Samuelson Effect in Commodity Futures Volatilities

Yuxiao Cheng

August 2020

Abstract

We investigate different forms of Samuelson parameterizations. We start with the simple exponential decay and the optimal Samuelson decay parameter B. Then we generalize our results to other forms of Samuelson. We use data of implied volatilities, futures data and nearby contracts data. We introduce a long term volatility term, and later a long term decay term into the original Samuelson model. It turns out that the model with long term volatility term using futures data better and is more stable. We show the application of Samuelson parameterization to the problem of swaption evaluation.

Contents

1 Introduction

Empirical analysis of price returns is an essential component in the valuation methods in any asset classes. Hedging strategies, construction of investment portfolio and models selection for derivatives valuation rely on statistical estimates of moments and models of return processes. Energy commodities present some unique challenges: seasonality and inventory are crucial for covariance structure; forward price volatility increases dramatically while approaching contract's expiration, the famous Samuelson effect; returns are highly non-normal. Trading activity in correlation products is minimal.

Moreover, commodity futures and options liquidity is concentrated at short tenors, options with strikes near the money. This fact makes the term structure of volatility and correlations very important in hedging decisions. Those structures should be also taken into account while valuing early expiration options, for example swaptions, which are quite popular and widely traded in Energy Markets.

The subject of study of this project will be commodity volatilities and correlations. The extraordinary Spring 2020 makes the research even more challenging and exciting. Inevitably, a new reality has affected volatilities and correlations. We would like to understand and analyze how.

1.1 Literature Review

We use the well-known books on Commodities Derivatives, wrote by A. Eydeland, "Energy and Power Risk Management" and other. The book says that the crude oil market is the largest commodity market in the world. The most significant trading hubs are New York, London, and Singapore. These markets trade crude oil as well as refined products such as gasoline and heating oil.And we will be focusing on such kind of important commodity: crude oil in our paper.

The book also introduces Samuelson effect for us: The volatility of any given forward contract is generally a decreasing function of the time to expiration. This property of the volatility structure is called the Samuelson effect.

We also study the calibration method from paper by R. Galeeva and Th. Haversang "Parameterized Calendar Correlations: Decoding Oil and Beyond" published in JOD, Spring 2020 edition which calibrates the implied oil calendar correlations to the model of growth and decay.

The authors suggests parametric families to model calendar correlations. They capture the empirical properties of historical realized calendar correlations: the growth of correlations with time to the expiration of earlier contract, and decay with time between two contracts. And the growth of correlations represents the "Samuelson effect" for commodity futures and is captured through dynamics of instantaneous correlations, which is similar to the object of our paper.

1.2 Goal of Project

In this project, we will concentrate on the following objectives:

1. Investigate different forms of Samuelson parameterization, start with the simple exponential decay. Analyze different forms, develop calibration procedures, run on futures data and nearby contracts data for crude oil, and market implied volatilities, develop extrapolation procedures. Analyze results on different time periods, including Spring 2020.

2. The goal is to describe methodologies to parameterize the well known Samuelson effect of increase of volatilities, both implied volatility and historical volatility of futures data and nearby contracts data, as contracts get closer to their maturity. This is a well understood and documented phenomenon across many physical commodities.

2 Theory

2.1 Notation explanation

- t: "Today"
- T_i: Expiration date of the WTI future contract i
- τ_i: Expiration date of an option on future i
- $\sigma_{0,i}$: Normalization constant
- B: Samuelson parameter
- $\sigma_i(t, T_i)$: Instantaneous volatility of the future contract i with expiration T_i at current time t
- $\sigma_{i,implied}$: Implied volatility of the future contract i with expiry at T_i at current time t
- $\sigma_{i,his}$: Historical volatility of the future contract i
- $\sigma_{i,nearby}$: Instantaneous volatility of the nearby future contract i
- σ_∞: long term volatility
- b: long term decay
- C_∞: long term volatility constant
- $Var([t, \tau_i])$: Variance for the future contract i with expiration T_i, realized between t and τ_i

2.2 Assumptions

- $\sigma_{i,implied}$ is the at-the-money (ATM) implied volatility.
- The total realized variance $Var([t, \tau_i])$ should match the total implied variance from the market $\sigma_{i,implied}(\tau_i - t)$, total variance:$Var([t, \tau_i]) = \sigma^2_{i,implied} \times (\tau_i - t)$
- The normalization constant $\sigma_{0,i}$ is the same for all contracts.

$$\sigma_{i,0} = \sigma_0$$

2.3 Samuelson Two Decays Volatility Model

This is the most general model. B is Samuelson parameter,C_∞ is long term volatility, b is long term decay.

$$\sigma_i(t,T_i) = \sigma_{i,0}(e^{-B(T_i-t)} + C_\infty e^{-b(T_i-t)}) \tag{1}$$

