Comparison of two different approaches to stock portfolio analysis

Filip Tošenovský[1]

Abstract. The article presents a portfolio analysis of selected shares traded on Prague Stock Exchange, using two approaches. One approach is represented by the standard procedure which defines a portfolio on the basis of optimization of the trade-off between the first and second moments of the considered share prices, with the moments being estimated from the available data by their common sample counterparts. The second approach to the analysis is based on the estimates capturing the time dynamics of the first and second moments of the share prices that are typically generated by nonstationary processes. To capture the dynamics, time series models are used. The two procedures are compared as to which of them provides an expected portfolio return that is closer to the true return.

Keywords: stock portfolio, moment estimates, time series.

JEL Classification: C13, C61
AMS Classification: 91G10, 91G70

1 Introduction

Stock market has long been in the spotlight of many investors, as it enables flexible changes in the financial structure of private entities, and provides room for investments. Investments especially boost a lot of interest in part due to theories on stock portfolio optimization. The theories try to reduce the risk resulting from the fluctuation of stock prices. There is more than one approach to achieving such objective. The procedures often use stochastic calculus, which is the approach we adopt in this article as well.

Creating a portfolio falls in the category of optimization, so that the selection of shares forming the portfolio must necessarily follow certain criteria based on which the portfolio is optimized. One should therefore pay attention to these criteria for their irrational selection may provide the portfolio owner with dissatisfactory results despite the fact that the rest of the portfolio formation was taken care of by exact mathematics. If we focus on the classical way of creating a portfolio, as proposed by Markowitz, the optimization criteria used are the unconditional expected value and variance of the portfolio yield. However, in reality their estimates must be used, which means to estimate unconditional expected values of the stock yields and covariances of these yields.

The commonly used estimates are calculated as in the case of estimation of stationary process moments. The question then is what is actually calculated by the common estimates, especially when stock prices are typically generated by nonstochastic processes. Apart from the way these estimates are obtained, their inertia is assumed as well for the time period during which the portfolio to be optimized is held by its owner. These are the reasons why we digress to other portfolio criteria in this paper – the predicted conditional expected values and variances of the stock yields with the predictions carried out using time series models. The alternative approach tries to overcome the inertia assumption, taking into account the latest dynamics of the market. We shall use both the standard and alternative approach for the portfolio optimization, and compare the yields of such portfolios.

In this paper, we shall draw the analysis on the classical Markowitz's approach to portfolio optimization. This approach aims to minimize the estimate of the portfolio unconditional variance $var[f(w_1, w_2, \ldots, w_n)] = \sum_{i,j=1}^{n} w_i w_j \, s_{ij}$ with respect to the variables w_i's. Here, $f(w_1, w_2, \ldots, w_n) = \sum_{i=1}^{n} w_i \, x_i$ represents the portfolio yield, x_i is a yield of the *i*-th stock included in the portfolio, w_i is the weight of the *i*-th stock in the portfolio and s_{ij} denotes an estimate of the covariance between the yields of the *i*-th and *j*-th stock. The estimate of the unconditional expected return of the portfolio equals $\sum_{i=1}^{n} w_i \, \bar{x}_i$, $\bar{x}_i$ being an estimate of the unconditional average yield of the *i*-th stock. The extreme of the function $var\, f$ is sought in the set *M* defined by (in)equalities: $\sum_i w_i = 1$, $w_i \geq 0$ for *i* = 1, 2,..., *n* and $\sum_{i=1}^{n} w_i \, \bar{x}_i \geq k$, where *k* is defined arbitrarily.

[1] Silesian University in Opava/School of Business Administration in Karvina, Department of Mathematical Methods in Economics, Univerzitní náměstí 1934/3, 733 40 Karviná, the Czech Republic, tosenovskyfilip@opf.slu.cz.

