

High-Frequency Financial Econometrics: Final Exam

Wanxin Chen

December 5, 2018

0.1 Exercise 1

0.1.1 A

$t = 430$ in both of my stocks, HD and VZ, for the date September 16, 2008. Since I'm in the morning of September 16, 2008 and I don't what will happen, I can not use that day's data to calculate RV or TV to estimate IV. I can only use past data to forecast IV.

0.1.2 B

1. For stock HD, TV of September.15th, 2008, one day before Sep.16th, was 0.0011, which was quite big. TVs for a week ago, from Sep.8th to Sep.12th were relatively small. Looking back at that time, we find that Lehman Brothers went bankruptcy on Sep.15th. Thus, I prefer the forecasting model only use the TV yesterday. I may only use yesterday, Sep. 15th, TV to forecast today's IV (White-Noise model). Or I may choose to use past one year's data to get AR or ARQ model to forecast IV (Instead of using RV_{t-1} as before we may use TV_{t-1}). To simplify the analysis, we use the White-Noise model here. The $TV_{t-1} = 0.0011$. This TV was quite big and it is the approxiamtion of IV_{t-1} . It gives $AnnualizedTV = 100 * \sqrt{0.0011 * 252} = 52.65$. 52.65% volatility is quite high and it may due to some bad events.

2. For stock VZ, TV of September.15th, 2008, one day before Sep.16th, was $6.4635 * 10^{-4}$, which was relatively big. TVs for a week ago, from Sep.8th to Sep.12th were relatively small. Looking back at that time, we find that Lehman Brothers went bankruptcy on Sep.15th. Thus, I prefer the forecasting model only use the TV yesterday. I may only use yesterday, Sep. 15th, TV to forecast today's IV (White-Noise model). Or I may choose to use past one year's data to get AR or ARQ model as described to forecast IV (Instead of using RV_{t-1} as before we may use TV_{t-1}). To simplify the analysis, we use the White-Noise model here. The $TV_{t-1} = 6.4635 * 10^{-4}$. This TV was relatively big and it is the approxiamtion of IV_{t-1} . It gives $AnnualizedTV = 100 * \sqrt{6.4635 * 10^{-4} * 252} = 40.36$. 40.36% volatility is high and it may due to some bad events.

$$AR : \hat{TV}_t = \hat{\beta}_0 + \hat{\beta}_1 TV_{t-1}$$

$$ARQ : \hat{TV}_t = \hat{\beta}_0 + \hat{\beta}_1 TV_{t-1} + \hat{\beta}_{1Q} Q \hat{IV}_{t-1}^{1/2} TV_{t-1}$$

0.1.3 C

This is the table of estimate of probabilities of today's HD and VZ stock return less than -2 percent and -4 percent. We know return of the day follows normal distribution with mean 0 and variance IV_t and we use TV_{t-1} to forecast IV_t , so we can use the following equation to get the estimates.

$$P(r_t^c < -q) = P\left(\frac{r_t^c - 0}{\sqrt{TV_{t-1}}} < \frac{-q}{\sqrt{TV_{t-1}}}\right)$$

Stock	q	Estimate	Lower bound	Upper bound
HD	2%	27.65%	17.04%	32.01%
HD	4%	11.77%	2.84%	17.50%
VZ	2%	21.57%	15.80%	25.19%
VZ	4%	5.78%	2.24%	9.06%

Table 1: Stock Estimate of $P(r_t^c \leq -q)$

For stock HD, for probability of losing 2 million on 100 million of dollars, the confidence interval width is 0.1497, which indicates the estimate is not accurate but may still be useful for some applications. For example, we get to know that the today's return gonna be less than -2 percent, or say loss 2 million of dollars, with probability at most 32.01% with 95% confidence. For probability of losing 4 million, the confidence interval width is 0.1466, which indicates the estimate is not accurate but may still be useful for

some applications. For example, we get to know that the today's return gonna be less than -4 percent with probability at least 2.84% with 95% confidence.

For stock VZ, for probability of losing 2 million on 100 million of dollars, the confidence interval width is 0.0939, which indicates the estimate is not really accurate but is useful for some applications. For example, we get to know that the today's return gonna be less than -2 percent, or say loss 2 million of dollars with probability at most 25.19% with 95% confidence. For probability of losing 4 million, the confidence interval width is 0.0682, which indicates the estimate is accurate and will be useful for many applications. For example, we get to know that the today's return gonna be less than -4 percent with probability at least 2.24% with 95% confidence.

0.1.4 D

This is the table of estimate of VaR at 1 percent probability. Actually, the returns I used are all in decimal so the equations are following. Also, as convention that the VaR is positive, I computed the absolute value of $Q_{0.01}^{100}$ to get VaR.

$$P(r_t^c * V \leq Q_{0.01}^{100}) = p$$

$$P\left(\frac{V * (r_t^c - 0)}{\sqrt{TV_{t-1}}} < \frac{Q_{0.01}^{100}}{\sqrt{TV_{t-1}}}\right) = p$$

$$P\left(\frac{r_t^c - 0}{\sqrt{TV_{t-1}}} < \frac{Q_{0.01}^{100}}{V * \sqrt{TV_{t-1}}}\right) = p$$

-

Stock	p	VaR(millions)	Lower bound VaR(millions)	Upper bound VaR(millions)
HD	1%	7.8408	4.8834	9.9553
VZ	1%	5.9135	4.6387	6.9585

Table 2: Stock Estimate of VaR at $p = 1\%$

As we can see the VaR spread for stock HD is 5.0719 million of dollars, which I think is relatively small ($\frac{5.0719}{100} = 5.0719\%$). Thus, I think the estimate of VaR is accurate. It may be useful in some scenarios. For example, we know that we are not going to lose more than 9.9553 million of dollars with the probability 99% with 95% confidence if we put all the 100 million dollar into stock HD. My boss may get some intuitions of our possible loss through these estimates and make some decisions.

