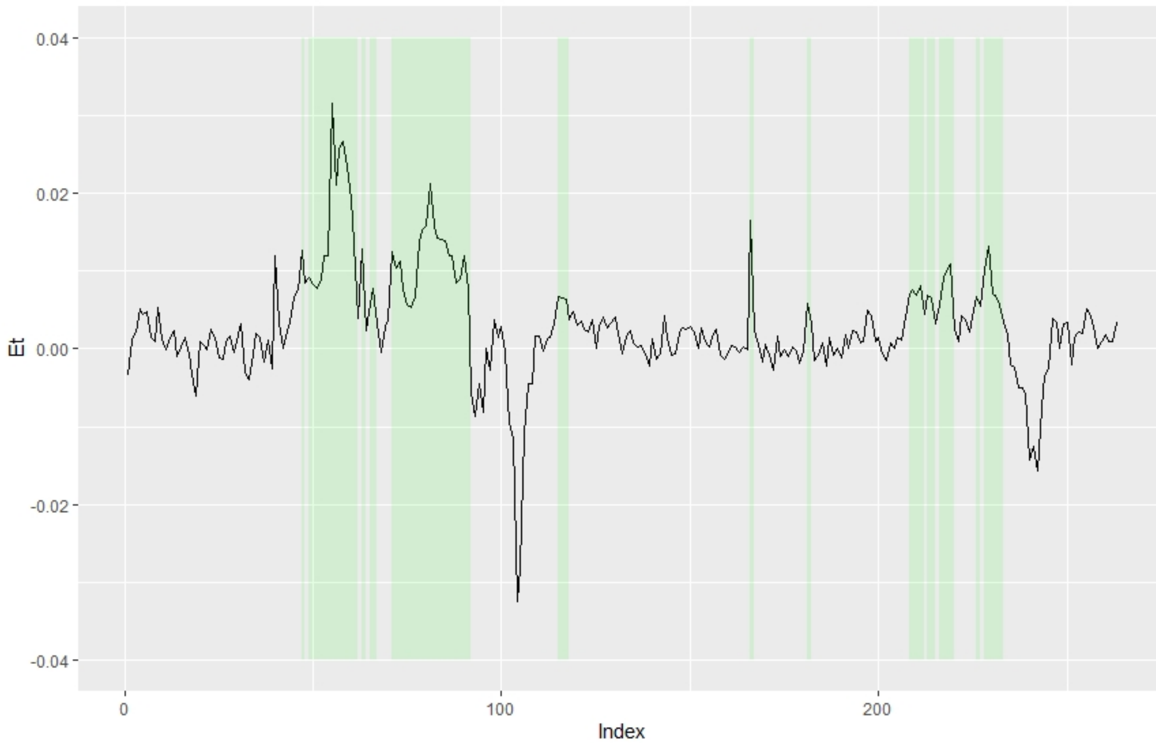

Figure 8: AR(1) model timing of E_t

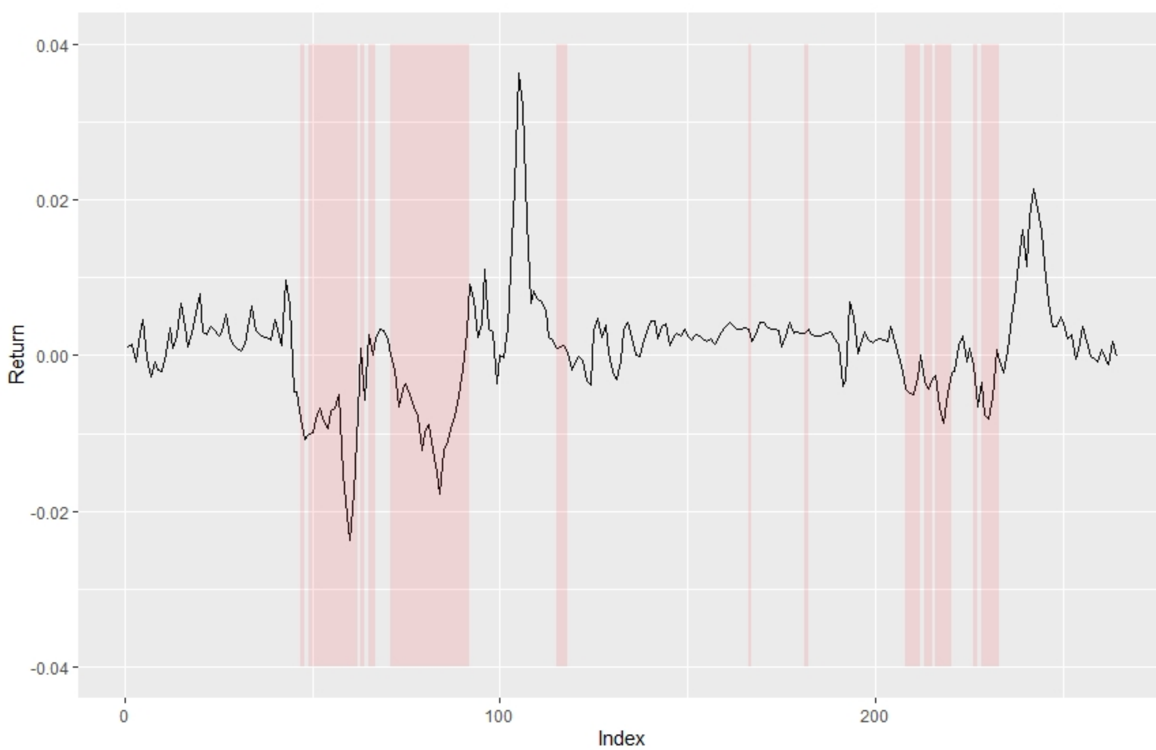

Figure 9: AR(1) model timing of Hong Kong carry trade

The refit-monthly AR(1) model appears to time most of the rises in E_t well without staying invested for too much of the subsequent falls. By the same token, the refit-monthly AR(1) model appears to enter into the sparse Hong Kong portfolio at times when the USD/HKD carry trade is about to underperform. An investor may consider shorting the Hong Kong risk-free asset to finance the purchase of the sparse Hong Kong portfolio to further enhance returns.

6 Combined Performance of Sparse US Portfolio with Timed Sparse Hong Kong Portfolio

6.1 Combined Portfolio Allocation Schema

We seek to analyze the sparse Hong Kong portfolio timing strategy further by examining the compatibility of the sparse Hong Kong portfolio and the sparse US portfolio. We first apply (2) at the portfolio level, treating each portfolio as one of two total assets. We constrain the allocations to fall between 0 and 1 to avoid unnecessary extreme weighting schema at the portfolio level; each portfolio already has plenty of shorts within. The portfolio-level covariance matrices need not be modified with (1) as they are already positive semidefinite. We call the resulting output the combined portfolio.

We find that the allocation schema generated are all boundary solutions, meaning that we either invest 100% into one portfolio or the other. We believe this to be the case because we originally set the same ceiling on volatility for each portfolio, so the one with the higher return in month t should demand the entire allocation in month $t+1$. Furthermore, the allocation schema remain highly variable.

