Accepted Manuscript

Forecasting Intraday Volume: Comparison of two early models

SZŰCSBalázs Árpád

Please cite this article as: SZŰCSBalázs Árpád, Forecasting Intraday Volume: Comparison of two early models, *Finance Research Letters* (2016), doi: 10.1016/j.frl.2016.11.018

Highlights

- **Aim:** Compare intraday volume forecasting models of the literature

- **Models:** Bialkowski, Darolles, Le Fol (2008) and Brownlees, Cipollini, Gallo (2011)

- **Intraday data:** 11 years of 33 NYSE and NASDAQ shares

- **Findings:** The former model is more accurate and much faster to estimate

Forecasting Intraday Volume:

Comparison of two early models

SZŰCS, Balázs Árpád*†‡

Abstract

There are few intraday volume forecasting models in the literature, and they do not reflect on each other regarding forecast performance. This paper compares two models that are often referenced: the model of *Bialkowski, Darolles and Le Fol (2008)* to that of *Brownlees, Cipollini and Gallo (2011)* using intraday data that covers 11 years of 33 NYSE and NASDAQ shares. The former is found to produce more accurate forecasts, while its estimation is faster by several orders of magnitude.

Keywords: intraday, volume, forecasting, comparison

1   Introduction

Research covering stock exchanges usually focuses on the price; therefore much less attention is paid to the trading volume. Consequently, our knowledge on volume is much narrower. However, the volume itself plays a significant role in the trading process. A future trade of large volume is likely to have a price effect, thus causing potentially substantial loss to its submitter. Order splitting and, in general, order execution strategies are hence of high importance in everyday trading, and they cannot be implemented without some forecast of the volume for that day. This is especially the

---

case in recent days as algorithmic trading is becoming increasingly dominant on the market, since most algorithms require volume as input[1]. Furthermore, VWAP strategies (see [Madhavan, 2002], [Kissell et al., 2004]) that amount for a significant portion of overall institutional trading can be executed solely based on intraday volume forecasts.

Rather few publications can be found on the subject of forecasting volume in general, and even fewer on forecasting intraday volume. This paper compares two relatively early models that both decompose the intraday U-shape of volume, but in a very different approach. [Bialkowski et al., 2008] suggests an intraday volume forecasting model for stocks using an *additive* decomposition of the intraday U-shape. This model uses both cross-sectional and time series data for each share. [Brownlees et al., 2011] builds an intraday volume forecasting model for ETFs, which is based on a *multiplicative* decomposition of the intraday U-shape. This is a purely time series model, but it uses both daily and intra-daily frequencies at the same time. Both models outperform a simple benchmark commonly used in practice, but the latter does not reflect on the former regarding forecast performance. It is, therefore, impossible to tell how they perform compared to each other, since they are both evaluated on different samples, using different error measures.

The contribution of this paper is twofold. First, it reflects on the estimation process of these models. Second, it compares their forecasts (to each other as well as to their common benchmark) on the same sample, using identical error measures in an attempt to decide which model could be considered preferable when it comes to intraday volume forecasting of shares.

Identifying the better model is of high relevance to practitioners because any increase in forecasting accuracy can be directly converted to monetary gains through the avoided price effect. Additionally, it is also important to know which model is quicker to be estimated, because in the world of high-frequency trading being just a fraction faster than others can be crucial. From a theoretical point of view, knowing which specification is more effective may contribute to creating an even better model of volume later on.

---

[1]Technical traders also tend to monitor the volume. ([Frömmel and Lampaert, 2016])

Observations per day	26
Number of days	2 668
Observations per share	69 368
Number of shares	33
Observations in the sample	2 289 144

Table 1: Overview of the final sample
Source: Own editing

2 Data

The database[2] used in the analysis contains stocks included in the Dow Jones Industrial Average (DJIA or Dow 30) index that covers significant companies listed on exchanges in the United States. The index has been computed since 1896. The actual shares included in it somewhat varied since the introduction of the index, which is why the database contains not 30, but 36 tickers. Most of them, namely 33 are listed on the NYSE, and the remaining 3 are listed on NASDAQ. The date of the first data point is 01/02/1998, except for stocks that were introduced to the exchange later, in which case the date of the IPO is the first data point. The date of the last data point is 07/13/2012 uniformly for all tickers.

