Modeling and Studying the Stylized Facts of VIX Index

Lin Zhu

F1 Mansion 27, Lane 218, Wuyang East Road, Jiangqiao Town, Jiading Area, Shanghai, China

Zhulin869@163.com

Keywords: VIX; Stylized facts; HAR model

Abstract: VIX Index is computed on a real-time basis through every working days (Whaley 2009). It was also called “investor fear gauge”, since we can use it to judge the level of investors’ expectation and fear. The VIX can also help investors forecast the market trend, and adjust technical analysis and trading strategy, since it can reflect the fluctuation of the S&P 500 (SPX) index. This paper will study the stylized facts of VIX and then using heterogeneous autoregressions (HAR) model for modeling with VIX data.

1. Introduction

Firstly, VIX was introduced by the CBOE (Chicago Board Options Exchange) in 1993. Actually, starting in 2003, the CBOE reported two kinds of market volatility index. One of them is the VXO which implied volatility of S&P 100 index options, and another kind of market volatility index is the VIX[1].

With the rapid development of modern financial markets, the property income becomes more important. It is crucial for investors to know more about all kinds of derivatives, and stock is one of the most popular derivatives. Participants in the stock market include individual investors and larger trader investors, such as insurance companies and hedge funds. Setty, Rangaswamy and Subramanya (2010) suggests stock market can reflect a country's economic strength and development[2]. Market volatility is essential for investors to study regularity of stock and making beneficial financial decisions. One of the most popular market volatility is VIX(Volatility Index), As a barometer of investor fear, VIX is a useful way to measure market expectations of volatility (McAleer et al. 2011) [3].

VIX becomes important in modern financial markets due to it is forward-looking. In attempting to understand VIX, Whaley (2009) emphasizes that VIX is not backward-looking, it can measure volatility that investors expect to know, instead of measuring volatility that has been recently realized. This paper will introduce the background for the VIX index, then studying the stylized facts of VIX and show the descriptive statistics. After ensuring properties of VIX, we use heterogeneous autoregressions (HAR) model for modeling with VIX data.

2. Stylized facts of VIX

According to the data, I use Stata to study the stylized facts of VIX. Firstly, I collect samples of documentation and process these data. I download the data from daily VIX index and SPX from CBOE website. The study runs from 01/02/2004 to 12/01/2017. Each sample includes 3505 observations. The descriptive statistics for the VIX and SPX are presented in Table 1.

Table 1 descriptive statistics

	VIX	SPX
Number of observations	3505	3505
Mean	18.67616	790.8824
Standard deviation	9.003793	367.8831
Minimum	9.25	295.46
Maximum	78.665	1527.46

  DOI: 10.25236/isrme.2019.020

Then importing the data and plotting VIX and SPX in Figure 1 and 2. Figure 1 reflects the time-series plot of the VIX index from 2004 to 2017. The time series fluctuate dramatically in the short term. However, they generally keep steady in the long term. In the most time, VIX fluctuates between 10 and 20. It shows the mean reversion of VIX (Drimus, and Farkas 2013) [4]. In other words, VIX index will revert to the normal mean of value after it deviates from the steady value. On the other hand, VIX seems to fluctuates over the whole period, it is more stable when index values are low and it is quite unstable with high index values (Fernandes, Medeiros and Scharth 2014) [5]. Figure 2 shows that the negative correlation between VIX and SPX is strong (Szado 2009) [6]. VIX increases at a lower rate when the SPX rises than when it falls. Szado (2009) suggests that the negative correlation between VIX and SPX is strongest in large down moves. In 2008, as a result of the Financial Crisis, the SPX dropped significantly, and the VIX increased to a peak which equals approximately to 80. In this period, the financial market of America suffers serious destruction especially stock market. Foerster (2014) pointed out that the VIX usually increases during recoveries, however these increases are temporary, and then the VIX decreased to the previous level[7].

