

DYNAMIC VS STATIC AUTOREGRESSIVE MODELS FOR
FORECASTING TIME SERIES

Chris Xie

*Polytechnic Institute
New York University (NYU), NY*

chris.xie@toprenergy.com

Phone: 905-931-0577

June 1, 2008

ABSTRACT

I build an innovative Dynamic Autoregressive Model (DAR) in forecasting time series, and make comparison with a Static Autoregressive Model (SAR). DAR model requires re-evaluating optimal orders and coefficients at each period, while SAR models simply treats them as constants. The optimal length of the base (historical data for building autoregressive models) has been also investigated. Results show that on average DAR models outperform SAR models by about 1% up to double digits percent. With increase of the length of the base, adjusted R-Squares for both models are diminishingly increasing, and errors (in percentage) differences are vanishing.

Key Words:

Dynamic Autoregressive Model (DAR); Static Autoregressive Model (SAR); Base Data (BD) also called the base; Optimal Number of Orders (p)

JEF Classifiers: C32, C53

I. INTRODUCTION

The autoregressive model is one of powerful tools to forecast time series. Autoregressive models differ from standard linear regression models, because they do not regress on independent variables, but on a subset of the dependent variables (i.e., its lagged values). The p^{th} order autoregressive model can be mathematically expressed as follows:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t$$

Here,

Y_t is the dependent variable at the time period t ;

Y_{t-i} is the dependent variable at the time period $t-i$ ($i = 1, 2, \dots, p$)

β_i is the coefficients ($i = 0, 1, 2, \dots, p$);

ε_i is the residual or random error;

p is the number of orders or autoregression rank;

Supposed that we know $Y_1, Y_2, \dots, Y_p, \dots, Y_{p+m}$, then we have following linear equations:

$$\begin{aligned} Y_{p+1} &= \beta_0 + \beta_1 Y_p + \beta_2 Y_{p-1} + \beta_3 Y_{p-2} + \dots + \beta_p Y_1 + \varepsilon_{p+1} \\ Y_{p+2} &= \beta_0 + \beta_1 Y_{p+1} + \beta_2 Y_p + \beta_3 Y_{p-1} + \dots + \beta_p Y_2 + \varepsilon_{p+2} \\ Y_{p+3} &= \beta_0 + \beta_1 Y_{p+2} + \beta_2 Y_{p+1} + \beta_3 Y_p + \dots + \beta_p Y_3 + \varepsilon_{p+3} \quad \dots\dots\dots (1) \\ &: \\ Y_{p+m} &= \beta_0 + \beta_1 Y_{p+m-1} + \beta_2 Y_{p+m-2} + \beta_3 Y_{p+m-3} + \dots + \beta_p Y_m + \varepsilon_{p+m} \end{aligned}$$

$$\text{Let } Y = \begin{pmatrix} Y_{p+1} \\ Y_{p+2} \\ Y_{p+3} \\ \dots \\ Y_{p+m} \end{pmatrix}_{m \times 1}, \quad X = \begin{pmatrix} 1 & Y_p & Y_{p-1} & Y_{p-2} & \dots & Y_1 \\ 1 & Y_{p+1} & Y_p & Y_{p-1} & \dots & Y_2 \\ 1 & Y_{p+2} & Y_{p+1} & Y_p & \dots & Y_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & Y_{p+m-1} & Y_{p+m-2} & Y_{p+m-3} & \dots & Y_m \end{pmatrix}_{m \times (p+1)},$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{pmatrix}_{m \times 1}, \quad \varepsilon = \begin{pmatrix} \varepsilon_{p+1} \\ \varepsilon_{p+2} \\ \varepsilon_{p+3} \\ \dots \\ \varepsilon_{p+m} \end{pmatrix}_{m \times 1}$$

Then we have $Y = X\beta + \varepsilon$. To minimize sum of squared errors ($\sum \varepsilon_i^2$), we can obtained

the optimal solutions for β , which could be expressed as follows:

$$\hat{\beta} = (X^T X)^{-1}(X^T Y). \quad \dots\dots\dots (2)$$

So given X, Y, we could easily calculate coefficients (betas) for this p-order autoregressive model.

Now we define the base data (BD) as follows:

$$BD = \begin{pmatrix} Y_{p+1} & Y_p & Y_{p-1} & Y_{p-2} & \dots & Y_1 \\ Y_{p+2} & Y_{p+1} & Y_p & Y_{p-1} & \dots & Y_2 \\ Y_{p+3} & Y_{p+2} & Y_{p+1} & Y_p & \dots & Y_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ Y_{p+m} & Y_{p+m-1} & Y_{p+m-2} & Y_{p+m-3} & \dots & Y_m \end{pmatrix}_{m \times (p+1)}$$

on which the coefficients (betas) will be calculated, and m, the number of rows in BD is called as the length of the base data (BD).

Therefore, at the time period T, as we know Y_i ($i = 1, 2, \dots, T$), we could estimate Y_{T+1} , as $\hat{Y}_{T+1} = \beta_0 + \beta_1 Y_T + \beta_2 Y_{T-1} + \dots + \beta_p Y_{T-p}$. Based on this linear regressive model, other key figures such as adjusted R-square can also be calculated.

II DYNAMIC AUTOREGRESSIVE MODEL (DAR)

Dynamic Autoregressive Model (DAR) is an autoregressive model with dynamically re-evaluating all betas (coefficients) and the number of lags (p-order) with respect to a rolling base for forecasting asset prices in order to achieve the best result by autoregressive models. Comparing to a standard autoregressive model (SAR), which treats all coefficients and the number of lags (p-order) initially generated by the initial base data as constants, DAR model not only treats all coefficients and the number of lags (p) as variables (time-subscribed), but also has to choose the optimal length (m) of base data (BD) to minimize the error terms such as the sum of errors-squared.

Given all historical figures (daily closed prices) before December 31, 2007 known for all stocks of DJI, in forecasting their daily closed prices during January 1 to May 30, 2008, our investigations show that DAR models outperform SAR models significantly on average at lower number (less than 350) of the length of the base, while with increasing the length of the base, both models have almost similar result (see details in the section of Dynamic Autoregressive Model vs Static Autoregressive Model).

The detailed processes or steps are elaborated as follows.