The sample remaining after the data cleaning process ranges from 10/10/2001 to 07/13/2012, a period that is 130 months, nearly 11 years long. The number of tickers remaining in the sample is 33 because the three tickers with the shortest registered period were excluded (for details, see Table A.1 in the Appendix). The original frequency of observations was 1 minute, but I aggregated the data into 15-minute bins, to comply with the literature. This resulted in 26 observations every day for each ticker (exchanges open at 9:30 a.m. and close at 4:00 p.m.). The stocks remaining in the sample were liquid enough, meaning that every stock had trades and thus a volume record larger than zero in every 15-minute interval. The database finally used for analysis thus contains 2.29 million observations (see Table 1).

The volume data was converted to turnover according to the following: $x_t = V_t/TSO_t$, where x stands for turnover, V for volume, and TSO for the total shares outstanding. The TSO data was downloaded from a Bloomberg terminal.
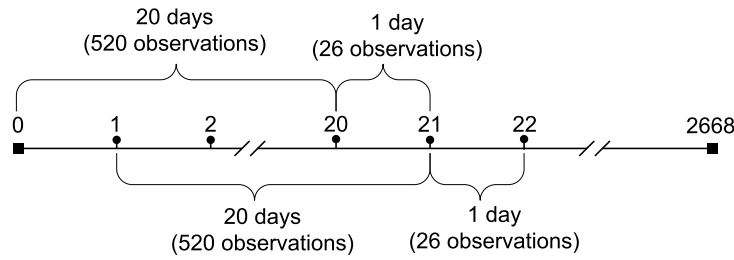
---

[2]Obtained from kibot.com.

4

Figure 1: Data handling
Source: Own editing

3   Data handling and error measures

For the sake of better comparability, data should be handled similarly throughout the different estimations. Although this naturally means the usage of the same (previously presented) sample, some further conventions have to be made.

The estimation period is chosen to be 20 days (following [Bialkowski et al., 2008]), which corresponds to one calendar month (approximately 20 trading days). This results in 520 observations in the estimation period. Forecasts are then to be made for the following 1 day (26 observations, following [Bialkowski et al., 2008] again). This period is used to evaluate the forecasts.

The parameters of the models are thus re-estimated daily, using a 20-day moving window, and the forecast is always produced for the following day, as illustrated in Figure 1. Consequently, 2648 different parameter estimations and forecasts for 2648 days are produced for each of the 33 shares in the sample.

While parameters are updated daily, the information base for the forecast is updated every 15 minutes. This is because 26 data points are to be forecasted each day. Although the parameters of the models are unchanged during the day, it makes sense to take advantage of the actuals that unfold during the day. This approach is often called one-step-ahead forecasting.

Another important issue to cover is that of the error measures. Among the various possibilities, two of the most common error measures are selected here, those that simply measure the deviation of the actuals and the forecasts. The first one is the Mean Squared Error (MSE):

5

$$MSE = \frac{\sum\limits_{t=1}^{N} \left(Y_t - Y_t^f \right)^2}{N} \tag{1}$$

where Y^f denotes the forecasted value of Y.

The second measure is the Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{\sum\limits_{t=1}^{N} \left| \frac{Y_t - Y_t^f}{Y_t} \right|}{N} \tag{2}$$

Both measures are calculated for each share, and also in the average of all shares. A model may be considered better than the other according to an error measure, if it gives lower average value, and also gives lower values on a higher number of individual shares.