Generally, the problem with estimates in the portfolio optimization is as follows: if we had perfect estimates of the moments of the future stock yields, the optimization based on Markowitz would mean that we could rely on having a *decent realized* investment yield not far from the expected yield of the portfolio. This is assured by setting up the portfolio so that its riskiness/variance is minimized while a certain predefined decent level of expected yield is assured. However, the estimates will almost surely not be perfect. Thus, we minimize $\sum_{i,j=1}^{n} w_i w_j\, s_{ij}$ instead of $\sum_{i,j=1}^{n} w_i w_j\, \sigma_{ij}$ on the set M differing from the original set N by the last inequality: $\sum_{i=1}^{n} w_i\, \bar{x}_i \geq k$ instead of $\sum_{i=1}^{n} w_i\, E(x_i) \geq k$, the latter defining together with the (in)equalities $\sum_i w_i = 1$, $w_i \geq 0$ the set N. This means that instead of the correct vector of weights $\boldsymbol{w}^{opt}$ we get a vector $\breve{\boldsymbol{w}}^{opt}$. Of course, it is very likely that $\breve{\boldsymbol{w}}^{opt}$ will not optimize $\sum_{i,j=1}^{n} w_i w_j\, \sigma_{ij}$ in N - if $\breve{\boldsymbol{w}}^{opt}$ belongs to N, that is. It may happen that $\breve{\boldsymbol{w}}^{opt}$ does not belong to N at all because $\sum_{i,j=1}^{n} \breve{w}_i^{opt} E(x_i) < k$. The better the moment estimates, the more the portfolio should be under control in the sense that its *realized* yield should be closer to its expected yield.

In light of what was just said, we will now present results of an empirical study that compares the *two approaches* mentioned in the introduction. To do so, we shall draw on the data from the Prague Stock Exchange.

2 Classical approach

For the purpose of our analysis, we shall now employ stocks of five companies with the highest market value listed in the SPAD trade system. These are, as implied by the Prague Stock Exchange data, Central European Media Enterprises (CEME), ČEZ, Erste Bank, Telefónica and NWR. We will use the closing prices of the stock from the start of 2012 to April of that year. The data are at www.penize.cz. Table 1 shows some of the oldest and latest prices of the stock for the given period.

Date	CEME	CEZ	ERSTE	NWR	TELEFÓNICA
27.04.12	146.06	758.0	426.0	126.53	379.0
26.04.12	139.50	759.9	423.3	126.10	378.0
25.04.12	134.10	764.9	435.6	125.20	378.1
…	…	…	…	…	…
04.01.12	134.60	805.0	353.7	142.95	391.5
03.01.12	138.00	800.9	366.5	140.80	385.5
02.01.12	133.64	791.0	359.0	139.29	382.5

Table 1 The latest and oldest closing prices of the selected stock (in Czech crowns)

We shall assume the investor always wants to set up the portfolio for a week. More precisely, the portfolio is created on Friday, and kept until next Friday when its realized profit/loss is confronted. Then the investor sets up a new weekly portfolio, taking into account the new data on the market development. Thus a series of portfolios whose composition differs in time is created. To optimize the portfolios, we first calculate the standard estimates of the first and second moments of the stock yields, i.e. sample averages and sample variances of the five-day yields of each stock and sample covariances between the yields, using historical data. To get reasonable estimates, we start with the time series of the oldest fifty values. Thus, the first and the oldest portfolio is optimized on the estimates calculated from the data from the start of 2012 to March 9. The second portfolio uses estimates calculated from the data from the start of 2012 to March 16, and so on. Table 2 shows sample averages of the yields used for seven subsequently optimized weekly portfolios.

Week	CEME	CEZ	ERSTE	NWR	TELEF.
1.	0.003457	0.003061	0.035187	0.006552	0.003778
2.	0.007675	0.002128	0.029835	0.001891	0.001979
3.	0.005171	0.002481	0.028431	-0.00546	0.001668
4.	0.003155	0.001549	0.02236	-0.00266	0.001223
5.	0.007269	0.00062	0.014641	-0.00529	-0.00052
6.	0.003431	-0.00032	0.011721	-0.00596	-0.00201
7.	0.003148	-0.00167	0.0114	-0.00578	-0.00101

Table 2 Sample averages of five-day yields to be used successively for seven weekly portfolios

Table 3 describes sample variances of the yields for different weeks, table 4 shows the covariance structure of the yields for the first constructed portfolio (covariances for other weeks were obtained in a similar fashion).