Given the asset prices (for example, the daily closed price) for a certain period of time as Y_1, Y_2, \dots, Y_T known and assuming the optimal length (m) of the base data (BD) already known, we try to forecast the asset price $Y_{T+1}, Y_{T+2}, \dots, Y_{T+i}$ (at the period T+1, T+2, ...T+i).

Forecast the asset price Y_{T+1} at the time period $T+1$

Step 1: Calculate initial p_1 (order) value by using the given known figures (Y_1, Y_2, \dots, Y_T)
 Keep adding additional lags until the adjusted R^2 stops increasing, or increase the number of lags (p) until Akaike Information Criterion (AIC) reaches the minimum

value. Here AIC can be expressed as
$$AIC = \log \left(\frac{\sum \varepsilon_i^2}{n} \right) + \frac{2p}{n}$$

Step 2: Determine the initial base data (BD) based on given known figures (Y_1, Y_2, \dots, Y_T), since the p_1 (order) value and m (the length of base) are known

The initial base data (BD) is set as follows:

$$B_1 = \begin{pmatrix} Y_{T-m+1} & Y_{T-m} & Y_{T-m-1} & Y_{T-m-2} & \dots & Y_{T-m-p_1+1} \\ Y_{T-m+2} & Y_{T-m+1} & Y_{T-m} & Y_{T-m-1} & \dots & Y_{T-m-p_1+2} \\ Y_{T-m+3} & Y_{T-m+2} & Y_{T-m+1} & Y_{T-m} & \dots & Y_{T-m-p_1+3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ Y_T & Y_{T-1} & Y_{T-2} & Y_{T-3} & \dots & Y_{T-p_1} \end{pmatrix}_{m \times (p_1+1)}$$

Step 3: Build an autoregressive model base on the above initial base data (B_1)

Treat the 1st column as the dependent variable and all other p columns as independent variables, and the dependent variable can be expressed in form of a linear combination of all p dependent variables, or that

$$\begin{aligned} Y_{T-m+1} &= \beta_{(1,0)} + \beta_{(1,1)}Y_{T-m} + \beta_{(1,2)}Y_{T-m-1} + \beta_{(1,3)}Y_{T-m-2} + \dots + \beta_{(1,p_1)}Y_{T-m-p_1+1} + \varepsilon_{(1,1)} \\ Y_{T-m+2} &= \beta_{(1,0)} + \beta_{(1,1)}Y_{T-m+1} + \beta_{(1,2)}Y_{T-m} + \beta_{(1,3)}Y_{T-m-1} + \dots + \beta_{(1,p_1)}Y_{T-m-p_1+2} + \varepsilon_{(1,2)} \\ Y_{T-m+3} &= \beta_{(1,0)} + \beta_{(1,1)}Y_{T-m+2} + \beta_{(1,2)}Y_{T-m+1} + \beta_{(1,3)}Y_{T-m} + \dots + \beta_{(1,p_1)}Y_{T-m-p_1+3} + \varepsilon_{(1,3)} \\ &\vdots \\ Y_T &= \beta_{(1,0)} + \beta_{(1,1)}Y_{T-1} + \beta_{(1,2)}Y_{T-2} + \beta_{(1,3)}Y_{T-3} + \dots + \beta_{(1,p_1)}Y_{T-p_1} + \varepsilon_{(1,m)} \end{aligned}$$

To minimize sum of squared errors ($\sum_{i=1}^m \varepsilon_{(1,i)}^2$), we can obtained the optimal solutions for β_1

by either using some linear programming packages or using the formula (2) stated in the INTRODUCTION.

Step 4: Forecast Y_{T+1}

Y_{T+1} can be estimated as follows:

$$\hat{Y}_{T+1} = \beta_{(1,0)} + \beta_{(1,1)}Y_T + \beta_{(1,2)}Y_{T-1} + \beta_{(1,3)}Y_{T-2} + \dots + \beta_{(1,p_1)}Y_{T-p_1+1}$$

Forecast the asset price Y_{T+2} at the time period $T+2$

Repeat the above steps except for some modifications

Step 1: Calculate p_2 (order) value using the given known figures ($Y_2, Y_3, \dots, Y_T, Y_{T+1}$) by adding Y_{T+1} as it becomes known and eliminating Y_1 as it deemed as obsolete.

Step 2: Determine the base (B_2) based on given known figures ($Y_2, Y_3, \dots, Y_T, Y_{T+1}$), since the p_2 (order) value and m (the length of base) are known

The base (B_2) is set as follows:

$$B_2 = \begin{pmatrix} Y_{T-m+2} & Y_{T-m+1} & Y_{T-m} & Y_{T-m-1} & \dots & Y_{T-m-p_2+2} \\ Y_{T-m+3} & Y_{T-m+2} & Y_{T-m+1} & Y_{T-m} & \dots & Y_{T-m-p_2+3} \\ Y_{T-m+4} & Y_{T-m+3} & Y_{T-m+2} & Y_{T-m+1} & \dots & Y_{T-m-p_2+4} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ Y_{T+1} & Y_T & Y_{T-1} & Y_{T-2} & \dots & Y_{T-p_2+1} \end{pmatrix}_{m \times (p_2+1)}$$

Step 3: Build an autoregressive model base on the above base data (B_2)

Similarly treat the 1st column as the dependent variable and all other p_2 columns as independent variables, and the dependent variable can be expressed in form of a linear combination of all p_2 dependent variables. To minimize sum of squared errors, we can obtained the optimal solutions for β_2 .

Step 4: Forecast Y_{T+2}

Y_{T+2} can be estimated as follows:

$$\hat{Y}_{T+2} = \beta_{(2,0)} + \beta_{(2,1)}Y_{T+1} + \beta_{(2,2)}Y_T + \beta_{(2,3)}Y_{T-1} + \dots + \beta_{(2,p_2)}Y_{T-p_2+2}$$

Keeping go this way to forecast the asset price Y_{T+i} at the time period $T+i$

Repeat the above steps except for some modifications

Step 1: Calculate p_i (order) value using the given known figures ($Y_i, Y_{i+1}, \dots, Y_T, Y_{T+1}, \dots, Y_{T+i-1}$) by adding Y_{T+i-1} as it becomes known and eliminating Y_{i-1} as it deemed as obsolete.