4 Estimation of the models

The benchmark of both [Bialkowski et al., 2008] and [Brownlees et al., 2011] are a simple average defined as:

$$\hat{y}_{p+1} = \frac{1}{L} \sum_{l=1}^{L} y_{p+1-m \cdot l} \tag{3}$$

where L stands for the number of days involved, while m denotes the number of intraday bins (which is 26 in our case). This formula is said to be commonly used in practice when it comes to intraday volume forecasts. It is hereinafter referred to as the *U-method* and is also estimated and compared to the other models. It clearly incorporates the well-known stylized fact of the intra-daily U-shape of volume, a feature that appears in the two other models as well.

4.1 The model of [Bialkowski et al., 2008]

The model of [Bialkowski et al., 2008] (in brief: *BDF model*, from the capital letters of the authors' names) uses an additive decomposition for the U-shape:

$$X = K + e \tag{4}$$

where X is a (PxN) matrix of turnovers, with P observations and N shares. The K common component is obtained using the factor analysis for large dimensions described in [Bai, 2003]. The forecast of the common component is obtained according to (3).

The forecast of the e specific component can be determined in two alternative ways. First, using an AR(1) model:

$$e_p = c + \theta_1 e_{p-1} + \varepsilon_p \tag{5}$$

where ε is white noise. Second, using a SETAR model:

$$e_p = (c_{1,1} + \theta_{1,2} e_{p-1,}) I(e_{p-1}) + (c_{2,1} + \theta_{2,2} e_{p-1})(1 - I(e_{p-1})) + \varepsilon_p \tag{6}$$

where

$$I(z) = \begin{cases} 1 & \text{if } z \leq \tau \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The two versions are denoted as BCG_AR and BCG_SETAR respectively, depending on the model of the specific component. The BDF model uses the information of the entire market (all shares) to forecast a different U-shape (common component) for each stock and then forecasts the specific component for each stock individually. The sum of these two is considered to be the forecast of the turnover.

Based on the above, the forecasts of the BDF model are obtained with no difficulties.

4.2 The model of [Brownlees et al., 2011]

The model of [Brownlees et al., 2011] (in brief: *BCG model*, from the capital letters of the authors' names) uses a multiplicative decomposition of the U shape:

7

$$x_{t\,i} = \eta_t \, \phi_i \, \mu_{t\,i} \, \varepsilon_{t\,i} \tag{8}$$

where $t \in \{1, \ldots, T\}$ denotes the number of days, $i \in \{1, \ldots, I\}$ the number of intraday bins (26 in our case), x the turnover, η the daily component, ϕ the intraday periodic component (the U-shape), and μ the intraday non-periodic component. The innovation term ε is i.i.d., nonnegative, with mean 1 and constant variance σ^2.

4.2.1 Specification

The specifications of each of these terms are the following. The daily component:

$$\eta_t = \alpha_0^{(\eta)} + \beta_1^{(\eta)} \eta_{t-1} + \alpha_1^{(\eta)} x_{t-1}^{(\eta)} \tag{9}$$

where

$$x_t^{(\eta)} = \frac{1}{I} \sum_{i=1}^{I} \frac{x_{t\,i}}{\phi_i \mu_{t\,i}} \tag{10}$$

The intraday periodic component:

$$\phi_{j+1} = exp\left\{ \sum_{k=1}^{K} [\delta_{1\,k} \cos(f\,k\,j) + \delta_{2\,k} \sin(f\,k\,j)] \right\} \tag{11}$$

where $f = \frac{2\pi}{I}$, $K = int(\frac{I}{2})$, $j = \{0, \ldots, I-1\}$ and finally $\delta_{2\,K} = 0$ if I is even.