We continue to question whether such a portfolio can be feasibly implemented. Comparing the months when the AR(1) model backtest suggests to enter into the sparse Hong Kong portfolio and the choice of the combined allocation schema, we see that almost every single AR(1) model choice belongs to the set of months where the combined allocation scheme chooses 100% to be invested in the sparse Hong Kong portfolio. However, the AR(1) backtest suggestion procudes longer holding periods, linking the gaps in the allocation schema. This gives validity to the use of AR(1) and its prediction to enter into the sparse Hong Kong portfolio, giving us a less variable schema for the combined portfolio. We choose to apply the same 100% allocation to the sparse Hong Kong portfolio when the AR(1) model backtest suggests we enter into it, keeping in line with the allocation schema (2) generates. Therefore, we conclude that the two portfolios are substitutes for one another.

6.2 Cumulative Performance of the Combined Portfolio

We assess the cumulative performance of the original US portfolio versus the combined portfolio, under both the allocation schema (2) generates and the allocation schema the AR(1) model backtest suggests. We observe in Figure 10 that the combined portfolio outperforms the US portfolio by a large margin, with a significantly more sparse set of stocks. However, since we believe the initial allocation schema to vary too highly to be feasible in practice, we examine in Figure 11 the cumulative performance difference between the AR(1) model backtest-suggested combined portfolio and the original US portfolio. We continue to see significant improvement in combined portfolio monthly performance over the long-run, with fewer switches into and out of the sparse Hong Kong portfolio.