Step 2: Determine the base (B_i) based on given known figures ($Y_i, Y_{i+1}, \dots, Y_T, Y_{T+1}, \dots, Y_{T+i-1}$), since the p_i (order) value and m (the length of base) are known

The base (\mathbf{B}_i) is set as follows:

$$B_i = \begin{pmatrix} Y_{T-m+i} & Y_{T-m+i-1} & Y_{T-m+i-2} & Y_{T-m+i-3} & \cdots & Y_{T-m+i-p_i} \\ Y_{T-m+i+1} & Y_{T-m+i} & Y_{T-m+i-1} & Y_{T-m+i-2} & \cdots & Y_{T-m+i-p_i+1} \\ Y_{T-m+i+2} & Y_{T-m+i+1} & Y_{T-m+i} & Y_{T-m+i-1} & \cdots & Y_{T-m+i-p_i+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ Y_{T+i-1} & Y_{T+i-2} & Y_{T+i-3} & Y_{T+i-4} & \cdots & Y_{T+i-p_i-1} \end{pmatrix}_{m \times (p_i+1)}$$

Step 3: Build an autoregressive model base on the above base data (B_i)

Again treat the 1st column as the dependent variable and all other p_i columns as independent variables, and the dependent variable can be expressed in form of a linear combination of all p_i dependent variables. To minimize sum of squared errors, we can obtained the optimal solutions for β_i .

Step 4: Forecast Y_{T+i}

Y_{T+i} can be estimated as follows:

$$\hat{Y}_{T+i} = \beta_{(i,0)} + \beta_{(i,1)}Y_{T+i-1} + \beta_{(i,2)}Y_{T+i-2} + \beta_{(i,3)}Y_{T+i-3} + \dots + \beta_{(i,p_i)}Y_{T+i-p_i}$$

From the above processes, at each period we re-evaluate p (the number of lags) and roll the base to keep it updated, except for m , the length of base, which we will discuss in the later section. The rationales behind DAR model are that the most recently asset prices are higher valuable as inputs for our model as we deem that asset prices have “short memories”, and that p value (the number of lags) should be dynamically changed as to make the best fitting of the model.

Next we applying DAR model to make forecasts for DJI index and S&P500 index. Assuming that it is just at the end of 2007, we want to predict the DJI index and S&P 500 index for the period of January 1 to May 30, 2008 by DAR model (the result shown in Figure 1 for DJI Index and Figure 2 for S&P 500 Index)

The results are summarized as follows:

DJI Index by DAR Model (m, the length of the base, set to 420)

| <u>Avg Adj R-Sq</u> | <u>Avg p(i)</u> | <u>Max p(i)</u> | <u>Min p(i)</u> | <u>Avg Abs Error (%)</u> | <u>Std Dev (Avg Abs Error (%))</u> |
|---------------------|-----------------|-----------------|-----------------|--------------------------|------------------------------------|
| 0.9766063 | 1.625 | 11 | 1 | 0.94% | 0.0080 |

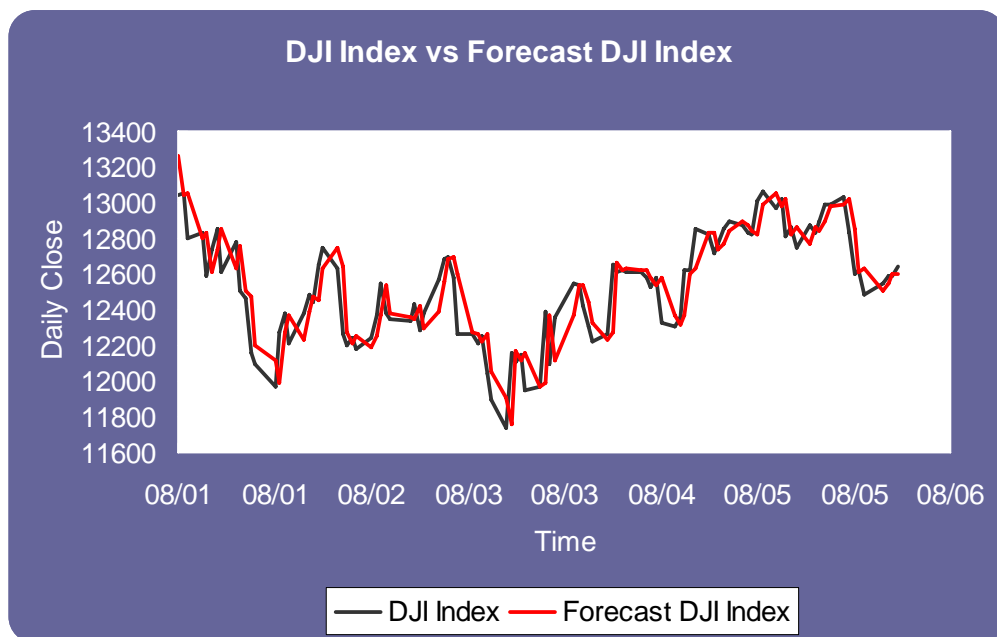
S&P 500 Index by DAR Model (m, the length of the base, set to 870)

| <u>Avg Adj R-Sq</u> | <u>Avg p(i)</u> | <u>Max p(i)</u> | <u>Min p(i)</u> | <u>Avg Abs Error (%)</u> | <u>Std Dev (Avg Abs Error (%))</u> |
|---------------------|-----------------|-----------------|-----------------|--------------------------|------------------------------------|
| 0.991316 | 1.94 | 2 | 1 | 1.03% | 0.0085 |

The average adjusted R-squares for both DJI and S&P500 indices are over 0.976, which