The intraday non-periodic component:

$$\mu_{t\,i} = \alpha_0^{(\mu)} + \beta_1^{(\mu)} \mu_{t\,i-1} + \alpha_1^{(\mu)} x_{t\,i-1}^{(\mu)} \tag{12}$$

where

$$x_{t\,i}^{(\mu)} = \frac{x_{t\,i}}{\eta_t \, \phi_i} \tag{13}$$

The parameter constraint $\alpha_0^{(\mu)} = 1 - \beta_1^{(\mu)} - \alpha_1^{(\mu)}$ should be applied in the course of the estimation. Also, some initial values are provided:

8

- $\eta_0 = x_0^{(\eta)} = \frac{1}{5I} \sum_{t=1}^{5} \sum_{i=1}^{I} x_{ti}$

- $\mu_{10} = x_{10}^{(\mu)} = 1$

- $\mu_{t0} = \mu_{t-1\,I}$

- $x_{t0}^{(\mu)} = x_{t-1\,I}^{(\mu)}$

The authors estimate 4 different variants of the model, all of which are further extensions to the above. They find the one with all the extensions to perform the best, so I only estimate this full version. This has the following extensions:

- Dummy at the first observation of each day in (12)

- Asymmetric effects (based on returns) in (9) and (12)

- A second lag in (12)

4.2.2 Estimation

The parameters are estimated via the Generalized Method of Moments (GMM) in one step. The authors provide the exact numerical optimization problem to be solved. However, the reader is still left with some uncertainties regarding the estimation of the model.

First, the initial values of some variables are not explicitly specified. These are, on the one hand, the return related extensions in (9) and (12). I assumed these to be zero, meaning that they have no initial effect. On the other hand, the second lag in (12) also lacks an initial value. I assumed it to be identical to that of the first lag.

Second, it is mentioned that the number of terms in (11) *may be considerably reduced* from 25, but it is not specified how exactly, nor the final choice. I decided to use 4, which reduces the total number of parameters to be estimated from 34 to 13. This makes the optimization much simpler, but still, 4 parameters should be enough to approximate the 26 points of the U-shape.

Third, although this is merely a technical issue, it remains unknown how the starting value of θ (the vector of the parameters to be estimated numerically) is specified during the optimization,

which turns out to be a key issue in the success of the estimation. After lengthy tests for an acceptable starting θ, I also inserted a grid search of θ before the optimization for each day.

BCG_0 The first estimation was run exactly as described above. This setting is denoted as BCG_0. The objective function provided for the GMM estimation did not appear to be smooth enough (on my sample) to allow for finding an acceptable solution within acceptable time. It took 60 days to run the estimation[3], which is much longer than the time needed for any other model I dealt with (the BDF model took about two hours to run). An acceptable solution would be a forecast the magnitude of which is comparable with the actuals observed later, which was not the case here, not even after the lengthy estimation described above. I tried different solver algorithms, increased numbers of iterations and function evaluations, but none of those helped, so I started experimenting with different settings.

BCG_1 The authors mention that parameters are rather stable across time and assets as well. Building on this, I altered the estimation as described below (BCG_1 setting). I looked up the first day for each share where the parameters found using BCG_0 yielded acceptable results, and used those parameters throughout the entire time scale of the sample, instead of re-estimating them daily. The forecasts obtained this way were acceptable. Even though the time needed for estimation is drastically reduced this way, keeping parameters unchanged for 11 years is questionable when it comes to intraday forecasting.

BCG_2 Returning to daily parameter estimation, the following is denoted as BCG_2. Due to the moving window depicted in Figure 1, the data of two consecutive estimation periods are 95% the same. We could make use of this observation by using the previously found θ as the initial value for the next optimization, instead of inserting a grid search for every day. This also saves some running time. Unfortunately, the results are even worse than with BCG_0. A possible explanation to this may be that once a bad θ hits in, it is inherited to all later days, thus reducing the chance of finding an acceptable one.

---

[3]This is expressed in machine time, which shows the theoretical waiting time using a single computer with the average performance of the 6 computers I had at my disposal. The actual waiting time was shorter due to the use of several computers.