As we can see the VaR spread for stock VZ is 2.3198 million of dollars, which I think is quite small ($\frac{2.3198}{100} = 2.3198\%$). Thus, I think the estimate of VaR is quite accurate. It can be helpful in many scenarios. For example, we know that we are not going to lose more than 6.9585 million of dollars with the probability 99% with 95% confidence if we put all the 100 million dollar into stock VZ. My boss can get accurate intuitions of our possible loss through these estimates and make important decisions.

0.2 Exercise 2

0.2.1 A

I chose 20150122_66, which has 66 put options prices on Jan.22nd 2015. I chose it randomly.

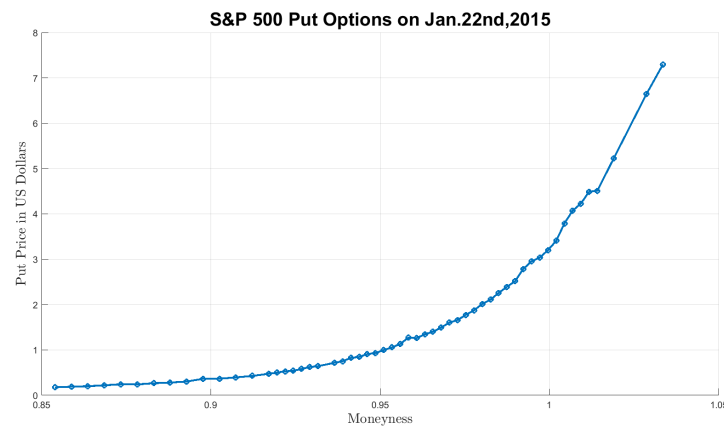
0.2.2 B

I used the 5min data of SPY and calculated RV for each day. Then I converted RV into annualized volatility σ .

$$\sigma = \sqrt{252 * RV}.$$

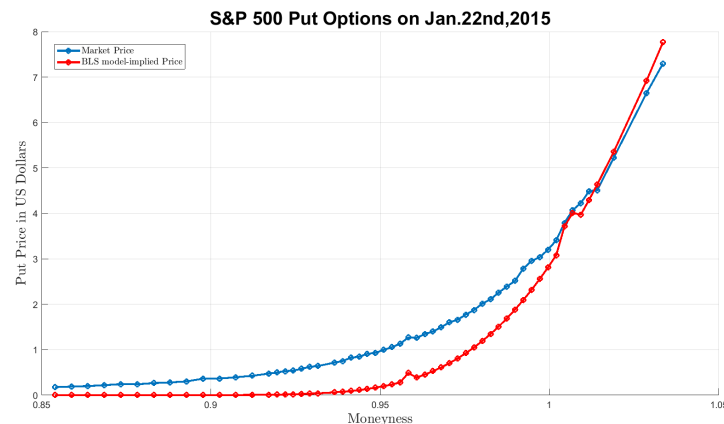
0.2.3 C

This is the market prices of SPY put options against their moneyness on Jan.22nd, 2015. We can spot that the put prices go up as put options go from out of the money to in the money. It is plausible because if the option goes far out of the money, which means strike price is much lower than today's price, then it's harder for stock to hit that strike price, so people won't want to pay much for the right to sell the stock at that low price and the put option price is low. On the contrary, if the option goes in the money, which means strike price is much higher than today's price, then it's also harder for stock to hit that strike price, so people want to pay much for the right to sell the stock at such a high price so the put option price is high.



0.2.4 D

This is the market prices and BLS model-implied prices of SPY put options against their moneyness on Jan.22nd, 2015. We notice that they have the same trend but BLS model-implied prices are lower than market prices when the put options are out of the money.

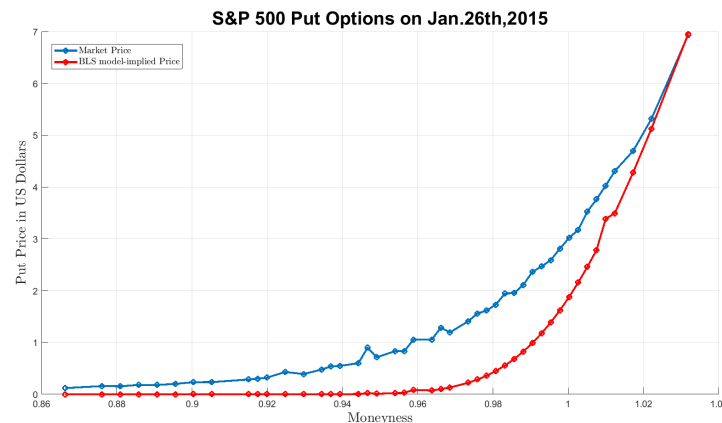


0.2.5 E

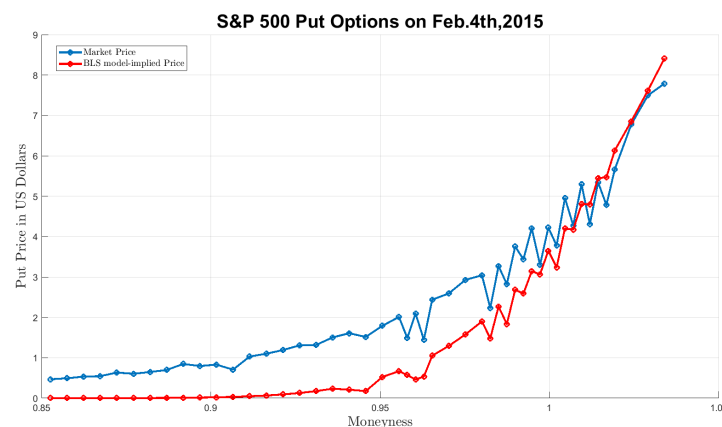
My plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Jan.22nd, 2015 is consistent with this finding. The BLS model-implied prices are lower than market prices when the put options are far out of the money. Also, when put options are close to money, which means moneyness close to one, the BLS model-implied prices are quite close to market prices. These two characters are consistent with the previous findings.