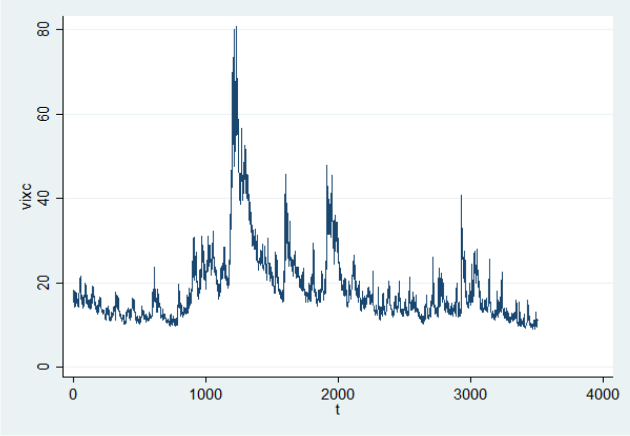

Figure 1

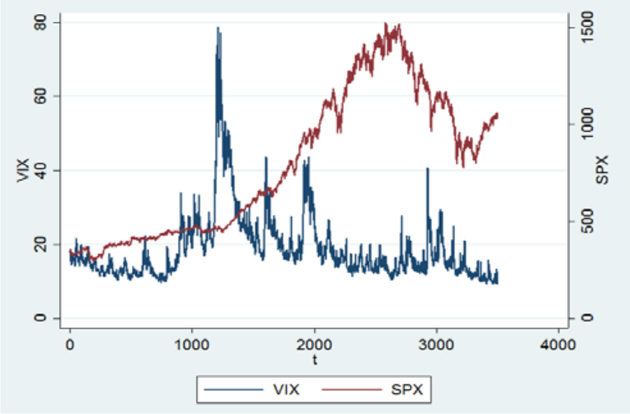

Figure 2

Then looking at the autocorrelation function of data to research the VIX index. Figure 3 shows that there is high persistence even 150 lags. It dies out slowly and this phenomenon lasts for a long period. Figure 4 shows that the square of VIX also dies out slowly, and with significant dependence.

Finally, we look at the empirical distribution of VIX. In this process, we use the Stata to plot (Figure 5 and 6). The VIX index is always positive, so the distribution of it is skewed to the right. Especially, in the second half of the data, the skewness coefficient increases. Comparing with it, using logarithm transformation to the VIX, the distribution of logarithm of VIX looks more symmetric. Although most of the excessive kurtosis have been solved, it is still skewed and non-normal. These histograms show that there are high level of skewness in the data of VIX.