Week	CEME	CEZ	ERSTE	NWR	TELEF.
1.	0.007939	0.000722	0.007244	0.003090	0.000811
2.	0.007477	0.000663	0.006820	0.003027	0.000763
3.	0.006895	0.000610	0.006244	0.003327	0.000699
4.	0.006410	0.000570	0.006165	0.003170	0.000643
5.	0.006200	0.000542	0.006434	0.003011	0.000654
6.	0.006095	0.000527	0.006300	0.002896	0.000670
7.	0.005690	0.000526	0.005899	0.002722	0.000643

Table 3 Sample variances of five-day yields to be used successively for seven weekly portfolios

	CEME	CEZ	ERSTE	NWR	TELEF.
CEME	0.007939	0.001483	0.004958	0.002092	-0.00065
CEZ	0.001483	0.000722	0.000449	0.000617	0.000187
ERSTE	0.004958	0.000449	0.007244	0.001194	-0.00038
NWR	0.002092	0.000617	0.001194	0.00309	0.000162
TELEF.	-0.00065	0.000187	-0.00038	0.000162	0.000811

Table 4 Sample covariances of five-day yields to be used for the first portfolio

Having the estimates, we performed the optimization in the Markowitz's sense, and we arrived at the results described in table 5. Expected return at the optimum for each week is on the far left of the table, realized return is the return the investor actually registered during the five-day period of holding the newest portfolio. Absolute difference of the two returns is contained in the third column of the table. The last five columns portray the portfolio weights at the optimum. As one can see, some of the differences in returns are quite severe.

Expected Return	Realized Return	Abs. Difference	w(1)	w(2)	w(3)	w(4)	w(5)
0.003	-0.0142	0.0174	0.000	0.495	0.045	0.033	0.428
0.010	0.0004	0.0096	0.000	0.182	0.412	0.000	0.406
0.010	-0.0284	0.0384	0.000	0.170	0.440	0.000	0.390
0.010	-0.0797	0.0897	0.000	0.000	0.594	0.000	0.406
0.005	0.0110	0.0060	0.048	0.341	0.484	0.000	0.127
0.005	0.0181	0.0131	0.000	0.416	0.584	0.000	0.000
0.005	0.0310	0.0260	0.000	0.000	0.615	0.000	0.385

Table 5 Portfolio optimization based on standard moment estimates, and the true development of the market

3 Alternative approach

We shall now employ the alternative procedure in the portfolio analysis. We will not assume for each weekly upgrade of the portfolio that the expected return of the stock yield and its variance do not undergo any dynamic change, as was assumed in the classical procedure. We shall assume the contrary, and we will describe the dynamics with ARIMA model family [1]. Each model used depicts the time development of the five-day yields of the given stock. The model was successively updated with the arrival of new data, however we updated it only every other week because one week of new data was a too short period to lead to a perceptible modification of the model. We found each model on the basis of Akaike information criterion [2], using autocorrelation and partial autocorrelation functions as the first hint for the selection of the model. There are usually several candidates for a good model. In most cases, an autoregressive model proved to be the best by the information criterion. In several cases, especially in the case of Erste Bank stock, moving average model proved to be better by the criterion, however the model was noninvertible with one or more of the roots of the moving average lag polynomial equal to one. Since such model would be difficult to use for forecasts, we selected also in this case an autoregressive model which was one of the best by the information criterion, although not the very best. The autoregressive models were applied to the differenced series of the five-day stock yields. Resulting models containing only statistically significant coefficients are as follows:

First and Second Week:
CEME: $\Delta z_t = -0.609\Delta z_{t-4} - 0.408\Delta z_{t-8} + 0.317\Delta z_{t-9} + \varepsilon_t$
ČEZ: $\Delta z_t = -0.251\Delta z_{t-1} - 0.471\Delta z_{t-4} - 0.354\Delta z_{t-5} + \varepsilon_t$
ERSTE: $\Delta z_t = -0.445\Delta z_{t-4} + \varepsilon_t$
NWR: $\Delta z_t = -0.46\Delta z_{t-4} - 0.471\Delta z_{t-8} + 0.41\Delta z_{t-12} + \varepsilon_t$
TEL : $\Delta z_t = -0.245\Delta z_{t-1} - 0.394\Delta z_{t-4} + \varepsilon_t$