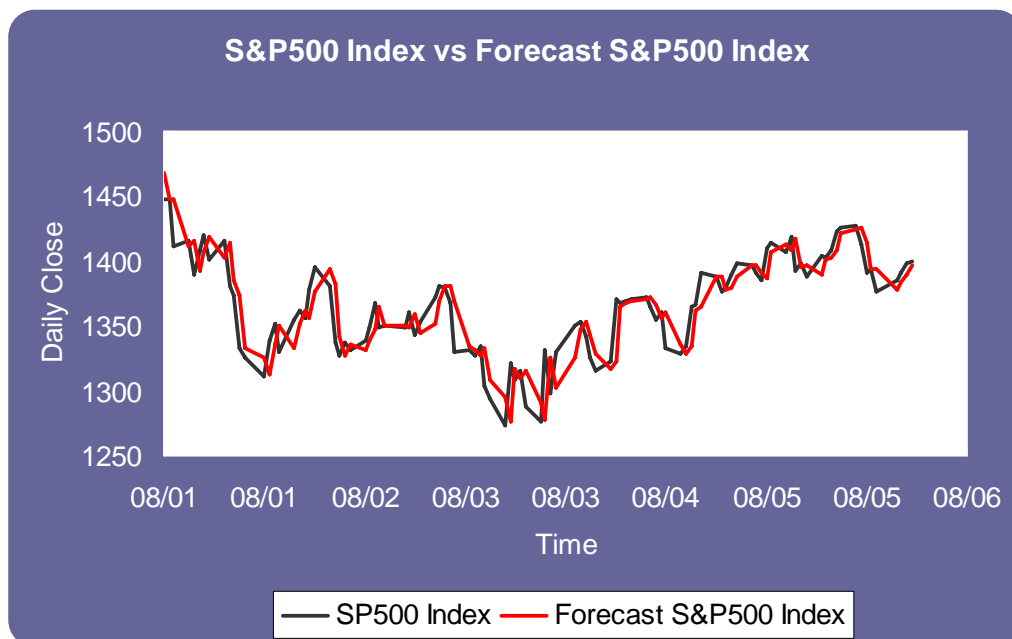
indicates a very good fitness for the DAR models. The average number of lags (p) for them are less than 2 with maximum numbers of lags of 11 for DJI index and of 2 for S&P500 index, which may explain a very “short memory”, as the indices at the time period $T+1$ are expressed in a linear combination of only a few most recent (previous) prices. The average of the absolute errors (residuals) in percentage is about 1%, which could be also deemed as a reasonably good result.

Figure 1: DJI Index (DAR model)



Note: The forecast period: January 1 – May 30, 2008

Figure 2: S&P500 Index (DAR model)



Note: The forecast period: January 1 – May 30, 2008

III CHOOSE APPROPRIATE LENGTH OF THE BASE (m)

Another key parameter for DAR model is m , the length of the base. What will be the optimal number for m ? Is the bigger m , the better outcome or the lower error? After investigations of all stocks of DJI, as well as DJI and S&P500 Indices by applying our DAR model, we find that the first question is hard to answer, and generally speaking the optimal number for m , the length of the base would be around 100-980 depending on the historical prices movements with the average of 470. But it is clear that the answer for the second question is “NO”, or that, we don’t need to make m as big as possible.

The Figure 3 & 4 shows the average absolute errors (in %) for DJI and S&P500 indices (daily close) for the period of January 1 to May 30, 2008 respectively, given all known historical indices figures by the end of year 2007, respect to m (the length of the base).

Similarly the Figure 5 & 6 show the average sum of error-squares for DJI and S&P500 indices (daily close) for the period of January 1 to May 30, 2008 respectively, given all known historical indices figures by the end of year 2007, respect to m (the length of the base).

Figure 3 & 4: DJIA and S&P500-Average Absolute Error (%) vs the length of the base

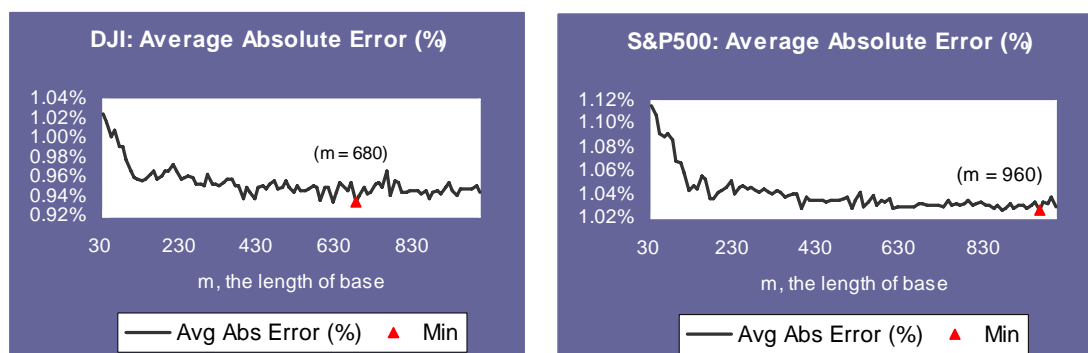
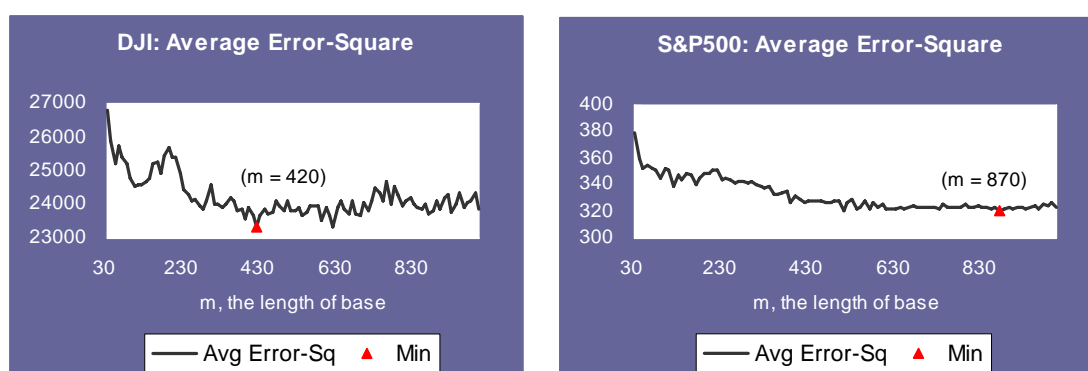


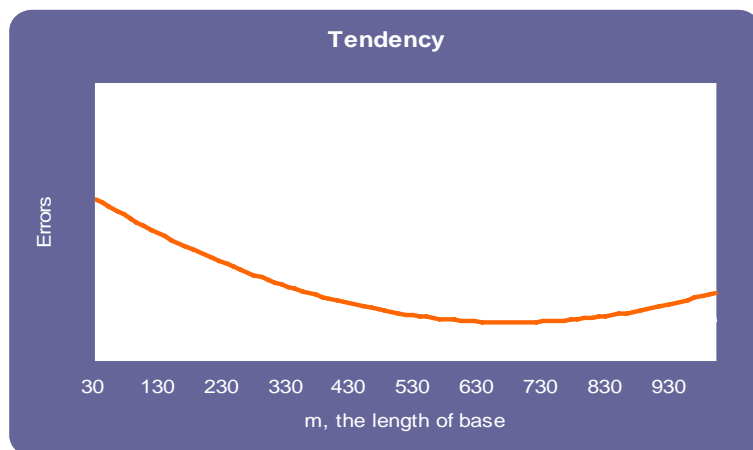
Figure 5 & 6: DJIA and S&P500-Average Error-Squares vs the length of the base