0.2.6 F

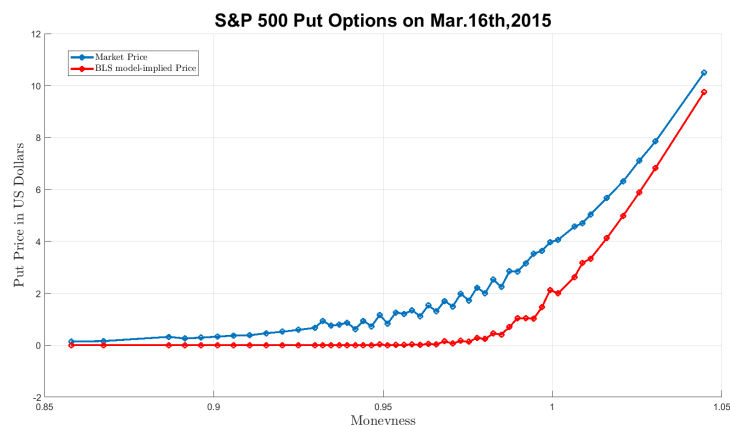
1. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Jan.26th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of the money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



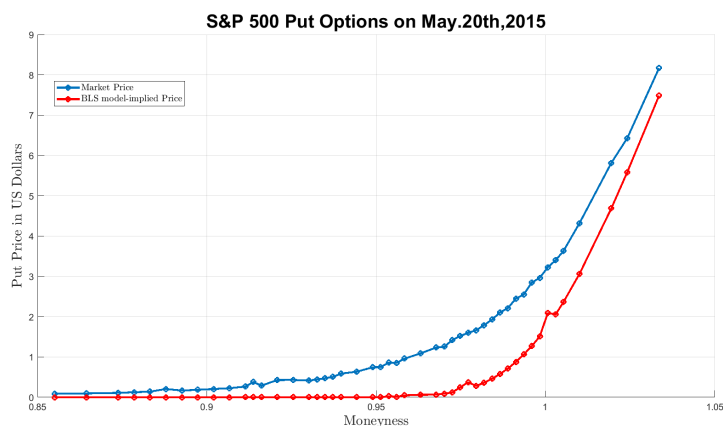
2. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Feb.4th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of the money. Also, when put options close to money, which means moneyness close to one, the BLS model-implied prices are close to market prices. These two characters are consistent with the previous findings.



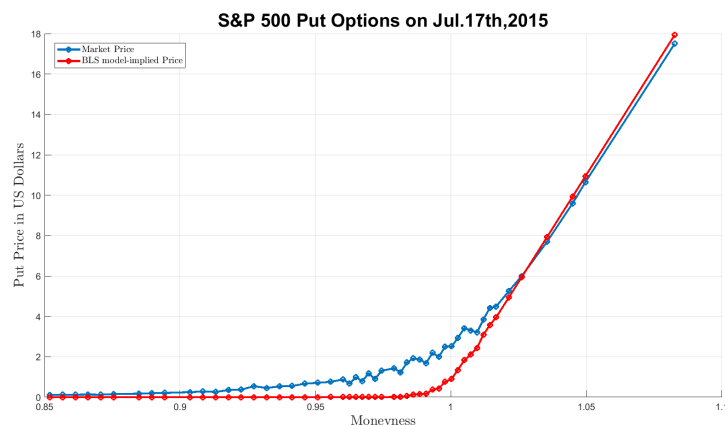
3. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Mar.16th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of the money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



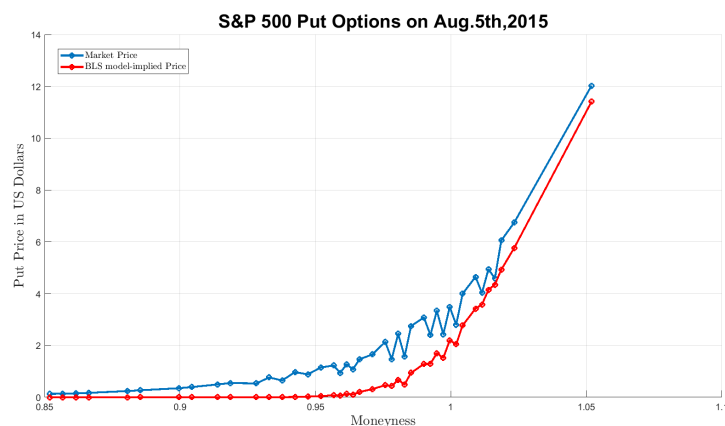
4. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on May.20th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are slightly lower than market prices when the put options are far out of money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



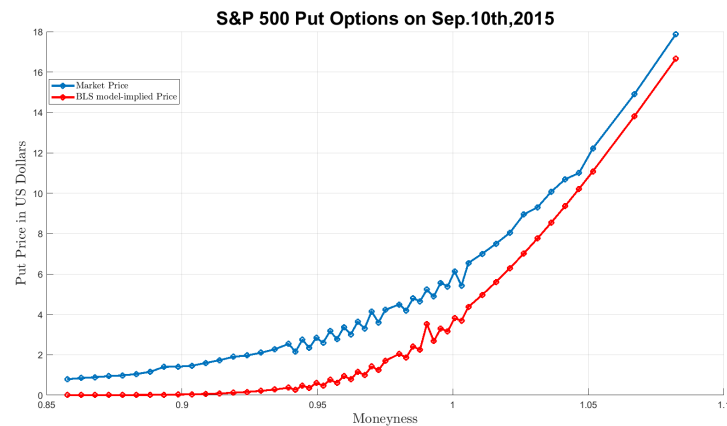
5. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Jul.17th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are slightly lower than market prices when the put options are far out of money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



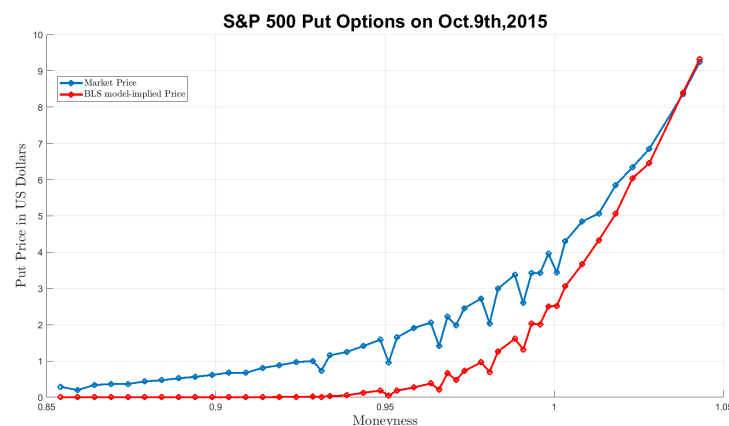
6. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Aug.5th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of money. Also, when put options close to money, which means moneyness close to one, the BLS model-implied prices are close to market prices. These two characters are consistent with the previous findings.