We will call this model: model 2

2.4 Samuelson Two Decays Variance Model

If we take square and remove the cross term, the Samuelson Two Decays Volatility Model becomes Samuelson Two Decays Variance Model.

$$\sigma_i^2(t,T_i) = \sigma_{i,0}^2(e^{-2B(T_i-t)} + C_\infty^2 e^{-2b(T_i-t)}) \tag{2}$$

And we will call this model: model 2a

2.5 Samuelson One Decay Volatility Model

For a particular cases, we assume long term volatility is a constant,We let b=0, and let C_∞=σ_∞, the Samuelson Two Decays Volatility Model becomes Samuelson One Decay Volatility Model

$$\sigma_i(t,T_i) = \sigma_{i,0}(e^{-B(T_i-t)} + \sigma_\infty) \tag{3}$$

And we will call this model:model 1

2.6 Samuelson One Decay Variance Model

If we take square and remove the cross term, the Samuelson One Decay Volatility Model becomes Samuelson One Decay Variance Model.

$$\sigma_i^2(t,T_i) = \sigma_{0,i}^2(e^{-2B(T_i-t)} + \sigma_\infty^2) \tag{4}$$

And we will call this model:model 1a

2.7 Samuelson Exponential Decay Model

The simplest case, we assume the long term volatility is zero, so we remove the σ_∞ term and the model becomes the simplest model:model 0. It becomes to parameterize the instantaneous volatility $\sigma_i(t,T_i)$ of a future contract i with expiration T_i at current time t by one exponential decay form:

$$\sigma_i(t,T_i) = \sigma_{0,i}e^{-B(T_i-t)} \tag{5}$$

And we will call this model:model 0

And in this report I will focus on Samuelson Two Decays Variance Model, Samuelson One Decay Variance Model and Samuelson Exponential Decay Model, which are model 2a, model 1a and model 0.

3 Calibration

3.1 Data pre-processing

We get futures information from Barchart.

3.2 Implied volatility

We get futures contracts implied volatilities from Barchart.

As mentioned in our assumption, the realized variance from "today" to options expiration should match the total implied variance from the market:

$$\sigma^2_{i,implied} \times (\tau_i - t) = Var([t, \tau_i]) \tag{6}$$

And we can get the realized variance by integrated the instantaneous volatility from "today" to the option expiration:

$$Var([t, \tau_i]) = \int_t^{\tau_i} \sigma_i^2(s, T_i) ds \tag{7}$$

Then we can calibrate this simple parameterization and try to match their structure by finding the best parameter B. Since we assume $\sigma_{0,i}$ is the same for all contracts, we could form ratios w.r.t the nearest future contract as:

$$\frac{\sigma^2_{k,implied} \times (\tau_k - t)}{\sigma^2_{1,implied} \times (\tau_1 - t)}, \quad k = 2, 3, \ldots$$

where $\sigma_{1,implied}$ is the implied ATM volatility of the nearest future contract, $\sigma_{k,implied}$ is the implied ATM volatility of the k-nearest future contract.

By taking ratios, the unknown $\sigma_{0,i}$ will be eliminated, and the only unknown parameter is a constant B.

In model 0, the realized variance can be calculated as

$$Var([t, \tau_i]) = \int_t^{\tau_i} \sigma^2_{0,i} e^{-2B(T_i - s)} ds = \frac{\sigma^2_{0,i}}{2B}(e^{-2B(T_i - \tau_i)} - e^{-2B(T_i - t)})$$

The ratios will be

$$\frac{\sigma^2_{k,implied} \times (\tau_k - t)}{\sigma^2_{1,implied} \times (\tau_1 - t)} = \frac{e^{-2B(T_k - \tau_k)} - e^{-2B(T_k - t)}}{e^{-2B(T_1 - \tau_1)} - e^{-2B(T_1 - t)}}, \quad k = 2, 3, \ldots$$

We call the left part as the realized ratios, and right part as the calculated ratios, since the left is got from market information, the right is calculated based on models.

Then we try to find the best value of B to minimum the difference of those two ratios. Root mean square error (RMSE) is used to measure the difference. By this way we can get the value of Samuelson parameter B.

The RMSE will be

$$\sqrt{\frac{\sum_{k=2}^{n}(\sigma_{k,implied} - \sigma_{1,implied} \times \frac{(\tau_1 - t)}{(\tau_k - t)} \times \sqrt{\frac{e^{-2B(T_k - \tau_k)} - e^{-2B(T_k - t)}}{e^{-2B(T_1 - \tau_1)} - e^{-2B(T_1 - t)}}})^2}{n}}, \quad k = 2, 3, \ldots$$

3.3 Historical volatility of futures data

Since the result we get from implied volatility is not stable, we move to historical volatility.