BCG_3 Finally, I use a blend of the above denoted as BCG_3. I use the grid first and check immediately whether the result is acceptable or not. If it is, I move on to the next day. Otherwise, the accepted parameters of the previous day are used. Thus we return to the elongated running time of BCG_0. This setting yields the best results among all four.

5 Results

This section compares the two versions of the BDF model, the BCG_3 model and the U-method. Since the BCG_3 unequivocally outperforms the other 3 estimation settings, I only keep this version for further analysis.

Table 2.a shows[4] that the BDF_AR model produced the lowest MSE values in the average of all shares. However, MSE is known to be scale-sensitive, which might be a problem when averaging over different shares. To mitigate this effect, I also calculated a modified average:

$$MSE^* = \frac{1}{N} \sum_{i=1}^{N} \frac{MSE_i}{\left(\frac{a_i}{a_{min}}\right)^2} \tag{14}$$

where MSE_i denotes the average MSE value of share i, while a_i is the average turnover of share i, and a_{min} equals the smallest a_i. Table 2.a shows that this correction does not change the fact that BDF_AR performs the best among the models.

It is also worth to compare the performances on a share-by-share basis. Table 2.b depicts a pairwise comparison of the models showing the number of shares where the model in the first column produces lower/higher average MSE compared to the model in the first row. For example, the BCG_3 model performed better than the U-method in 28 cases, and worse in 5 cases out of the 33 shares. We can conclude that the BDF_AR model is superior in this comparison to all of the other ones.

After conducting a similar analysis based on the MAPE measure, the BDF_SETAR model proves to be the best performing one: it has the lowest overall error (Table 2.a), and beats all the other models in pairwise comparison for the majority of the shares (Table 2.c).

---

[4]Notation: $aE + b \equiv a \cdot 10^b$, and $aE - b \equiv a \cdot 10^{-b}$

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	1,02E-03	6,49E-04	6,60E-04	6,77E-04
MSE*	3,65E-05	2,30E-05	2,38E-05	2,53E-05
MAPE	0,503	0,403	0,399	0,402

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	5 / 28
BDF_AR	33 / 0	-	27 / 6	31 / 2
BDF_SETAR	33 / 0	6 / 27	-	28 / 5
BCG_3	28 / 5	2 / 31	5 / 28	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	1 / 32	1 / 32	0 / 33
BDF AR	32 / 1	-	3 / 30	21 / 12
BDF SETAR	32 / 1	30 / 3	-	26 / 7
BCG 3	33 / 0	12 / 21	7 / 26	-

c)

Table 2. Comparison on the full sample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

12

According to the above, the appropriate version of the BDF model clearly outperforms both the BCG model and the U-method.

Since the 11-year-long sample is large enough to be divided into several shorter periods, it allows us to test for robustness of the results. Hence I divided the sample into 5 non-overlapping subperiods of equal length and repeated the same analysis for these. Detailed results are shown in the Appendix in Tables A.2-A.6.

For subperiods 2, 4 and 5, results are very similar to those of the full sample: BDF_AR is the best when measured with MSE, and BDF_SETAR is the best when measured with MAPE. In the case of subperiod 3, the only difference is that the overall MSE is lower for the BDF_SETAR model, but this is only due to the scale sensitivity, since the MSE* is still lower for the BDF_AR. These results are therefore essentially similar as well.

The only subperiod where results are different is the first one (Table A.2), where the BCG_3 model wins both overall and pairwise when evaluated with the MAPE measure.

The pairwise results of the subperiods could be aggregated, after which the number of total instances increases from 33 to $5 \cdot 33 = 165$ in each case. Table 3. shows that the BDF_AR is still the best based on the MSE, whereas the BDF_SETAR remains the best based on the MAPE measure.

Finally, let us compare the worst forecasts of the models using the 95th percentiles of the errors. Table 4. shows that the BDF_AR model produced the smallest 95th percentile for most of the shares on the full sample, as well as on each of the subsamples. This means that this model tends to make smaller extreme errors than the others.