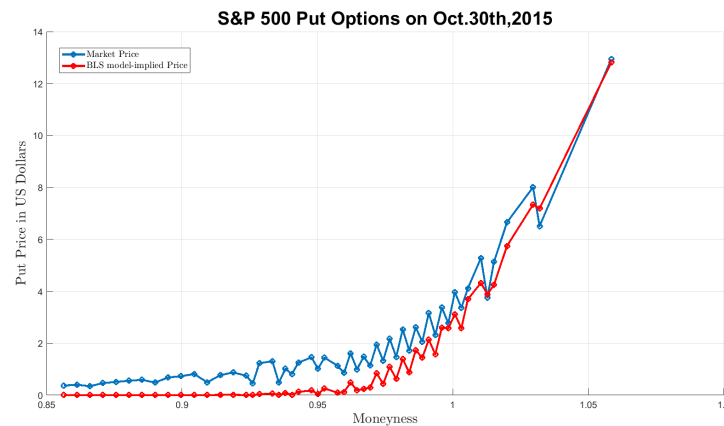
7. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Sep.10th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



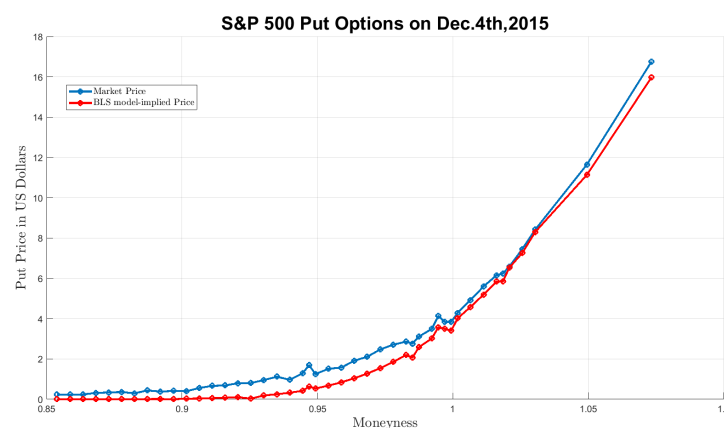
8. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Oct.9th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of money. Although when put options are close to money, which means moneyness close to one, the BLS model-implied prices are still lower than market prices, but this is also possible according to previous findings.



9. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyness on Oct.30th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are lower than market prices when the put options are far out of money. Also, when put options close to money, which means moneyness close to one, the BLS model-implied prices are close to market prices. These two characters are consistent with the previous findings.



10. This is the plot of market prices and BLS model-implied prices of SPY put options against their moneyiness on Dec.4th, 2015. The plot is consistent with previous finding. The BLS model-implied prices are slightly lower than market prices when the put options are far out of money. Also, when put options close to money, which means moneyiness close to one, the BLS model-implied prices are close to market prices. These two characters are consistent with the previous findings.



0.2.7 G

No. I do not think it is a good idea to suggest my boss writing a large amount of put options. Although according to BLS model, the put options are overvalued by market sometimes, the BLS model treats the volatility as constant, which may underestimate the probability of prices go down. For example, the volatility may change a lot before expiration and if the market price (SPY stock price) drops a lot, far below the strike price, we are gonna lose tons of money, much more than the money we lose if we input same amount of 10 million dollars into SPY stock market. Thus, I think writing 10 million dollars put options is quite risky. We can at most make money by collecting the premium but we have to burden so much risk. So I won't feel comfortable to talk to boss to write 10 million out of money put options if I only find the BLS model prices are lower than the market put prices. However, if I also has highly confidence in future market prices rising or has some hedging strategies, writing put options might be profitable.

0.3 Exercise 3

0.3.1 A

This is the table of the BLS implied prices and Hull-White implied prices of these put options with different strike prices.

K(USD)	Black-Scholes(USD)	Hull-White(USD)
42.5	1.0859	2.0181
44	1.4884	2.0716
48	3.0046	3.1192
50	4.0111	3.9941
52	5.1762	4.8955
55	7.1936	7.3048

Table 3: Put Price implied by Two Models

0.3.2 B

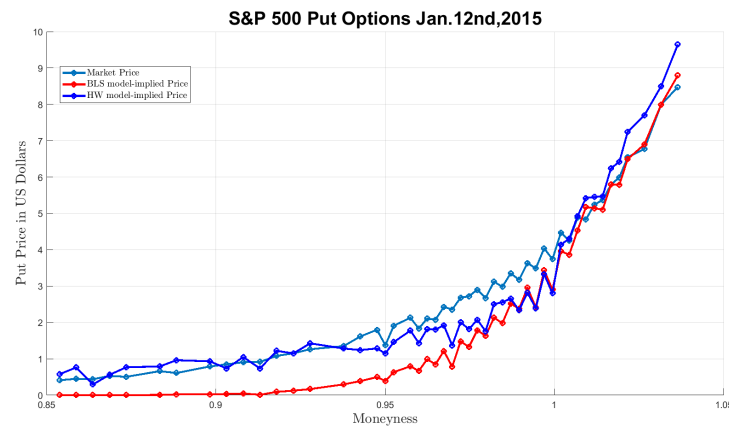
This is the plot of the prices implied by both models against the moneyness. HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are much higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