Crude Oil WTI Futures historical price data are got from Barchart. "Moving Windows" is used to get Samuelson parameters. For a specific window, $[t_1, t_2]$, we can get daily return for each contract in that period. In our case, we use 60 contracts in each window, for example, from January 31, 2018 to January 31, 2019, we use futures contracts, the expirations of which are from March 20, 2019 to February 20, 2024. And the daily return will be

$$r_n = \log(\frac{F_n/F_{n-1}}{\sqrt{t_n - t_{n-1}}})$$

where F_n is the price of the future on t_n, F_{n-1} is the price on t_{n-1}. Then we can get the historical volatilities (daily) σ_i for each contract in that window by taking the standard deviation on daily returns.

And the historical variance in $[t_1, t_2]$ can be calculated as $\sigma_i^2 \times (t_2 - t_1)$ Since our data range of historical data is 1 year. So we will fix $(t_2 - t_1) = 1$ and use 1 directly in our following fomula.

3.3.1 Model 0

In model 0, the calculated variance of the future contract i in that window,

$$Var_i([t_1, t_2]) = \int_{t_1}^{t_2} \sigma_{0,i}^2 e^{-2B(T_i - s)} ds = \frac{\sigma_{0,i}^2}{2B}(e^{-2B(T_i - t_2)} - e^{-2B(T_i - t_1)})$$

Then we get the ratios,

$$\frac{\sigma_k^2}{\sigma_1^2} = \frac{e^{-2B(T_k - t_2)} - e^{-2B(T_k - t_1)}}{e^{-2B(T_1 - t_2)} - e^{-2B(T_1 - t_1)}}, \quad k = 2, 3, \ldots$$

By calculated the RMSE of realized ratios and calculated ratios, we can get the optimal value of Samuelson parameter B and corresponding $\sigma_{0,i}$.

3.3.2 Model 1a

In model 1a, after adding the long term volatility term, the calculated variance becomes

$$\begin{aligned} Var_i([t_1, t_2]) &= \int_{t_1}^{t_2} \sigma_{0,i}^2 (e^{-2B(T_i - s)} + \sigma_\infty^2) ds \\ &= \frac{\sigma_{0,i}^2}{2B}(e^{-2B(T_i - t_2)} - e^{-2B(T_i - t_1)}) \\ &\quad + \sigma_{0,i}^2 \sigma_\infty^2 (t_2 - t_1) \end{aligned}$$

Then ratios will be

$$\frac{\sigma_k^2}{\sigma_1^2} = \frac{p(T_k)}{p(T_1)}, \quad k = 2, 3, \ldots$$

where

$$p(T_i) = (e^{-2B(T_i-t_2)} - e^{-2B(T_i-t_1)}) + 2B\sigma_\infty^2(t_2 - t_1), i = 1, 2, 3, \ldots$$

3.3.3 Model 2a

In model 2a, we add another parameter b, the calculated variance is

$$\begin{aligned} Var_i([t_1, t_2]) &= \int_{t_1}^{t_2} \sigma_{0,i}^2 [e^{-2B(T_i-s)} + C_\infty^2 e^{-2b(T_i-s)}] ds \\ &= \frac{\sigma_{0,i}^2}{2B}(e^{-2B(T_i-t_2)} - e^{-2B(T_i-t_1)}) \\ &\quad + \frac{\sigma_{0,i}^2 C_\infty^2}{2b}(e^{-2b(T_i-t_2)} - e^{-2b(T_i-t_1)}) \end{aligned}$$

Then ratios will be

$$\frac{\sigma_k^2}{\sigma_1^2} = \frac{f(T_k)}{f(T_1)}, \quad k = 2, 3, \ldots$$

where

$$\begin{aligned} f(T_i) = b(&e^{-2B(T_i-t_2)} - e^{-2B(T_i-t_1)}) \\ &+ C_\infty^2 B(e^{-2b(T_i-t_2)} - e^{-2b(T_i-t_1)}), i = 1, 2, 3, \ldots \end{aligned}$$

3.4 Historical volatility of nearby contract data

For a specific window $[t_j, t_{j+1}]$, we regard the standard deviation of the i^{th} nearby contract log-return in that period as the Historical volatility of nearby contract, $\sigma_i(t_j, T_i)$, T_i is the expiration date of the i^{th} nearby contract.

We also first get the daily log-return of the i^{th} nearby contract just like what we did in historical volatility by

$$r_n = \frac{\log(F_n/F_{n-1})}{\sqrt{t_n - t_{n-1}}}$$

where F_n is the price of the contract on t_n, F_{n-1} is the price on t_{n-1}.

Especially on switch dates, for example, when the $i+1^{th}$ nearby contract becomes the i^{th} nearby contract, we take the log-return w.r.t to the price of the previous day of the $i+1^{th}$ nearby contract, instead of i^{th} nearby contract, which means we never take log-returns on different contracts.