The optimal m for forecasting DJI index during the period of January 1 to May 30, 2008 with the objective to minimize the average absolute error (%) is 680, while with the objective to minimize the sum of error-squares the optimal m is 420. The optimal m for forecasting S&P500 indices during the same period with the objective to minimize the average absolute error (%) is 960, while with the objective to minimize the sum of error-squares the optimal m

is 870. All graphs show a tendency that when m moves from a smaller number up to around 400, the average errors both in terms of absolute values (in %) and sum of error-squares significantly reduced, then when m moves above around 400, the average errors in both terms appear a diminishingly decrease, and eventually increase. So the tendency curve seems to be a right-skewed “smile” shape (See Figure 7). These characteristics could be more manifest in the results of my investigations for some individual stocks by applying DAR model to forecast their prices (daily close) with the same period in the following paragraphs.

Figure 7: Tendency (right-skewed “smile” shape)



Applying DAR model to forecast stock prices (daily close) for CitiGroup (C), JP Morgan (JPM), Boeing (BA), and IBM for the period of January 1 to May 30, 2008, given all historical daily close prices before January 1, 2008, in comparison with the actual daily close price figures, we construct the graphs regarding the errors terms with respect to m , the length of the base in Figure 8 & 9.

The results are summarized as follows:

| <u>Company</u> | <u>Optimal m</u> | <u>Min Avg Abs Error (%)</u> | <u>Optimal m</u> | <u>Min Avg Error-Sq</u> |
|----------------|------------------|------------------------------|------------------|-------------------------|
| Citigroup | m = 240 | 2.66% | m = 120 | 0.7095 |
| JPMorgan | m = 630 | 2.32% | m = 130 | 1.8141 |
| Boeing | m = 420 | 1.43% | m = 690 | 2.1623 |
| IBM | m = 690 | 1.25% | m = 690 | 3.0404 |

Clearly, it is not the bigger m , the lower errors terms, which could be also interpreted that the older prices become less meaningful used for forecasting the future prices by using Dynamic Autoregressive Models (DAR). With the target of minimizing the average error terms (including average absolute error in %, and average error-squares), there will exist an optimal number for m (the length of the base), which, according to my investigations, would be in between 100 to 980, on average of around 470 (see details in the *Appendix 1*).

Figure 8

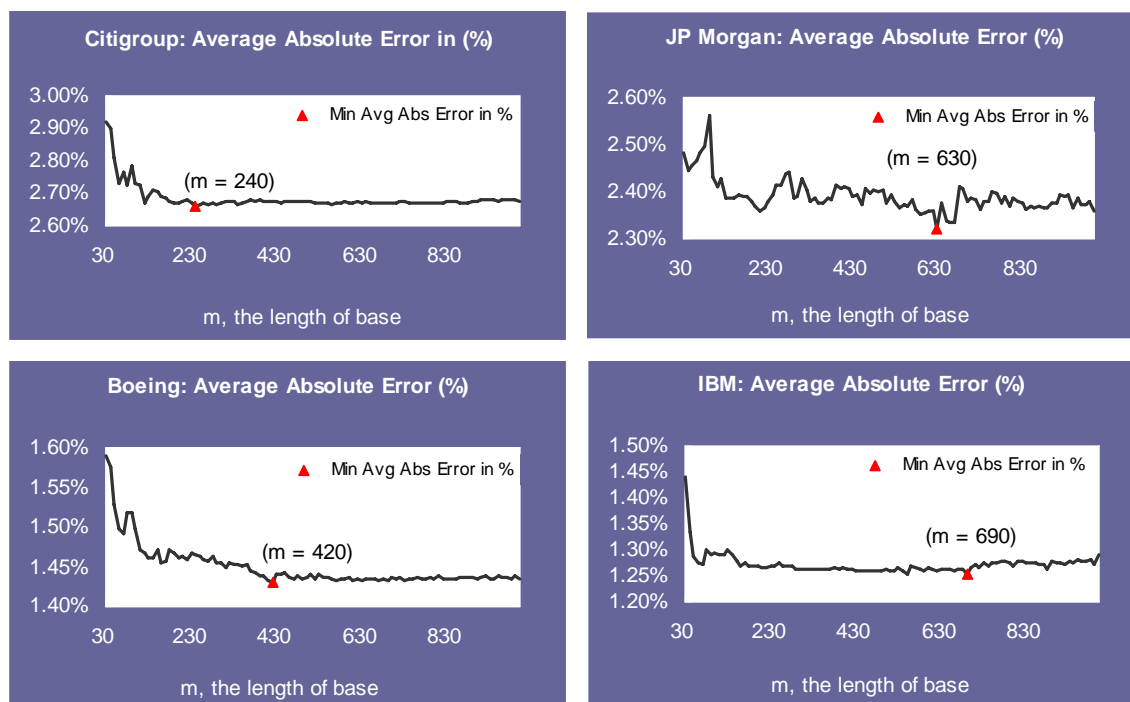
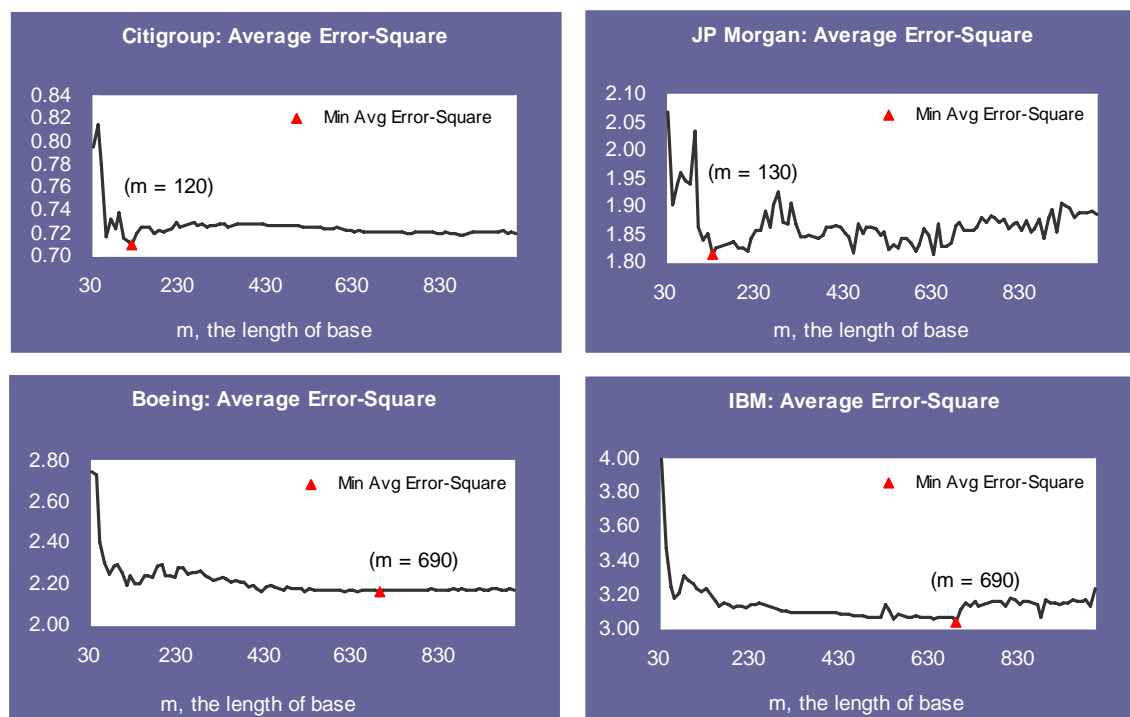