Third and Fourth Week:
CEME: $\Delta z_t = -0.627\Delta z_{t-4} - 0.338\Delta z_{t-8} + \varepsilon_t$
ČEZ: $\Delta z_t = -0.23\Delta z_{t-1} - 0.75\Delta z_{t-4} + 0.34\Delta z_{t-5} + 0.24\Delta z_{t-7} - 0.35\Delta z_{t-8} + \varepsilon_t$
ERSTE: $\Delta z_t = -0.516\Delta z_{t-4} - 0.286\Delta z_{t-8} + \varepsilon_t$
NWR: $\Delta z_t = -0.462\Delta z_{t-4} - 0.467\Delta z_{t-8} - 0.4\Delta z_{t-12} + \varepsilon_t$
TEL: $\Delta z_t = -0.29\Delta z_{t-1} - 0.61\Delta z_{t-4} - 0.47\Delta z_{t-8} - 0.3\Delta z_{t-11} - 0.54\Delta z_{t-12} + \varepsilon_t$

Fifth and Sixth Week:
CEME: $\Delta z_t = -0.623\Delta z_{t-4} - 0.326\Delta z_{t-8} + 0.214\Delta z_{t-9} + \varepsilon_t$
ČEZ: $\Delta z_t = -0.693\Delta z_{t-4} + 0.264\Delta z_{t-7} - 0.277\Delta z_{t-8} + \varepsilon_t$
ERSTE: $\Delta z_t = -0.522\Delta z_{t-4} - 0.308\Delta z_{t-8} + \varepsilon_t$
NWR: $\Delta z_t = -0.49\Delta z_{t-4} - 0.506\Delta z_{t-8} - 0.36\Delta z_{t-12} + \varepsilon_t$
TEL: $\Delta z_t = -0.29\Delta z_{t-1} - 0.61\Delta z_{t-4} - 0.46\Delta z_{t-8} - 0.31\Delta z_{t-11} - 0.55\Delta z_{t-12} + \varepsilon_t$

Seventh Week:
CEME: $\Delta z_t = -0.638\Delta z_{t-4} - 0.319\Delta z_{t-8} + 0.21\Delta z_{t-9} + \varepsilon_t$
ČEZ: $\Delta z_t = -0.74\Delta z_{t-4} + 0.24\Delta z_{t-7} - 0.319\Delta z_{t-8} + \varepsilon_t$
ERSTE: $\Delta z_t = -0.527\Delta z_{t-4} - 0.346\Delta z_{t-8} + \varepsilon_t$
NWR: $\Delta z_t = -0.4976\Delta z_{t-4} - 0.353\Delta z_{t-8} + \varepsilon_t$
TEL: $\Delta z_t = -0.25\Delta z_{t-1} - 0.63\Delta z_{t-4} - 0.5\Delta z_{t-8} - 0.29\Delta z_{t-11} - 0.54\Delta z_{t-12} + \varepsilon_t$

These models were used first to obtain predictions of the conditional expected values of the five-days-ahead yields of each stock. Assuming independence of ε_t's across time, the expected value of z_t conditional on all the past available information z_{t-1}, z_{t-2},..., z_1 equals in case of a zero-mean stationary process $\Delta z_t = \varphi_1\Delta z_{t-1} + \varphi_2\Delta z_{t-2} + \dots + \varphi_p\Delta z_{t-p} + \varepsilon_t$ to

$$\begin{aligned} E(z_t|z_{t-1}, z_{t-2}, \dots, z_1) &= z_{t-1} + \varphi_1\Delta z_{t-1} + \varphi_2\Delta z_{t-2} + \dots + \varphi_p\Delta z_{t-p} + E(\varepsilon_t|z_{t-1}, z_{t-2}, \dots, z_1) \\ &= z_{t-1} + \varphi_1\Delta z_{t-1} + \varphi_2\Delta z_{t-2} + \dots + \varphi_p\Delta z_{t-p} + E(\varepsilon_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) \\ &= z_{t-1} + \varphi_1\Delta z_{t-1} + \varphi_2\Delta z_{t-2} + \dots + \varphi_p\Delta z_{t-p} + E(\varepsilon_t) \end{aligned}$$

or

$$E(z_t|z_{t-1}, z_{t-2}, \dots, z_1) = z_{t-1} + \varphi_1\Delta z_{t-1} + \varphi_2\Delta z_{t-2} + \dots + \varphi_p\Delta z_{t-p}. \tag{1}$$