Given t, we set the window as $[t, t+30]$, and the expiration date of the first nearby contract as $T_0 = t+15$, $T_{i+1} - T_i = 30$. Then we can take the standard deviation of the i^{th} nearby contract log-return in that period as the Historical volatility of nearby contract data, and square it as the variance, $\sigma_i^2(t, T_i)$.

3.4.1 Model 0

In model 0, as mentioned before, we can get the realized variance for each nearby contract on a specific window $[t, t+30]$, $\sigma_i^2(t, T_i)$.

And the calculated variance of the future contract i in that window, $\sigma_{0,i}^2 e^{-2B(T_i - t)}$

Then we get the ratios,

$$\frac{\sigma_k^2}{\sigma_1^2} = \frac{e^{-2B(T_k - t)}}{e^{-2B(T_1 - t)}}, \quad k = 2, 3, \ldots$$

By calculated the RMSE of realized ratios and calculated ratios, we can get the optimal value of Samuelson parameter B and corresponding $\sigma_{0,i}^2 = \frac{\sigma_k^2}{e^{-2B(T_k - t)}}$.

3.4.2 Model 1a

We add long term volatility term σ_∞ into model 0, and the calculated variance becomes $\sigma_{0,i}^2(e^{-2B(T_i - t)} + \sigma_\infty^2)$

Then ratios will be

$$\frac{\sigma_k^2}{\sigma_1^2} = \frac{e^{-2B(T_i - t)} + \sigma_\infty^2}{e^{-2B(T_1 - t)} + \sigma_\infty^2}, k = 2, 3, \ldots$$

Then we do the calibration for both B and σ_∞ to get the optimal value of Samuelson parameter B, long term volatility term σ_∞

4 Results

4.1 Implied volatility

In model 0, we can find Samuelson parameter B is not stable, as is seen in Table 1 below.The date in the table are the expiration date of future contract. We use these dates to obverse the implied volatility.And all errors we listed in the tables are for fitted volatility. They are the RMSE of fitted volatility

01/22/2019 0.4447	02/20/2019 0.1602	03/20/2019 0.1172	04/18/2019 -0.0191	05/21/2019 -0.0272
06/20/2019 0.6338	07/22/2019 0.2412	08/20/2019 0.2963	09/20/2019 0.5037	10/22/2019 0.2622
11/20/2019 0.2575	12/19/2019 -0.0592	01/21/2020 0.2303	02/20/2020 0.6063	03/20/2020 8.9046
04/21/2020 1.4682	05/19/2020 1.8338			

Table 1: model 0 Samuelson parameter B

Also, RMSEs in Table 2 show that the fitting results are not good.

01/22/2019 0.0767	02/20/2019 0.0362	03/20/2019 0.0288	04/18/2019 0.0174	05/21/2019 0.0207
06/20/2019 0.0977	07/22/2019 0.0479	08/20/2019 0.0544	09/20/2019 0.0771	10/22/2019 0.0497
11/20/2019 0.0431	12/19/2019 0.0328	01/21/2020 0.0395	02/20/2020 0.0796	03/20/2020 0.2899
04/21/2020 0.1571	05/19/2020 0.1902			

Table 2: model 0 RMSE

As Figure 1 is shown, realized ratios change a lot and cannot be explained by one model. That is the reason why we use historical volatility later

Figure 1: model 0 Realized Ratios

4.2 Historical volatility

4.2.1 Model 0

In model 0, the Samuelson parameter B results of historical volatility become more stable than the results of implied volatility. And The date in the table are also the expiration date of future contract. We use these dates as the end of observation period. For example, in 04/21/2020, the t1 is 2019-02-28 and t2 is 2020-02-29 and the historical volatility is for this 1 year range.

03/20/2019	04/18/2019	05/21/2019	06/20/2019
0.1152	0.1176	0.1173	0.1181
07/22/2019	08/20/2019	09/20/2019	10/22/2019
0.1168	0.1297	0.1291	0.1359
11/20/2019	12/19/2019	01/21/2020	02/20/2020
0.1519	0.1557	0.1807	0.1820
03/20/2020	04/21/2020	05/19/2020	06/22/2020
0.1865	0.1901	0.4016	0.5852

Table 3: model 0 Samuelson parameter B

The values of RMSE also show that fitting results are improved after changing to historical volatility.

03/20/2019 0.0007	04/18/2019 0.0007	05/21/2019 0.0007	06/20/2019 0.0007
07/22/2019 0.0007	08/20/2019 0.0006	09/20/2019 0.0006	10/22/2019 0.0007
11/20/2019 0.0009	12/19/2019 0.0009	01/21/2020 0.0009	02/20/2020 0.0008
03/20/2020 0.0008	04/21/2020 0.0009	05/19/2020 0.0037	06/22/2020 0.0065

Table 4: model 0 RMSE

As Figure 2 and Figure 3 are shown, the ratios will converge to a constant, and model 0 cannot fit tails well. So we will add a long term volatility term in Samuelson One Decay Variance Model.