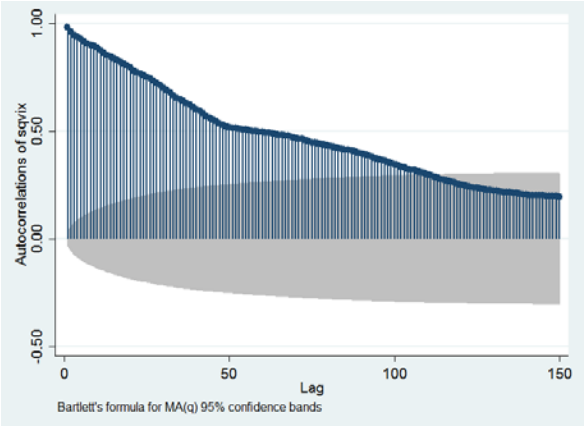

Figure 3

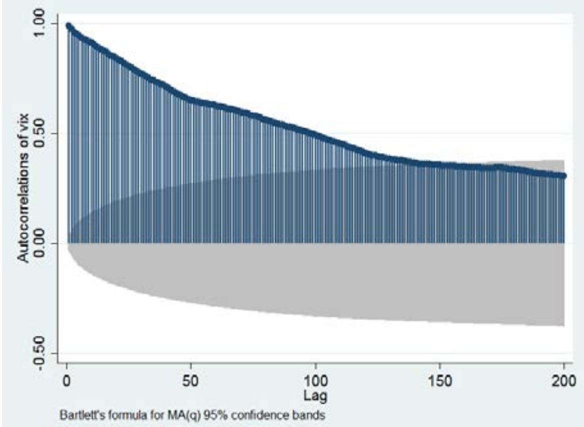

Figure 4

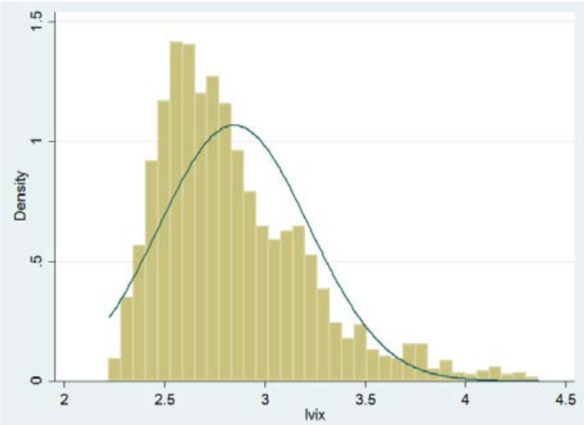

Figure 5

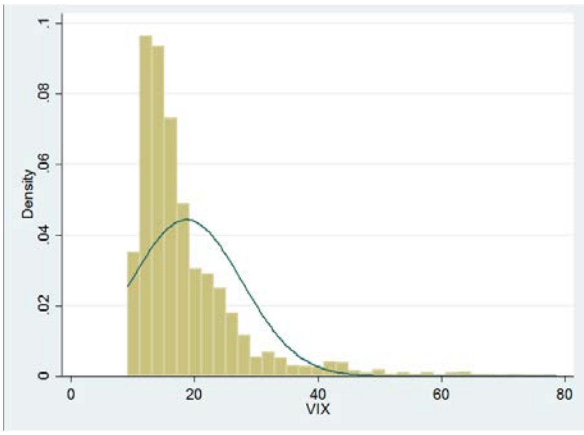

Figure 6

To summarize, this paper lists the following stylized facts for the VIX.
a.VIX index has the characteristic of mean reversion.
b.The negative correlation between VIX and SPX is strong.
c.VIX Index is persistent.
d.The distribution of VIX and logarithm of VIX are all skewed.

3. Modeling VIX Index

If VIX can reasonably be treated as observable, then it can be modeled using standard time series tools such as AR models. From figure 3, VIX is highly persistent and we would need a high order AR model for it, which is not a good way. A heterogeneous autoregressive model (HAR) can be used to study the realized volatility, and HAR model can capture the long-range dependence that arises from the asymmetric propagation of volatility between long and short horizons (Corsi 2009)[8].

Before modeling the VIX index and the logarithm of VIX index, I use Augmented Dickey-Fuller (ADF) test to check their stationarity. Results (Table 2 and 3) show that the null hypothesis can be rejected at the 99% statistical level for the full sample, since the p-values are all equal to 0, thus we conclude there is no unit root.

Table 2 ADF test (VIX)

VIX	Coefficient	Std.Err.	t	P>|t|
L1.	-0.0134	0.0026	-5.08	0
LD.	0.0980	0.0168	5.82	0

Notes: p-value for z(t) = 0

Table 3 ADF test (logarithm of VIX)

Log VIX	Coefficient	Std.Err.	t	P>|t|
L1.	-0.0123	0.0245	-5.04	0
LD.	0.1594	0.0167	9.55	0

Notes: p-value for z(t) = 0

HAR model is appropriate to modeling and forecasting the VIX index. Typically, when we calculate the current VIX index, we can use the data in the previous day, the average data of the previous trading week which means 5 days, and the average data of the previous trading month which means 22 days. Then we got the model Eq.1, where Eq.1 and Eq.2, or in logs we have Eq.4.

$$VIX_t = \varphi_0 + \varphi_1 VIX_{t-1} + \varphi_5 \overline{VIX}_{t-5} + \varphi_{22} \overline{VIX}_{t-22} + \varepsilon_t \tag{1}$$

$$\overline{VIX}_{t-5} = \frac{1}{5} \sum_{i=1}^{5} VIX_{t-i} \tag{2}$$

$$\overline{VIX}_{t-22} = \frac{1}{22} \sum_{i=1}^{22} VIX_{t-i} \tag{3}$$

$$log\,VIX_t = \varphi_0 + \varphi_1\, log\,VIX_{t-1} + \varphi_5\, \overline{log\,VIX}_{t-5} + \varphi_{22}\, \overline{log\,VIX}_{t-22} + \varepsilon_t \tag{4}$$

The results are respectively showed in Table 4, 5 and 6.