This is due to linearity of the expected value operator. Equation (1) and the law of iterated predictions [3] were used for the forecast of the conditional expected values of the five-days-ahead yields. Further, the models found were used to obtain estimates of the residuals $\hat{\varepsilon}_t$. The simple average of the second powers of these residuals was then employed to obtain the estimate of the variance $var(\varepsilon_t)$. Finally, using this estimate and the estimates of φ_i's, five-day-ahead conditional variances of the stock yields were calculated, using the usual operations applied to time series models. This means that if the model is $y_t = \varphi y_{t-4} + \varepsilon_t$, for instance, and one wants to calculate $var(y_{t+2}|y_t, y_{t-1}, \dots, y_1)$, then since $y_{t+2} = \varphi^2 y_t + \varphi\varepsilon_{t+1} + \varepsilon_{t+2}$ by recursive substitution, the conditional variance equals $var(\varepsilon_t)\cdot(1+\varphi^2) \cong \widehat{var}(\varepsilon_t)\cdot(1+\hat{\varphi}^2)$, as ε_t's are independent. The same logic was used in calculating $var(y_{t+5}|y_t, y_{t-1}, \dots, y_1)$. Since portfolio analysis also involves estimates of covariances, one might think of their dynamic description as well. A multivariate model might then be more appropriate for our problem, however the correct selection of such model is more involved and less certain. Therefore we adopted the standard estimates of the covariances here. Our alternative procedure is thus a hybrid one with some moments being estimated in a nonstandard way, while others being estimated in the standard way. The forecasts of conditional expected values and variances are in table 6 and 7.

Week	CEME	CEZ	ERSTE	NWR	TELEF.
1.	-0.022	0.001	-0.007	-0.017	-0.015
2.	0.051	-0.010	0.012	-0.024	-0.015
3.	-0.004	0.010	-0.009	-0.033	-0.009
4.	-0.007	-0.005	-0.026	-0.042	-0.012
5.	-0.016	-0.010	-0.067	-0.037	-0.019
6.	0.017	-0.015	-0.022	0.016	0.005
7.	-0.012	-0.016	0.013	-0.009	-0.012

Table 6 Forecasts of five-day-ahead conditional expected yields of the selected stock for different weeks

Week	CEME	CEZ	ERSTE	NWR	TELEF.
1.	0.002510773	0.000337216	0.002696325	0.001061261	0.000562437
2.	0.002510773	0.000337216	0.002696325	0.001061261	0.000562437
3.	0.002667569	0.000305385	0.002096976	0.000915414	0.000360972
4.	0.002667569	0.000305385	0.002096976	0.000915414	0.000360972
5.	0.002206435	0.000260801	0.001856409	0.000926423	0.000321658
6.	0.002206435	0.000260801	0.001856409	0.000926423	0.000321658
7.	0.001955696	0.000288055	0.001797117	0.000946821	0.000280025

Table 7 Forecasts of five-day conditional variances of the yields for different weeks

Using the forecasts of conditional expected values and variances of the stock yields for a week ahead, the portfolio for each week was again optimized and the expected and actual return provided by the portfolio were compared. The results in decimal form are contained in table 8 which is an analogy to table 5.

Expected Return	Realized Return	Abs. Difference	w(1)	w(2)	w(3)	w(4)	w(5)
0.001	-0.0076	0.0086	0.000	0.975	0.000	0.000	0.025
0.010	-0.0105	0.0205	0.380	0.000	0.000	0.000	0.620
0.010	-0.0130	0.0225	0.000	1.000	0.000	0.000	0.000
-0.005	-0.0232	0.0187	0.000	1.000	0.000	0.000	0.000
-0.010	-0.0096	0.0003	0.000	1.000	0.000	0.000	0.000
0.007	0.0088	0.0019	0.176	0.000	0.000	0.000	0.824
0.005	0.0349	0.0299	0.000	0.000	0.687	0.000	0.313

Table 8 Portfolio optimization based on alternative moment estimates, and the true market progress

Let us note here that in reality a rational investor would not use the alternative approach in the fourth and fifth week because of the negative expected returns forecast. Since all of these returns are negative, as can be seen in table 6, there is obviously no way how to optimize the portfolio and not experience a loss. We included the two weeks in our analysis for the sole purpose of finding out how the estimates behave mathematically, as compared to the true returns in the fourth and fifth week.

Finally, Figure 1 compares the absolute differences between expectations and reality, recorded both in the case of the classical estimates and the alternative estimates. It can be seen that the alternative estimates perform strikingly better than the classical ones in the sense that in their case the absolute differences are smaller, enabling the investor to have a greater control over their portfolios. The better performance of the alternative estimates suggest that they are closer to the real expected value and variance of the portfolio held by the investor.