Table 5. depicts the same for the absolute percentage errors. There is only one exception, the first subperiod, where the BCG_3 model reaches the smallest Q95 for 25 shares. In all other cases, the smallest extreme errors are most commonly attributed to the BDF_SETAR model.

6 Conclusions

In this paper, I compared the intraday volume forecasts of the BDF and the BCG models using data that covers 11 years of 33 NYSE and NASDAQ shares. Both models clearly outperform the

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	1 / 164	2 / 163	31 / 134
BDF_AR	164 / 1	-	126 / 39	140 / 25
BDF_SETAR	163 / 2	39 / 126	-	113 / 52
BCG_3	134 / 31	25 / 140	52 / 113	-

a)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	2 / 163	2 / 163	1 / 164
BDF_AR	163 / 2	-	49 / 116	92 / 73
BDF_SETAR	163 / 2	116 / 49	-	115 / 50
BCG_3	164 / 1	73 / 92	50 / 115	-

b)

Table 3. Aggregated results of the subsamples
a) Number of shares with lower/higher average MSE in a pairwise comparison. b) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

Sample	U	BDF_AR	BDF_SETAR	BCG_3
Full	0	29	3	1

a)

Sample	U	BDF_AR	BDF_SETAR	BCG_3
Sub_1	0	19	4	10
Sub_2	0	24	9	0
Sub_3	0	25	7	1
Sub_4	1	16	14	2
Sub_5	0	21	12	0
Sum	**1**	**105**	**46**	**13**

b)

Table 4. Number of shares with the lowest Q95 value calculated on Squared Errors
a) For the full sample. b) For the subsamples.
Source: Own editing

14

sample	U	BDF_AR	BDF_SETAR	BCG_3
full	0	0	24	9

a)

sample	U	BDF_AR	BDF_SETAR	BCG_3
sub_1	0	5	3	25
sub_2	0	4	18	11
sub_3	0	0	27	6
sub_4	0	2	16	15
sub_5	0	2	28	3
sum	0	13	92	60

b)

Table 5. Number of shares with the lowest Q95 value calculated on Absolute Percentage Errors
a) For the full sample. b) For the subsamples.
Source: Own editing

commonly used benchmark of the U-method. Depending on whether the MSE or MAPE criterion was used, either the BDF_AR or the BDF_SETAR version of the BDF model was the most promising.

Two remarks are to be made. First, the BCG model is originally tested for ETFs and not shares, but the volume patterns of the two asset classes can be considered similar. Second, the BCG model took longer to run by several orders of magnitude, and the success of the estimation on this data was fairly incidental itself.

According to these tests, the much shorter running time and the better forecast performance makes the BDF model recommendable over the BCG model for using on shares.

Further research could include comparison through different error measures, and to other models, such as [Humphery-Jenner, 2011], [SATISH et al., 2014], or [Manchaladore, 2010].

References

[Bai, 2003] Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71:135–171. DOI: 10.1111/1468–0262.00392.

[Bialkowski et al., 2008] Bialkowski, J., Darolles, S., and Le Fol, G. (2008). Improving VWAP strategies: A dynamic volume approach. *Journal of Banking & Finance*, 32:1709–1722. DOI:

15

10.1016/j.jbankfin.2007.09.023.

[Brownlees et al., 2011] Brownlees, C. T., Cipollini, F., and Gallo, G. M. (2011). Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics*, 9:489–518. DOI: 10.1093/jjfinec/nbq024.

[Frömmel and Lampaert, 2016] Frömmel, M. and Lampaert, K. (2016). Does frequency matter for intraday technical trading? *Finance Research Letters*, 18:177–183.

[Humphery-Jenner, 2011] Humphery-Jenner, M. (2011). Optimal VWAP trading under noisy conditions. *Journal of Banking and Finance*, 35(9):2319–2329.