Examining both methods of combining the sparse US portfolio and the sparse Hong Kong portfolio, we see significant improvement in portfolio performance over the original US portfolio at the same volatility level. We conclude there is evidence to support that switching between the sparse Hong Kong and the sparse US portfolios can improve performance over staying in the original US portfolio. At the very least, adding the sparse Hong Kong portfolio to one's equity holdings with the AR(1) model timing strategy proves to enhance returns without increasing volatility.

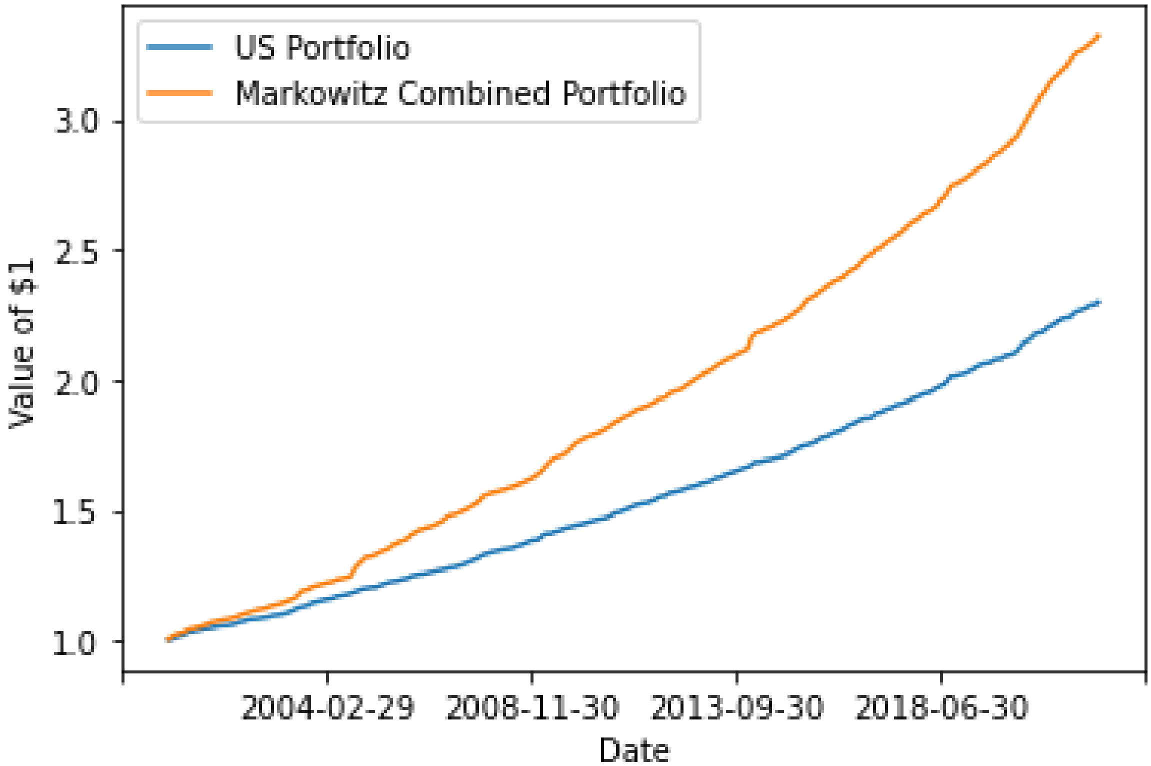

Figure 10: Cumulative performance of Markowitz combined portfolio versus original US portfolio