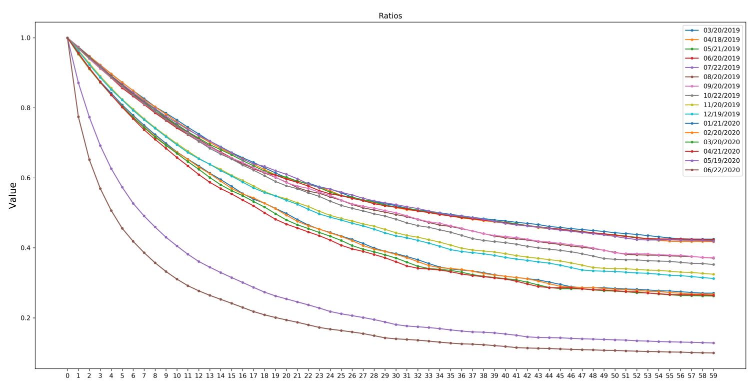

Figure 2: model 0 Realized Ratios

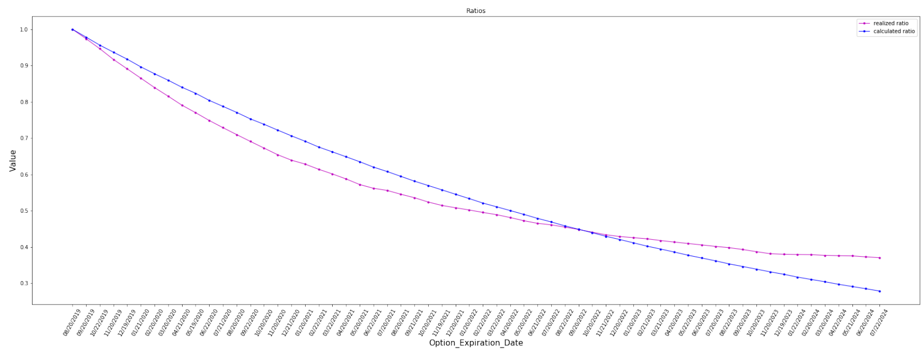

Figure 3: model 0 Realized and Calculated Ratios in Aug 2019

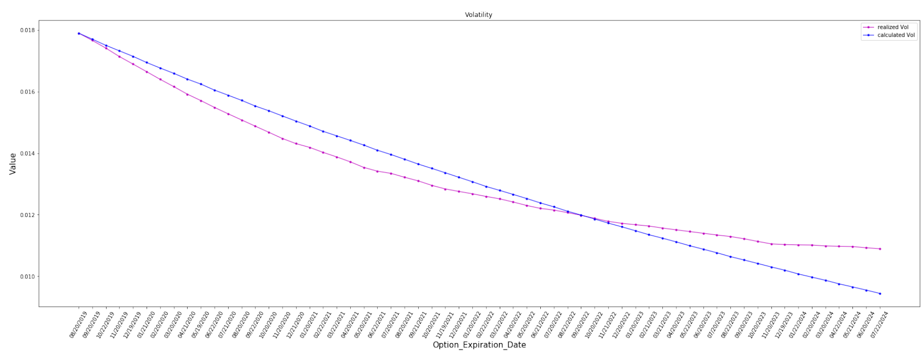

Figure 4: model 0 Realized and Calculated Volatility in Aug 2019

4.2.2 Model 1a

In model 1a, we first draw contour plot to get start values of Samuelson parameter B and long term volatility term σ_∞. Contour plots are a way to show a three-dimensional surface on a two-dimensional plane. It graphs two predictor variables X Y on the y-axis and a response variable Z as contours. And in our case, it graphs two predictor variables σ_∞ and B and a response variable:ratio:$\frac{\sigma_k^2}{\sigma_1^2}$ And from the contour plot, we can see that different lines intersect at point (0.3,0.65),which means the reasonable range of B is 0.2 to 0.4 and range of σ_∞ is 0.6 to 0.7. This can help us know the general range of the parameters.