Figure 9



It is also clear that the adjusted R-squared for DJI and S&P indices are increasing with the increase of m, the length of the base (see Figure 10)

Figure 10:



Very much similar results for the stocks of DJI, for example, Citigroup, JP Morgan, Boeing and IBM (see Figure 11).

Figure 11



IV DAR vs SAR MODEL

The Static Autoregressive Model (SAR) is a simple process for forecasting the asset price at the time period $T+1$, as described in Dynamic Autoregressive Model (DAR). We just treat the p value (the number of lags) and all coefficients (betas), which are obtained from the standard autoregressive model by using the initial base (B1), as constants, to forecast the asset prices at the time period $T+1, T+2, \dots, T+i, \dots$. The forecast values are written as follows:

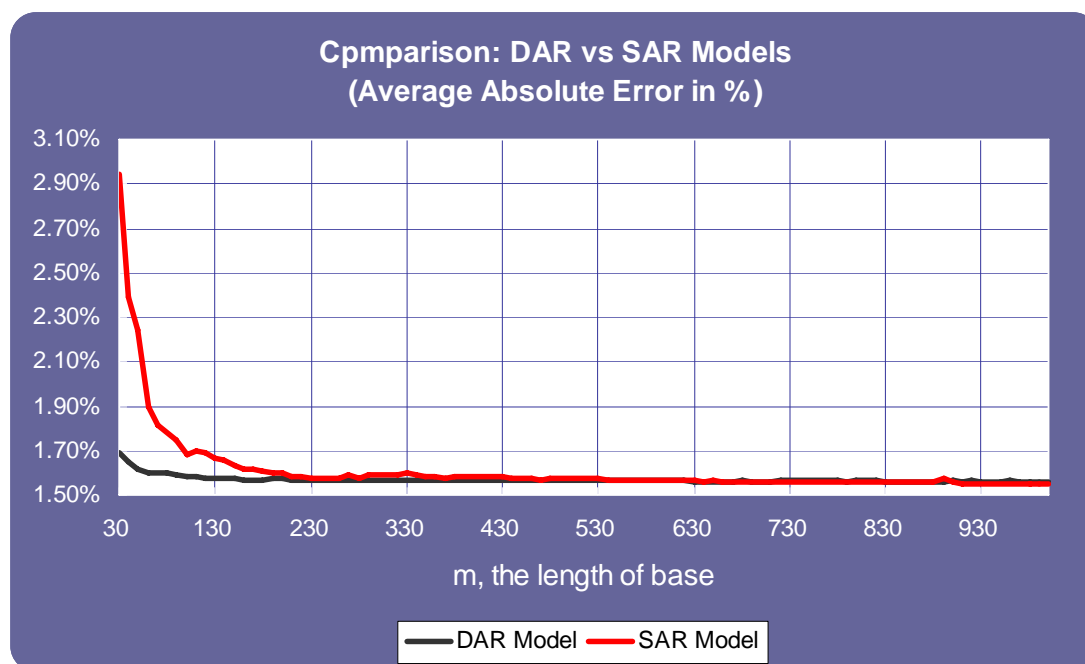
$$\hat{Y}_{T+i} = \beta_0 + \beta_1 Y_{T+i-1} + \beta_2 Y_{T+i-2} + \beta_3 Y_{T+i-3} + \dots + \beta_p Y_{T+i-p} \quad (i = 1, 2, 3, \dots)$$

Here

- p is the number of lags, from the DAR process in forecasting T+1;
- β_i ($i = 0, 1, 2, \dots, p$) are coefficients, obtained from DAR process in forecasting T+1;
- m, the length of the base, is pre-determined by analyzing the historical data, similarly when DAR model is applied.

Similarly, supposed that we are at the end of year 2007, and we have all historical data (daily close prices) for all stocks of DJI, we want to make forecasts for the future prices during the period of January 1 to May 30, 2008 for all of these stocks (also daily close price) by applying both DAR and SAR models, and then compare the estimated values generated from both DAR and SAR models to the actual daily close prices. We evaluate the results, finding that on average, DAR model outperforms SAR model by 0.8% to 40% in reducing the average of absolute error in percentage based on the lower range (30-350) of m, the length of the base, while at the upper range (350-1000) the results for both almost equal (see Figure 12). The data is given in *Appendix 2*.

Figure 12: DAR vs SAR model



Note:

All stocks of DJI (Dow Jones Industrial Average) used for the above analysis. Assuming that all stock prices (daily closed) before December 31, 2007 are known, we both apply DAR and SAR model to forecast the closed prices at the period from January 1, 2008 to May 22, 2008, and then compare with the actual daily closed prices.