[Kissell et al., 2004] Kissell, R., Glantz, M., and Malamut, R. (2004). A practical framework for estimating transaction costs and developing optimal trading strategies to achieve best execution. *Finance Research Letters*, 1:35–46.

[Madhavan, 2002] Madhavan, A. (2002). VWAP strategies. *Transaction Performance*, Spring:32–38.

[Manchaladore, 2010] Manchaladore, J. (2010). Wavelet decomposition for intra-day volume dynamics. *Quantitative Finance*, 10(8):917–930.

[SATISH et al., 2014] SATISH, V., SAXENA, A., and PALMER, M. (2014). Predicting intraday trading volume and volume percentages. *Journal of Trading*, 9(3).

Appendix

#	Ticker	Start date	Description	Exchange	Excluded
1	AA	1/2/1998	Alcoa, Inc.	NYSE	
2	AIG	1/2/1998	American International Group, Inc.	NYSE	
3	AXP	1/2/1998	American Express Company	NYSE	
4	BA	1/2/1998	Boeing Co.	NYSE	
5	BAC	1/2/1998	Bank of America Corporation	NYSE	
6	C	1/2/1998	Citigroup, Inc.	NYSE	
7	CAT	1/2/1998	Caterpillar, Inc.	NYSE	
8	CSCO	1/2/1998	Cisco Systems, Inc.	NASDAQ	
9	CVX	10/10/2001	Chevron Corporation	NYSE	
10	DD	1/2/1998	E.I. Du Pont de Nemours and Company	NYSE	
11	DIS	1/2/1998	Walt Disney Co.	NYSE	
12	GE	1/2/1998	General Electric Company	NYSE	
13	**GM**	**11/18/2010**	**General Motors Company**	**NYSE**	**Excluded**
14	HD	1/2/1998	The Home Depot, Inc.	NYSE	
15	HON	1/2/1998	Honeywell International, Inc.	NYSE	
16	HPQ	1/2/1998	Hewlett-Packard Company	NYSE	
17	IBM	1/2/1998	International Business Machines Corp.	NYSE	
18	INTC	1/2/1998	Intel Corporation	NASDAQ	
19	JNJ	1/2/1998	Johnson & Johnson	NYSE	
20	JPM	1/2/1998	JPMorgan Chase & Co.	NYSE	
21	KFT	6/14/2001	Kraft Foods, Inc.	NYSE	
22	KO	1/2/1998	The Coca-Cola Company	NYSE	
23	MCD	1/2/1998	McDonald's Corp.	NYSE	
24	MMM	1/2/1998	3M Co.	NYSE	
25	MO	1/2/1998	Altria Group Inc.	NYSE	
26	MRK	1/2/1998	Merck & Co. Inc.	NYSE	
27	MSFT	1/2/1998	Microsoft Corporation	NASDAQ	
28	PFE	1/2/1998	Pfizer Inc.	NYSE	
29	PG	1/2/1998	Procter & Gamble Co.	NYSE	
30	**PGP**	**5/26/2005**	**Pimco Global Stockplus & Incom**	**NYSE**	**Excluded**
31	T	1/2/1998	AT&T, Inc.	NYSE	
32	**TRV**	**2/27/2007**	**The Travelers Companies, Inc.**	**NYSE**	**Excluded**
33	UTX	1/2/1998	United Technologies Corp.	NYSE	
34	VZ	7/3/2000	Verizon Communications Inc.	NYSE	
35	WMT	1/2/1998	Wal-Mart Stores Inc.	NYSE	
36	XOM	12/1/1999	Exxon Mobil Corporation	NYSE	

Table A.1: Shares in the data base
Source: Kibot.com

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	4,71E-04	3,01E-04	3,20E-04	2,86E-04
MSE*	1,13E-05	8,21E-06	8,53E-06	8,35E-06
MAPE	0,479	0,385	0,386	0,379