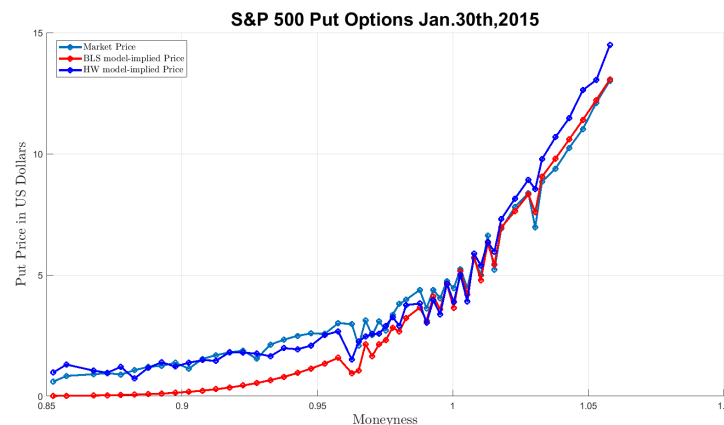
0.3.3 C

The following 10 files were chosen at random.

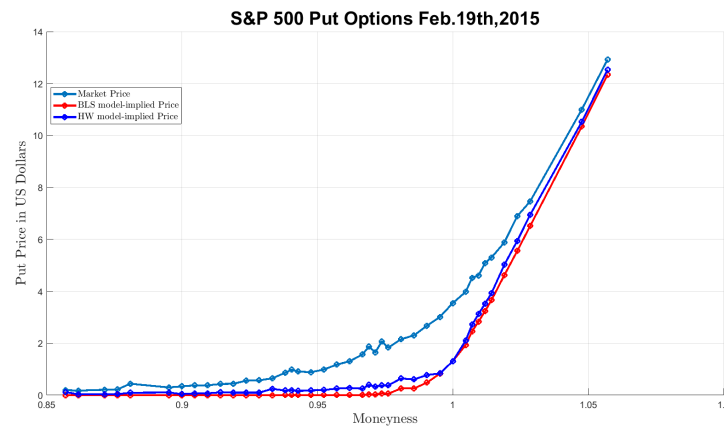
1. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Jan.12nd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices and get closer to actual market prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



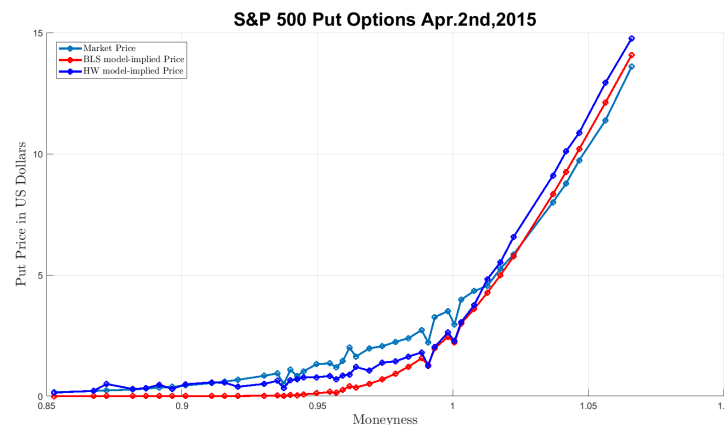
2. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Jan.30th, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices and get quite close to actual market prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



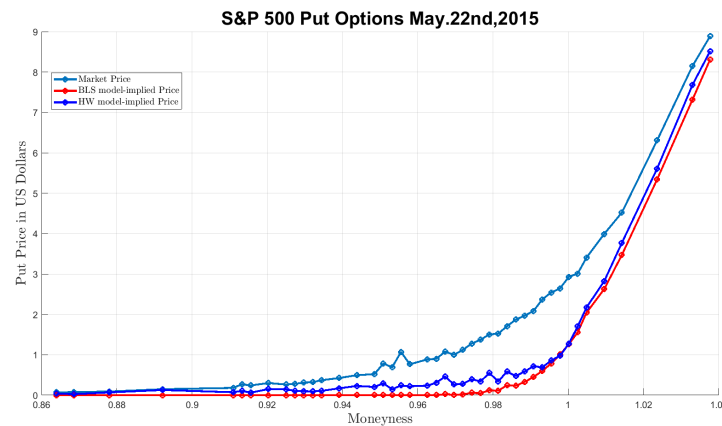
3. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Feb.19th, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



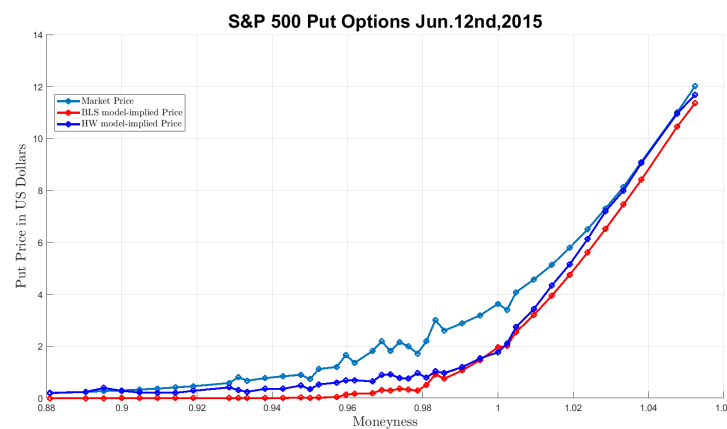
4. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Apr.2nd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices and get closer to actual market prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