V CONCLUSION

Even though the data we use for illustrations above is based on a special time period, the results for using DAR and SAR models to analyze all stocks of DJI as well as DJI & S&P indices show quite consistent and stable. We could reasonably conclude that DAR models will have advantages over SAR models in time-series forecasting by reducing the error terms, especially for the lower range of the length of the base (30-350). In other words, if there are not enough historical data available for analyzing, DAR models will most likely outperform over SAR model in time-series forecasting. With increasing the length of the base, these advantages for DAR models are diminishing and the results for both models are almost converging together. It is clear that there will be an optimal number for the length of the base (m), which means that it is not the bigger m , the better outcome. Via our investigations, on average, the length of the base should be around 400-500 (for analyzing daily close prices we may just need daily close prices for past two years). The curve of the error term with respect to the length of the base (m) shows a right skewed “smile” shape, of which the lowest point is the optimal number for the length of the base. With increase of the length of the base for both models, the R-squares show a diminishing increasing, as they tend to increase significantly within the lower range (30-350), and then tend to be flat and gradually approaching to 1 (however, 1 can never be reached).

VI. REFERENCES

- (i) *Financial Risk Manager Handbook (GARP)*, Philippe Jorion, 2007
- (ii) *The Complete Guide to Capital Markets for Quantitative Professionals*, Alex Kuznetsov, 2006
- (iii) *Statistical Analysis of Financial Data in S-Plus*, Rene A. Carmona, 2004
- (iv) *Lectures Notes*, Prof. F. Novomestky, Polytechnic Institute of New York University, 2008

APPENDICES

Appendix 1: Optimal Bases for DAR & SAR models

| Company | DAR: Opt.Base (m) | SAR: Opt.Base (m) |
|---------|-------------------|-------------------|
| AA | 340 | 210 |
| AIG | 280 | 700 |
| AXP | 660 | 950 |
| BA | 690 | 70 |
| C | 120 | 70 |
| CAT | 640 | 250 |
| DD | 400 | 110 |
| DIS | 880 | 520 |
| GE | 760 | 990 |
| GM | 120 | 70 |
| HD | 130 | 40 |
| HON | 100 | 180 |
| HPQ | 210 | 190 |
| IBM | 690 | 490 |
| INTC | 980 | 530 |
| JNJ | 400 | 530 |
| JPM | 130 | 140 |
| KO | 120 | 120 |
| MCD | 250 | 190 |
| MMM | 880 | 880 |
| MRK | 930 | 970 |
| MSFT | 450 | 640 |
| PFE | 270 | 950 |
| PG | 910 | 960 |
| T | 750 | 1000 |
| UTX | 410 | 700 |
| VZ | 810 | 770 |
| WMT | 160 | 990 |
| XOM | 170 | 150 |
| Average | 470 | 495 |
| Min | 100 | 40 |
| Max | 980 | 1000 |

Appendix 2:

Average R-squares and Absolute Error (%) Comparisons for DAR and SAR Models

| Length of Base m | Average R-Square | | Average Absolute Error in Percentage | | |
|---------------------|------------------|--------|--------------------------------------|-------|----------------|
| | DAR | SAR | DAR | SAR | Outperform (%) |
| 30 | 0.6834 | 0.6330 | 1.69% | 2.94% | 42.5% |
| 40 | 0.7441 | 0.7210 | 1.65% | 2.39% | 31.1% |
| 50 | 0.7854 | 0.7686 | 1.62% | 2.25% | 28.0% |
| 60 | 0.8173 | 0.8387 | 1.60% | 1.89% | 15.3% |
| 70 | 0.8427 | 0.8578 | 1.60% | 1.82% | 11.9% |
| 80 | 0.8627 | 0.8679 | 1.60% | 1.78% | 10.1% |
| 90 | 0.8784 | 0.8772 | 1.59% | 1.75% | 8.7% |
| 100 | 0.8908 | 0.8817 | 1.58% | 1.68% | 6.2% |
| 110 | 0.8985 | 0.8760 | 1.58% | 1.70% | 6.9% |
| 120 | 0.9036 | 0.8891 | 1.57% | 1.69% | 6.6% |
| 130 | 0.9082 | 0.8978 | 1.57% | 1.67% | 5.8% |
| 140 | 0.9128 | 0.9045 | 1.57% | 1.66% | 5.0% |
| 150 | 0.9172 | 0.9116 | 1.57% | 1.63% | 3.6% |
| 160 | 0.9211 | 0.9174 | 1.57% | 1.62% | 3.0% |
| 170 | 0.9241 | 0.9225 | 1.57% | 1.61% | 2.9% |
| 180 | 0.9270 | 0.9311 | 1.57% | 1.60% | 2.2% |
| 190 | 0.9303 | 0.9388 | 1.57% | 1.60% | 1.6% |
| 200 | 0.9336 | 0.9437 | 1.57% | 1.60% | 1.6% |
| 210 | 0.9371 | 0.9487 | 1.57% | 1.58% | 0.8% |
| 220 | 0.9406 | 0.9497 | 1.57% | 1.58% | 0.8% |
| 230 | 0.9437 | 0.9514 | 1.57% | 1.58% | 0.7% |
| 240 | 0.9465 | 0.9531 | 1.56% | 1.58% | 0.8% |
| 250 | 0.9492 | 0.9550 | 1.57% | 1.58% | 0.6% |
| 260 | 0.9518 | 0.9562 | 1.57% | 1.58% | 0.7% |
| 270 | 0.9541 | 0.9564 | 1.57% | 1.59% | 1.4% |
| 280 | 0.9563 | 0.9572 | 1.56% | 1.58% | 0.7% |
| 290 | 0.9580 | 0.9581 | 1.56% | 1.59% | 1.5% |
| 300 | 0.9594 | 0.9599 | 1.57% | 1.59% | 1.5% |
| 310 | 0.9606 | 0.9615 | 1.57% | 1.59% | 1.4% |
| 320 | 0.9617 | 0.9636 | 1.57% | 1.59% | 1.4% |
| 330 | 0.9628 | 0.9658 | 1.57% | 1.60% | 1.7% |
| 340 | 0.9640 | 0.9683 | 1.57% | 1.59% | 1.2% |
| 350 | 0.9652 | 0.9699 | 1.57% | 1.58% | 0.7% |
| 360 | 0.9663 | 0.9715 | 1.56% | 1.59% | 1.3% |
| 370 | 0.9674 | 0.9727 | 1.57% | 1.58% | 0.6% |
| 380 | 0.9685 | 0.9736 | 1.57% | 1.58% | 1.0% |
| 390 | 0.9696 | 0.9747 | 1.57% | 1.58% | 1.2% |
| 400 | 0.9707 | 0.9754 | 1.57% | 1.58% | 0.8% |