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	0 / 33
BDF_AR	33 / 0	-	27 / 6	18 / 15
BDF_SETAR	33 / 0	6 / 27	-	12 / 21
BCG_3	33 / 0	15 / 18	21 / 12	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	0 / 33
BDF_AR	33 / 0	-	26 / 7	9 / 24
BDF_SETAR	33 / 0	7 / 26	-	6 / 27
BCG_3	33 / 0	24 / 9	27 / 6	-

c)

Table A.2: Comparison on the first subsample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	3,59E-04	2,30E-04	2,47E-04	2,34E-04
MSE*	1,11E-05	7,47E-06	7,87E-06	7,77E-06
MAPE	0,502	0,403	0,400	0,409

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	1 / 32
BDF_AR	33 / 0	-	25 / 8	26 / 7
BDF_SETAR	33 / 0	8 / 25	-	17 / 16
BCG_3	32 / 1	7 / 26	16 / 17	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	1 / 32	1 / 32	1 / 32
BDF_AR	32 / 1	-	8 / 25	20 / 13
BDF_SETAR	32 / 1	25 / 8	-	24 / 9
BCG_3	32 / 1	13 / 20	9 / 24	-

c)

Table A.3: Comparison on the second subsample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

19

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	1,14E-03	1,10E-03	1,03E-03	1,16E-03
MSE*	3,18E-05	2,42E-05	2,43E-05	2,57E-05
MAPE	0,474	0,370	0,363	0,370

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	1 / 32	1 / 32	1 / 32
BDF_AR	32 / 1	-	26 / 7	32 / 1
BDF_SETAR	32 / 1	7 / 26	-	25 / 8
BCG_3	32 / 1	1 / 32	8 / 25	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	0 / 33
BDF_AR	33 / 0	-	3 / 30	18 / 15
BDF_SETAR	33 / 0	30 / 3	-	30 / 3
BCG_3	33 / 0	15 / 18	3 / 30	-

c)

Table A.4: Comparison on the third subsample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	2,57E-03	1,26E-03	1,33E-03	1,28E-03
MSE*	4,80E-05	3,18E-05	3,23E-05	3,65E-05
MAPE	0,590	0,492	0,488	0,463

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	10 / 23
BDF_AR	33 / 0	-	22 / 11	31 / 2
BDF_SETAR	33 / 0	11 / 22	-	29 / 4
BCG_3	23 / 10	2 / 31	4 / 29	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	1 / 32	1 / 32	0 / 33
BDF_AR	32 / 1	-	8 / 25	15 / 18
BDF_SETAR	32 / 1	25 / 8	-	24 / 9
BCG_3	33 / 0	18 / 15	9 / 24	-

c)

Table A.5: Comparison on the fourth subsample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing

21

	U	BDF_AR	BDF_SETAR	BCG_3
MSE	5,35E-04	3,52E-04	3,73E-04	4,22E-04
MSE*	2,35E-05	1,67E-05	1,78E-05	2,20E-05
MAPE	0,472	0,368	0,360	0,388

a)

MSE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	1 / 32	19 / 14
BDF_AR	33 / 0	-	26 / 7	33 / 0
BDF_SETAR	32 / 1	7 / 26	-	30 / 3
BCG_3	14 / 19	0 / 33	3 / 30	-

b)

MAPE	U	BDF_AR	BDF_SETAR	BCG_3
U	-	0 / 33	0 / 33	0 / 33
BDF_AR	33 / 0	-	4 / 29	30 / 3
BDF_SETAR	33 / 0	29 / 4	-	31 / 2
BCG_3	33 / 0	3 / 30	2 / 31	-

c)

Table A.6: Comparison on the fifth subsample
a) Average error measures for each model. b) Number of shares with lower/higher average MSE in a pairwise comparison. c) Number of shares with lower/higher average MAPE in a pairwise comparison.
Source: Own editing