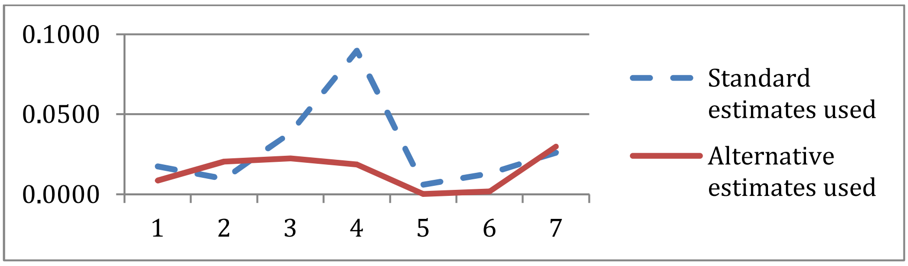

Figure 1 Absolute differences between expected returns and true returns for the period of seven weeks

Stock investments are accompanied by transaction costs which consist of the amount paid to the broker who carries out the trade, and the spread, i.e. the difference between the price at which the stock is bought, and the price at which the stock is sold. The spread is accounted for in the portfolio yields described in table 5 and table 8, so let us focus on the commission paid to the broker. The way the commission is calculated may vary among banks, but a common practice is to take it as a fixed percentage of the total amount invested in stock, regardless of whether the stock is bought or sold. Let S denote the initial amount of money invested in stock, and let $100p$ express the percentage put away for the commission, $1 > p > 0$. If the investor updates the portfolio once a week, the portfolio bought in the first week is valued at $S/(1+p)$. Depending on the development of the market, the value then changes to $[S/(1+p)]\cdot(1+i_1)$ where i_1 denotes the weekly yield of the portfolio. Updating the portfolio for the following week now means selling the current portfolio, i.e. obtaining the amount

$$[S/(1+p)]\cdot(1+i_1) - p[S/(1+p)]\cdot(1+i_1) = (1-p)\cdot[S/(1+p)]\cdot(1+i_1), \tag{2}$$

where the substraction occurs due to the transaction costs, and then reinvesting this amount to obtain a new portfolio of the desired composition, valued due to the transaction costs again at

$$\left(\frac{1-p}{1+p}\right)\cdot\left(\frac{S}{1+p}\right)\cdot(1+i_1). \tag{3}$$

Continuing this way, it is easy to see that after n weeks the investor registers in cash the amount

$$S\cdot\left(\frac{1-p}{1+p}\right)^n\cdot\prod_{k=1}^{n}(1+i_k), \tag{4}$$

where i_k denotes the portfolio weekly yield in the kth week of the investment. It is obvious from (4) that transaction costs play a role when one is concerned about the performance of the portfolio, but has no effect on the conclusion which of the *two portfolio strategies* we described in this article performs better, as the latter statement depends solely on the term $\prod_{k=1}^{n}(1+i_k)$. Using the real returns from table 5 and table 8, the term $\prod_{k=1}^{n}(1+i_k)$ equals 0.936 and 0.979, respectively, confirming that the alternative approach gives a better result.

4 Conclusion

We presented an analysis of two different approaches to optimizing a stock portfolio that would be always held for a period of five days. The optimization was performed as suggested by Markowitz. One of the approaches was the classical one, calculating the commonly used sample averages and covariances of five-day stock yields from the historical data and using them for the optimization. The alternative approach modelled the historical data with ARIMA equations, and calculated five-day forecasts of conditional averages and variances of the analyzed stock yields. The remaining second moments necessary for the optimization were estimated as in the classical case, i.e. using historical stock yields without using more involved ARIMA equations. The two approaches were compared as to which of them provides an expected return that is closer to the trully realized return. It turns out the alternative approach, benefiting from the description of the actual dynamics of the market, leads to an optimized portfolio with a more realistic expected return estimate. Investor holding such a portfolio has a greater control over the investment, and can rely to a greater extent on experiencing a true profit that is closer to the expected profit.

References

[1] Box, G. E. P., Jenkins, G. M., and Reinsel G.C.: *Time Series Analysis*. John Wiley & Sons, 2008.
[2] Brockwell P.J., and Davis, R.A.: *Time Series: Theory and Methods*. Springer-Verlag New York, 1991.
[3] Hamilton, J.D.: *Time Series Analysis*, Princeton University Press, 1994.