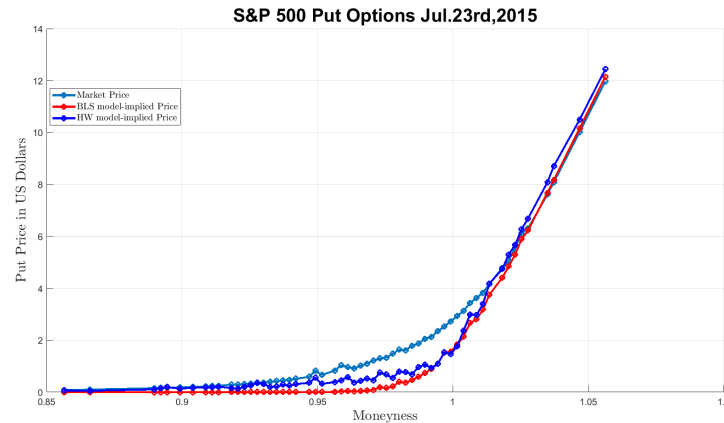
5. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on May.22nd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



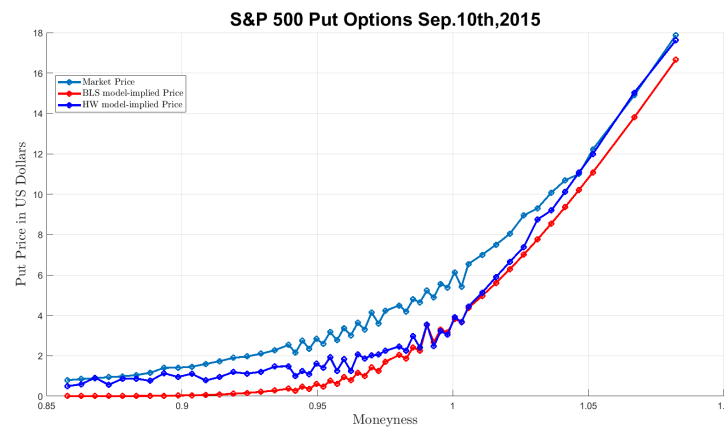
6. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Jun.12nd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



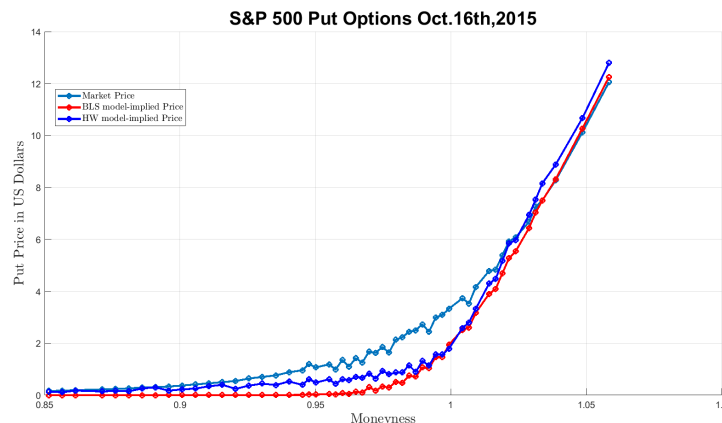
7. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Jul.23rd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



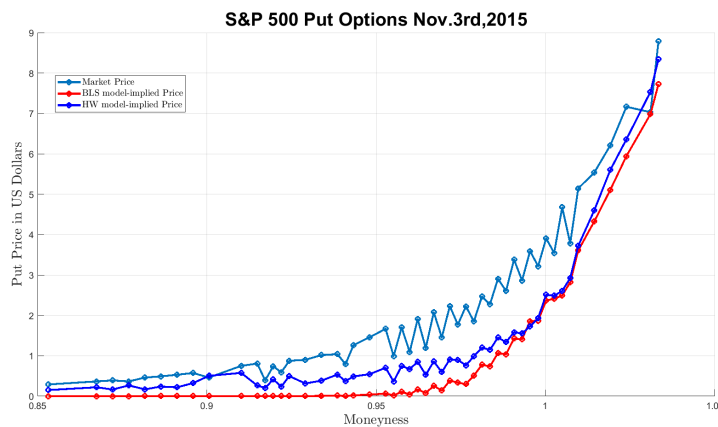
8. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Sep.10th, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



9. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Oct.16th, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



10. This is the plot of the market prices, BLS model-implied prices, HW model-implied prices of SPY put options against their moneyness on Nov.3rd, 2015. From the plot, we can see that HW put prices are higher than the BLS prices at most times but sometimes the BLS prices may be higher than HW prices when the put option is close to the money, or say the moneyness is close to 1. When the put options are far out of the money, HW put prices are higher than BLS prices, avoiding the underpricing. Thus, I think the random volatility, assumed in HW model, helps explain the underpricing of the BLS model.



0.4 Exercise 4

0.4.1 A

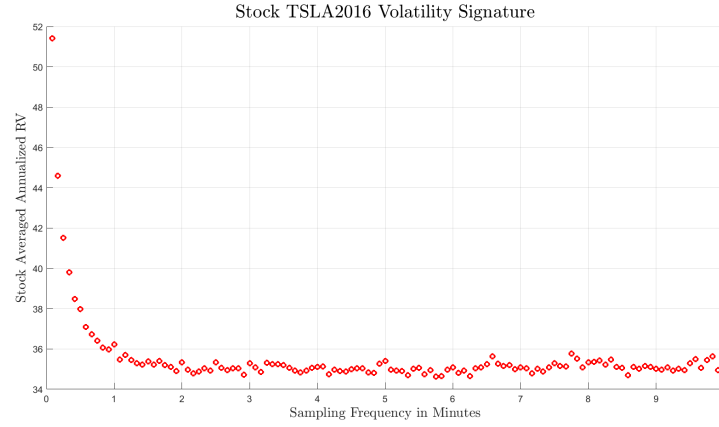
Inputting $k_n = 60$, we can get usual 5-min RV of Stock TSLA in 2016 and IBM in 2014. Inputting different k_n , we can get different coarse sampling RV.