Table 4 Regression results

Variables	Coefficient	Std.Err.	t	P>|t|
VIXL1	0.9565	0.0158	60.62	0.000
VIX_5	0.0007	0.0202	0.04	0.972
VIX_{22}	0.0335	0.0106	3.17	0.000
Constant	0.1723	0.0570	3.03	0.000

Notes: R-squared=0.9753, Adj R-squared=0.9753

In table 4, we can find that the p-value of $\overline{VIX}_{t-5} = 0.972$. This means that it is not significant. So we add another one variable RVt-10 to study this model, and it represents the biweekly components.

Table 5 Regression results

Variables	Coefficient	Std.Err.	t	P>|t|
$VIXL_1$	0.9711	0.0160	60.65	0.000
VIX_5	-0.1156	0.0312	-3.70	0.000
VIX_{10}	0.1609	0.0330	4.87	0.000
VIX_{22}	-0.0265	0.0162	-1.64	0.102
Constant	0.1870	0.0568	3.29	0.001

Notes: R-squared=0.9755, Adj R-squared=0.9754

In table 5, the p-value of $\overline{VIX}_{t-22} = 0.1$ is approximately equals to 0.1. The results mean that it is significant at the level of 10% significance.

Table 6 Regression results

Variables	Coefficient	Std.Err.	t	P>|t|
$LVIXL_1$	1.0015	0.0152	65.85	0.000
$LVIX_5$	-0.0639	0.0193	-3.30	0.001
$LVIX_{22}$	0.0537	0.0103	5.23	0.000
Constant	0.0246	0.0079	3.13	0.002

Notes: R-squared=0.9758, Adj R-squared=0.9758

Table 6 presents the results of modelling of logarithms of VIX. The p-value of all parameters are similar to 0. It means that at the level of 5% significance, all parameters are significant. On the other hand, R-squared (coefficient of determination) and Adj R-squared (adjusted coefficient of determination) exceed 0.95, which represents a great fit of the sample regression to the data. Last but not least, only the coefficient of $logVIX_{t-1}$ is approximately 1. It means the data on the previous day have the most significant effects.

4. Comments and further research

Although this paper has studied the time-series properties of the VIX index, there are still some insufficiency about researching market volatility and VIX index.

CBOE can also calculate several other volatility indices which as a hedging instrument. For example, Psaradellis and Sermpinis (2016) propose another two kinds of volatility index, the VXN and the VXD [9]. As forward-looking indicators, VIX, VXN and VXD can all represent the expected future market volatility. We can compare VIX with VXN and VXD for further research about market volatility in the future.

According to the research of Corsi (2009), we use HAR model for modeling. HAR model is just a linear autoregression. An ARFIMA model can also be used to study capture the properties of the volatility index. ARFIMA (1, d, 1) performs better than VAR models and other simple linear models based on economic variables for forecasting the US implied volatility indices (Konstantinidi et al. 2008) [10]. We can also use the HAR-GASVR (res) model to research the VIX index. As the most accurate techniques for forecasting the VIX, it can achieve better results than the HAR methods

References

[1] R. E. Whaley: Journal of Portfolio Management, Vol. 35 (2009) No.3, p.98.

[2] D. V. Setty, T. M. Rangaswamy and K. N. Subramanya: International Journal of Computer Applications, Vol. 1 (2010) No.3, p.88.

[3] M. McAleer, C. L. Chang, J. A. Jimenez-Martin and T. Pérez-Amaral: The North American Journal of Economics and Finance (2011) No.25, p.151.

[4] G. Drimus and W. Farkas: Review of Derivatives Research, Vol. 16 (2013) No.3, p.267.

[5] M. Fernandes, M. C. Medeiros and M. Scharth: Journal of Banking and Finance, (2014) No.40, p.1.

[6] E. Szado: The Journal of Alternative Investments, Vol.12 (2009) No.2, p.68.

[7] A. Foerster: Economic Review, Vol.19 (2014) No.5, p.551.

[8] F. Corsi: Journal of Financial Econometrics, (2009) No.7, p.174.

[9] I. Psaradellis and G. Sermpinis: International Journal of Forecasting, Vol. 32 (2016) No.4, p.1268.

[10] E. Konstantinidi, G. Skiadopoulos and E. Tzagkaraki: Journal of Banking and Finance, Vol.32 (2008) No.11, p.2401.