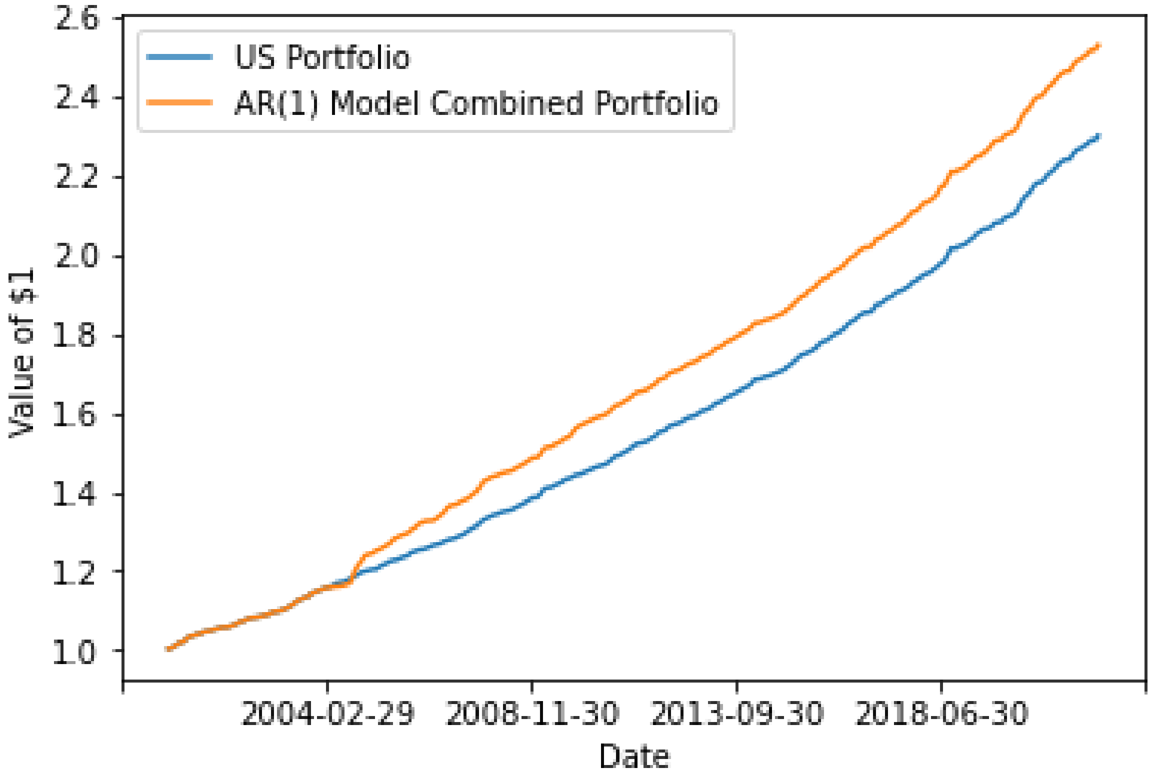

Figure 11: Cumulative performance of AR(1) model combined portfolio versus original US portfolio

7 Conclusion

We have explored the possibility of incorporating stocks from Asian stock markets into a portfolio of US stocks to improve overall portfolio performance. We used price data from 2000 to 2022 of stocks in the S&P 500 Index, the Hong Kong Hang Seng Index, the Japanese Nikkei 225 Index, and the Singaporean Straits Times Index.

We applied the Markowitz mean-variance portfolio optimization framework to the US and Asian stocks, with modifications to the covariance matrices and the measure of risk in the optimization problem. Given the difficulty of implementing portfolios containing too many stocks, we then found a sparse replication of these two optimal portfolios with lasso regression. Additionally, in constructing the optimal Asian portfolio, we found that the stocks on the Hang Seng Index dominate the monthly return contributions over the stocks on the other two exchanges, giving us reason to drop the Japanese and Singaporean stocks.

Finally, we sought to find a strategy for entering into the sparse Hong Kong portfolio and combining it with the sparse US portfolio. We attempted two methods to accomplish this goal. We used another round of Markowitz mean-variance optimization at the two-portfolio level. We also used an AR(1) model to time high-enough sparse Hong Kong portfolio equity risk premium above the USD/HKD carry trade (and/or the Hong Kong risk-free asset). However, since both sparse portfolios have similar volatility, the allocation schema always favored one portfolio entirely over the other each month. Both methods suggested similar windows of opportunity for entering into the sparse Hong Kong portfolio and also both methods outperformed the original US portfolio. The Markowitz method had slightly better cumulative performance, but the allocation schema varied too highly for implementation, while the AR(1) method was more ideal for implementation but had slightly worse cumulative performance.

In conclusion, we find significant evidence that incorporating Hong Kong stocks into an optimized US stock portfolio improves overall portfolio performance, given the same level of risk set in (2). Future works may include improving the methods of sparse portfolio replication beyond lasso, stress-testing the risk thresholds in (2) for the sparse Hong Kong portfolio to produce nontrivial allocation schema at the portfolio level, and developing an algorithm beyond AR(1) with higher-frequency data to develop a timing strategy that works intra-month.