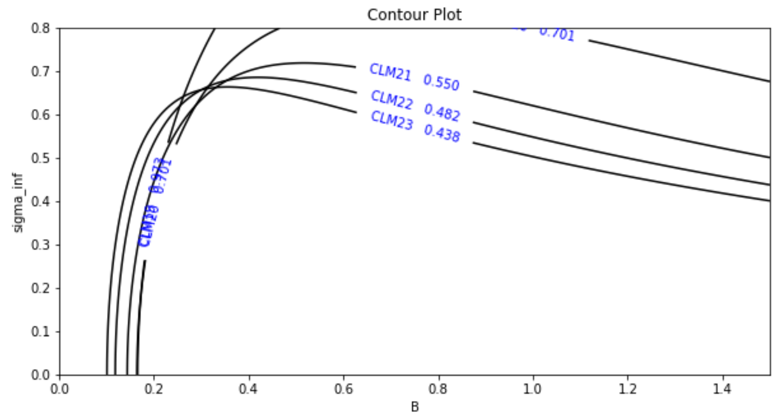

Figure 5: model 1a Contour plot in April 2019

After adding the long term volatility term, Samuelson parameter B becomes much more stable, also RMSE decreases a lot which means model 1a fits better than model 0. The results are shown in Table 5. Also figure 6 and figure 7 supports that model 1a fits the ratios well.

Contract	B	σ_∞	RMSE
03/20/2019	0.3081	0.6705	5.99E-05
04/18/2019	0.3069	0.6568	5.35E-05
05/21/2019	0.3181	0.6646	4.48E-05
06/20/2019	0.3339	0.6680	4.23E-05
07/22/2019	0.3028	0.6581	3.74E-05
08/20/2019	0.2795	0.5848	3.92E-05
09/20/2019	0.2815	0.5893	4.04E-05
10/22/2019	0.2749	0.5551	3.77E-05
11/20/2019	0.3270	0.5372	7.45E-05
12/19/2019	0.3208	0.5218	6.84E-05
01/21/2020	0.3294	0.4531	6.41E-05
02/20/2020	0.3310	0.4511	6.13E-05
03/20/2020	0.3364	0.4441	5.15E-05
04/21/2020	0.3579	0.4478	4.40E-05
05/19/2020	0.7088	0.2773	5.70E-04
06/22/2020	1.0379	0.2152	1.28E-03

Table 5: model 1a Results Summary

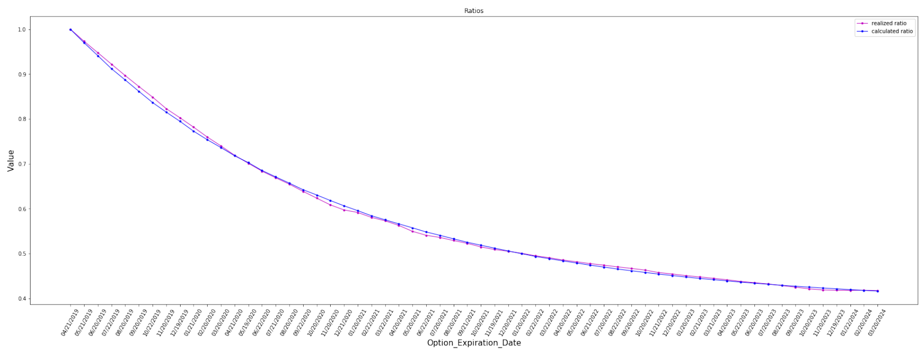

Figure 6: model 1a Realized and Calculated Ratios in April 2019

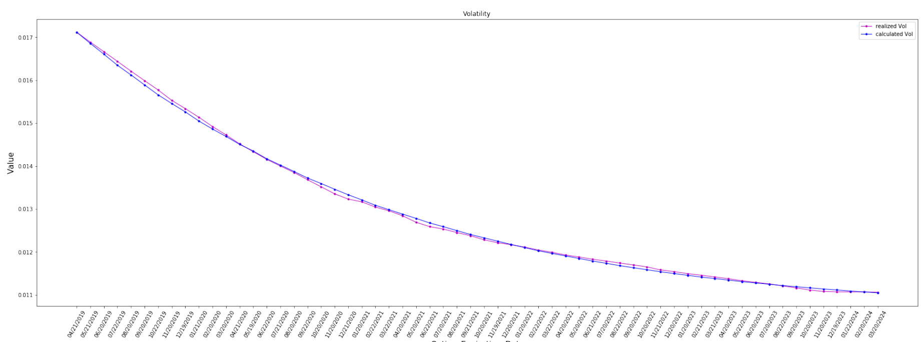

Figure 7: model 1a Realized and Calculated Volatility in April 2019

4.2.3 Model 2a

Although Samuelson parameter B in Samuelson One Decay Variance Model is more stable than model 0, the another parameter σ_∞ we introduced, which changes a lot. This time, we try model the long term decay like, so introduce one more parameter b. Before doing that, we first fix $B = 0.8$, based on what we got in model 1a. Then we draw the contour plots for C_∞ and b to get the start values, as is shown below. We can see the general range of C_∞ and b should be 0.6 to 0.8 and 0 to 0.02.