0.4.2 B

1. This is the Volatility signature plot for TSLA-2016. When the sampling frequency is higher than 1 minute ($k_n \leq 12$), the averaged annualized RV goes higher as sampling frequency goes higher or say k_n goes lower. The averaged annualized RV is stable after k_n is bigger than 12. The reason is that the ultra high frequency data, like 5-second data, is contaminated by noise. According to the following equation, average RV will explode as $n(k_n)$ goes very big, or say as k_n goes very small since $n(k_n) = \lceil n_{high}/k_n \rceil$ (for 5-sec data

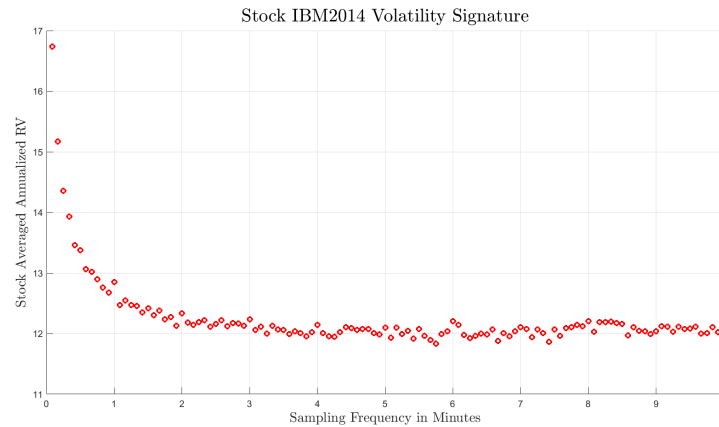
$n_{high} = 4620$).

$$\sum_{i=1}^{n(k_n)} (Y_i^{n(k_n)} - Y_{i-1}^{n(k_n)})^2 \approx \sum_{i=1}^{n(k_n)} (\Delta_i^{n(k_n)} X)^2 + \sum_{i=1}^{n(k_n)} (\chi_i^2 + \chi_{i-1}^2) = IV + 2n(k_n)\sigma_\chi^2.$$



2. This is the Volatility signature plot for IBM-2014. When the sampling frequency is higher than 2 minute ($k_n \leq 24$), the averaged annualized RV goes higher as sampling frequency goes higher or say k_n goes lower. The averaged annualized RV is stable after k_n is bigger than 24. The reason is that the ultra high frequency data, like 5-second data, is contaminated by noise. According to the following equation, average RV will explode as $n(k_n)$ goes very big, or say as k_n goes very small since $n(k_n) = [n_{high}/k_n]$ (for 5-sec data $n_{high} = 4620$).

$$\sum_{i=1}^{n(k_n)} (Y_i^{n(k_n)} - Y_{i-1}^{n(k_n)})^2 \approx \sum_{i=1}^{n(k_n)} (\Delta_i^{n(k_n)} X)^2 + \sum_{i=1}^{n(k_n)} (\chi_i^2 + \chi_{i-1}^2) = IV + 2n(k_n)\sigma_\chi^2.$$



0.4.3 C

The average $\hat{\sigma}_\chi^2$ of TSLA in 2016 is $1.1349 * 10^{-7}$ and the average $\hat{\sigma}_\chi^2$ of IBM in 2014 is $1.2029 * 10^{-8}$. As n goes really high, the daily RV we calculated will mainly be noises. As the following equation suggests, we can use the highest frequency data to get daily RV, which is contaminated by noise, and divide it by $2n_{high}$

to get the variance of daily noise.

$$\sum_{i=1}^n (Y_i^n - Y_{i-1}^n)^2 \approx \sum_{i=1}^n (\Delta_i^n X)^2 + \sum_{i=1}^n (\chi_i^2 + \chi_{i-1}^2) = IV + 2n\sigma_\chi^2 \approx 2n\sigma_\chi^2.$$

$$\hat{\sigma}_\chi^2 = \frac{1}{2n} \sum_{i=1}^n (Y_i^n - Y_{i-1}^n)^2.$$

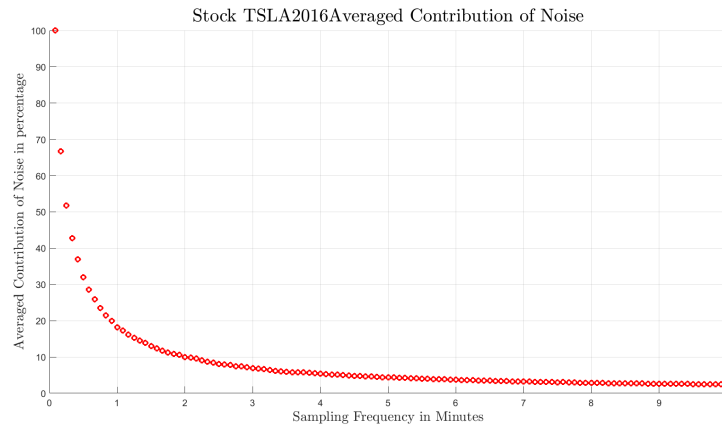
0.4.4 D

The following equations give the logic. The first equation describes the definition. The second equation is just putting k_n to the denominator. The third equation holds because RVs sampled at the same frequency of a particular day are similar, so $k_n RV_0(k_n)$ is an approximation of $\sum_{i=0}^{k_n-1} RV_i(k_n)$. The last equation holds because $\sum_{i=0}^{k_n-1} RV_i(k_n)$ actually contains all high frequency prices, so it is an approximation of $\sum_{i=1}^n (Y_i^n - Y_{i-1}^n)^2$, which can be divided into IV and Noise. Thus, the *contribution*(k_n) is a measure of $\frac{Noise}{IV+Noise}$ at frequency k_n .

$$contribution(k_n) \equiv \frac{2 \frac{n}{k_n} \sigma_\chi^2}{RV_0(k_n)} = \frac{2n\sigma_\chi^2}{k_n RV_0(k_n)} \approx \frac{2n\sigma_\chi^2}{\sum_{i=0}^{k_n-1} RV_i(k_n)} \approx \frac{Noise}{IV + Noise}.$$