Dynamic vs Static Autoregressive Models for Forecasting Time Series

| | | | | | |
|-----|--------|--------|-------|-------|-------|
| 410 | 0.9717 | 0.9758 | 1.57% | 1.58% | 0.9% |
| 420 | 0.9727 | 0.9761 | 1.56% | 1.58% | 1.0% |
| 430 | 0.9735 | 0.9765 | 1.57% | 1.58% | 0.8% |
| 440 | 0.9743 | 0.9772 | 1.56% | 1.58% | 0.8% |
| 450 | 0.9750 | 0.9776 | 1.56% | 1.57% | 0.6% |
| 460 | 0.9756 | 0.9781 | 1.57% | 1.57% | 0.3% |
| 470 | 0.9761 | 0.9785 | 1.57% | 1.57% | 0.2% |
| 480 | 0.9766 | 0.9788 | 1.57% | 1.57% | 0.1% |
| 490 | 0.9770 | 0.9790 | 1.57% | 1.57% | 0.3% |
| 500 | 0.9774 | 0.9793 | 1.57% | 1.57% | 0.3% |
| 510 | 0.9777 | 0.9797 | 1.57% | 1.57% | 0.3% |
| 520 | 0.9781 | 0.9801 | 1.57% | 1.57% | 0.3% |
| 530 | 0.9785 | 0.9807 | 1.57% | 1.57% | 0.3% |
| 540 | 0.9789 | 0.9811 | 1.57% | 1.57% | 0.2% |
| 550 | 0.9793 | 0.9814 | 1.57% | 1.57% | 0.2% |
| 560 | 0.9796 | 0.9816 | 1.57% | 1.56% | -0.2% |
| 570 | 0.9800 | 0.9818 | 1.56% | 1.57% | 0.1% |
| 580 | 0.9803 | 0.9821 | 1.56% | 1.56% | 0.1% |
| 590 | 0.9807 | 0.9822 | 1.57% | 1.56% | -0.1% |
| 600 | 0.9810 | 0.9823 | 1.57% | 1.57% | 0.0% |
| 610 | 0.9813 | 0.9825 | 1.56% | 1.57% | 0.2% |
| 620 | 0.9816 | 0.9827 | 1.56% | 1.57% | 0.1% |
| 630 | 0.9817 | 0.9813 | 1.56% | 1.57% | 0.5% |
| 640 | 0.9818 | 0.9818 | 1.56% | 1.56% | -0.1% |
| 650 | 0.9818 | 0.9806 | 1.56% | 1.56% | 0.2% |
| 660 | 0.9818 | 0.9813 | 1.56% | 1.56% | -0.2% |
| 670 | 0.9818 | 0.9816 | 1.56% | 1.56% | -0.1% |
| 680 | 0.9818 | 0.9819 | 1.56% | 1.56% | -0.4% |
| 690 | 0.9818 | 0.9821 | 1.56% | 1.56% | -0.3% |
| 700 | 0.9818 | 0.9822 | 1.56% | 1.55% | -0.3% |
| 710 | 0.9818 | 0.9826 | 1.56% | 1.56% | 0.0% |
| 720 | 0.9818 | 0.9829 | 1.56% | 1.56% | -0.2% |
| 730 | 0.9819 | 0.9831 | 1.56% | 1.56% | 0.0% |
| 740 | 0.9821 | 0.9833 | 1.56% | 1.56% | -0.3% |
| 750 | 0.9823 | 0.9835 | 1.56% | 1.56% | -0.2% |
| 760 | 0.9826 | 0.9837 | 1.56% | 1.56% | -0.2% |
| 770 | 0.9828 | 0.9836 | 1.56% | 1.56% | -0.2% |
| 780 | 0.9830 | 0.9837 | 1.56% | 1.56% | -0.2% |
| 790 | 0.9832 | 0.9839 | 1.56% | 1.56% | -0.3% |
| 800 | 0.9833 | 0.9840 | 1.56% | 1.56% | -0.2% |
| 810 | 0.9835 | 0.9840 | 1.56% | 1.56% | -0.3% |
| 820 | 0.9836 | 0.9841 | 1.56% | 1.55% | -0.6% |
| 830 | 0.9837 | 0.9843 | 1.56% | 1.56% | -0.1% |

Dynamic vs Static Autoregressive Models for Forecasting Time Series

| | | | | | |
|------|--------|--------|-------|-------|-------|
| 840 | 0.9838 | 0.9846 | 1.56% | 1.55% | -0.4% |
| 850 | 0.9840 | 0.9849 | 1.56% | 1.56% | -0.1% |
| 860 | 0.9841 | 0.9850 | 1.56% | 1.56% | -0.2% |
| 870 | 0.9842 | 0.9851 | 1.56% | 1.56% | -0.1% |
| 880 | 0.9844 | 0.9853 | 1.56% | 1.56% | -0.1% |
| 890 | 0.9845 | 0.9819 | 1.56% | 1.57% | 0.9% |
| 900 | 0.9844 | 0.9840 | 1.56% | 1.56% | -0.4% |
| 910 | 0.9844 | 0.9848 | 1.56% | 1.55% | -0.7% |
| 920 | 0.9845 | 0.9854 | 1.56% | 1.55% | -0.6% |
| 930 | 0.9847 | 0.9857 | 1.56% | 1.55% | -0.7% |
| 940 | 0.9848 | 0.9860 | 1.56% | 1.55% | -0.8% |
| 950 | 0.9849 | 0.9863 | 1.56% | 1.55% | -0.6% |
| 960 | 0.9851 | 0.9864 | 1.56% | 1.55% | -0.7% |
| 970 | 0.9852 | 0.9865 | 1.56% | 1.55% | -0.6% |
| 980 | 0.9853 | 0.9866 | 1.56% | 1.55% | -0.7% |
| 990 | 0.9855 | 0.9867 | 1.56% | 1.55% | -0.7% |
| 1000 | 0.9857 | 0.9867 | 1.56% | 1.55% | -0.7% |