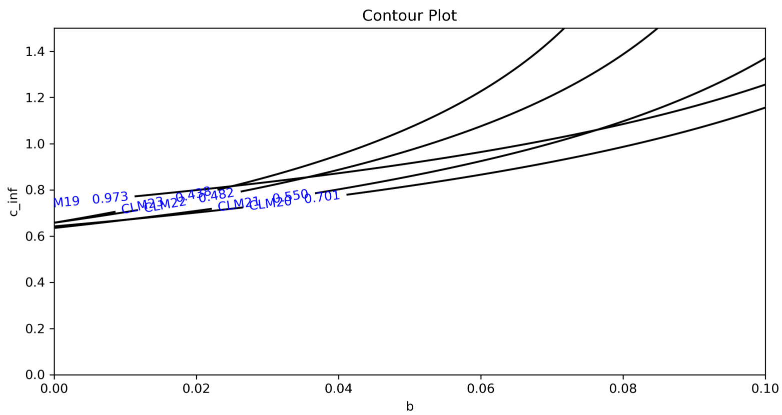

Figure 8: Model 2a Contour Plot in Apr 2019

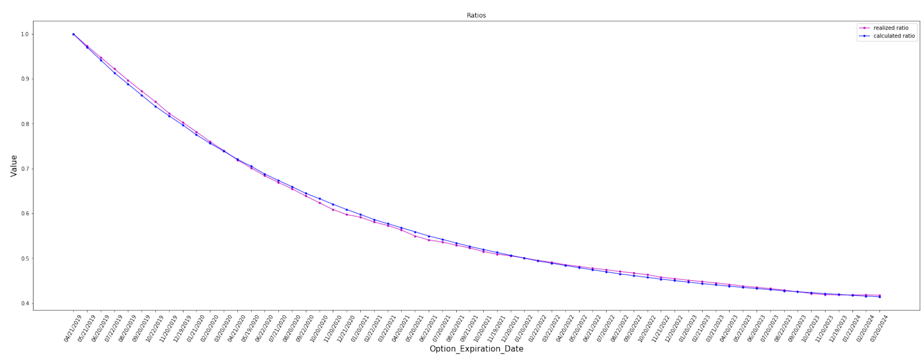

Figure 9: model 2a Realized and Calculated Ratios in April 2019

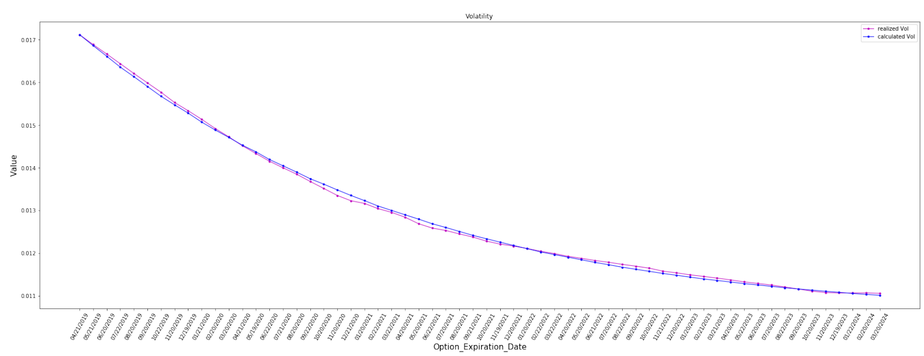

Figure 10: model 2a Realized and Calculated Volatility in April 2019

Then do the calibration, the results are shown in Table 6. We can find that b is so small that the results do not improved to much with a new parameter included.

Contract	b	σ_∞	RMSE
03/20/2019	9.26E-08	0.6658	6.63E-05
04/18/2019	3.36E-10	0.6528	5.83E-05
05/21/2019	6.43E-10	0.6548	6.66E-05
06/20/2019	6.99E-12	0.6512	9.71E-05
07/22/2019	1.84E-13	0.6564	3.79E-05
08/20/2019	1.22E-02	0.6513	4.07E-05
09/20/2019	1.15E-02	0.6512	3.67E-05
10/22/2019	1.62E-02	0.6395	4.62E-05
11/20/2019	2.33E-13	0.5194	1.17E-04
12/19/2019	5.72E-12	0.5071	9.61E-05
01/21/2020	9.01E-12	0.4299	1.42E-04
02/20/2020	5.63E-15	0.4266	1.33E-04
03/20/2020	9.33E-09	0.4152	1.56E-04
04/21/2020	3.02E-11	0.4063	2.38E-04
05/19/2020	7.49E-02	6.49E-06	2.82E-03
06/22/2020	1.31E-02	-1.03E-05	5.15E-03