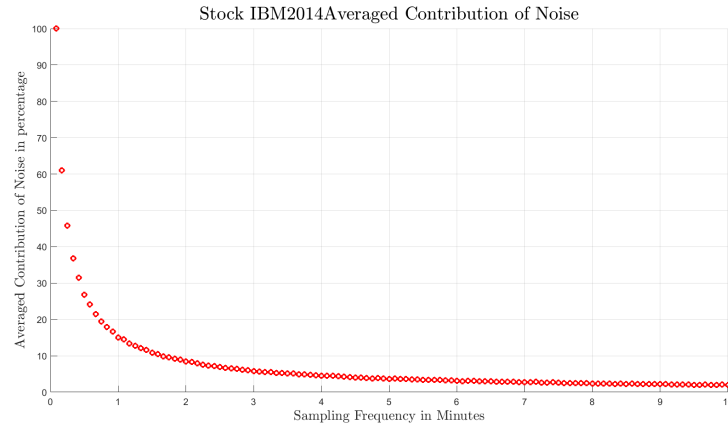
0.4.5 E

1. This is plot of the average contribution of the Noise to the Total RV against the frequency k_n of TSLA in 2016.



The noises do dominate RV at the very high frequencies since the average contribution of the noise are pretty high when k_n is small, or say at 5-sec and 10-sec. The noise is unimportant at low frequency since it contributes less than 10 percent of the total variation at 5-min and 8-min.

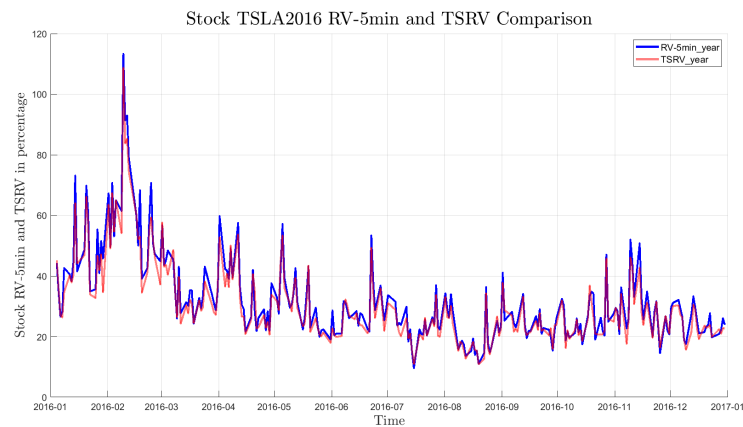
2. This is plot of the average contribution of the Noise to the Total RV against the frequency k_n of IBM in 2014.



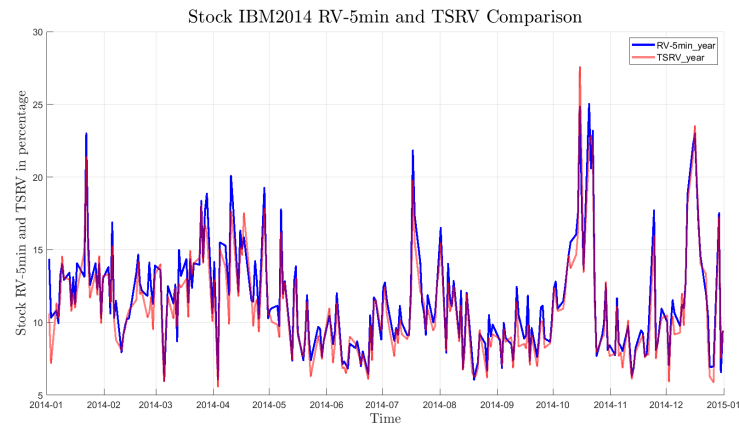
The noises do dominate RV at the very high frequencies since the average contribution of the noise are pretty high when k_n is small, or say at 5-sec and 10-sec. The noise is unimportant at low frequency since it contributes far less than 10 percent of the total variation at 5-min and 8-min.

0.4.6 G

1. This is plot of the TSRV based on $k_n = 60$ and RV based on 5-min coarse sampling of TSLA in 2016. We notice that TSRV and RV are quite closed to each other.

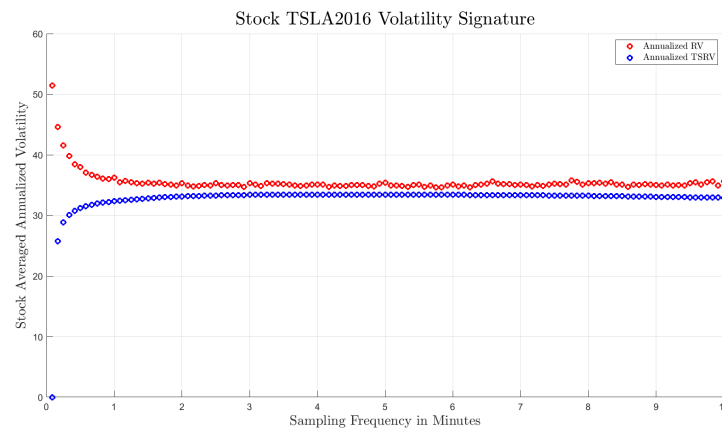


2. This is plot of the TSRV based on $k_n = 60$ and RV based on 5-min coarse sampling of IBM in 2014. We notice that TSRV and RV are quite closed to each other.



0.4.7 H

1. This is the Volatility signature plot including average RV and average TSRV of TSLA in 2016. The TSRV deals with the microstructure noise successfully since the curve is relatively flat even when the sampling frequency increases. The TSRV is smaller than usual RV estimator since TSRV subtracted the noises and TSRVs are more stable than usual RV estimator at different sampling frequency since TSRVs incorporated and averaged subsampled RVs by using high frequency data.



2. This is the Volatility signature plot including average RV and average TSRV of IBM in 2014. The TSRV deals with the microstructure noise successfully since the curve is relatively flat even when the sampling frequency increases. The TSRV is smaller than usual RV estimator since TSRV subtracted the noises and TSRVs are more stable than usual RV estimator at different sampling frequency since TSRVs incorporated and averaged subsampled RVs by using high frequency data.