Table 6: model 2a Results Summary

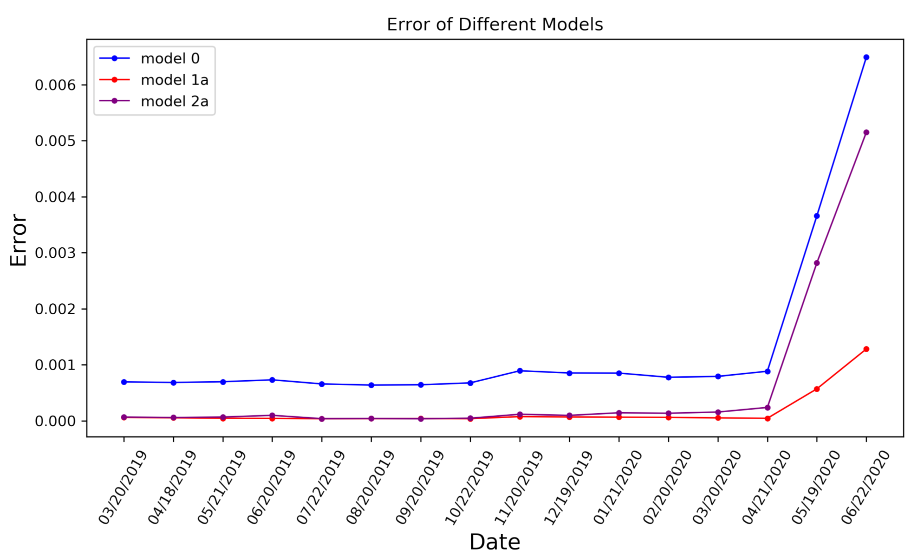

Figure 11: RMSE of three different models

From the results from plot 11. We can find that model 1a performs much better than other two models. Especially in the most volatile May and June, it can still control the RMSE, which means the introduction of variable long term volatility is reasonable and suitable.

4.3 Samuelson calibration using nearby contracts

4.3.1 Model 0

In model 0, the Samuelson parameter B is relatively stable except the last three contracts. And the date in this table is different. We just pick the day near expiration date on January 2019, and then use rolling window of 30 working days to decide our next date until June 2020.

01/17/2019 0.1678	03/04/2019 0.2087	04/15/2019 0.3946	05/29/2019 0.1449
07/11/2019 0.1855	08/22/2019 0.1628	10/04/2019 0.3965	11/15/2019 0.1254
12/31/2019 0.2631	02/13/2020 0.2847	03/27/2020 0.6900	05/11/2020 3.6294
06/23/2020 0.4794			

Table 7: model 0 Samuelson parameter B

The values of RMSE also show that fitting results are fine except the last three contracts.

01/17/2019 0.0006	03/04/2019 0.0010	04/15/2019 0.0015	05/29/2019 0.0002
07/11/2019 0.0003	08/22/2019 0.0008	10/04/2019 0.0027	11/15/2019 0.0004
12/31/2019 0.0011	02/13/2020 0.0013	03/27/2020 0.0116	05/11/2020 0.0243
06/23/2020 0.0049			

Table 8: model 0 RMSE

As Figure 12 and Figure 13 are shown, the ratios will converge to a constant, and model 0 cannot fit tails well. So we will add a long term volatility term in model 1a.

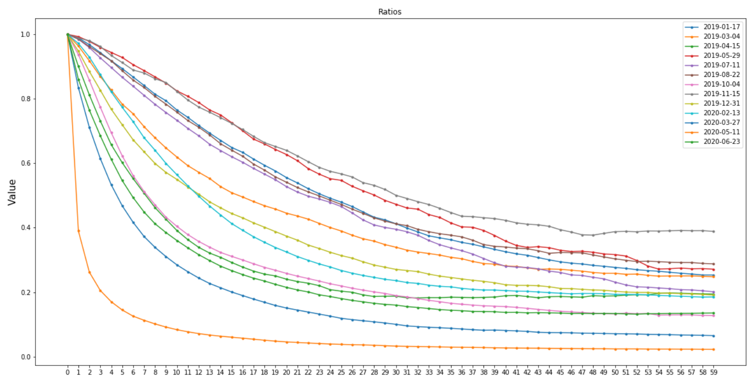

Figure 12: model 0 Realized Ratios

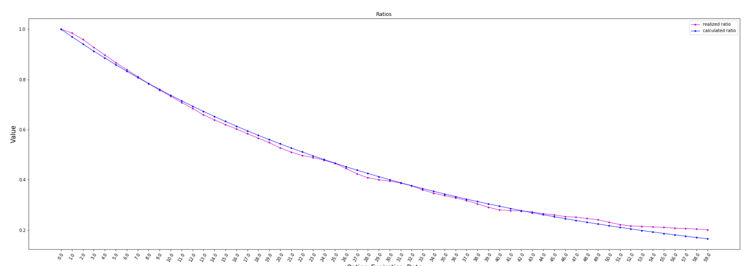

Figure 13: model 0 Realized and Calculated Ratios in Jul 2019

,

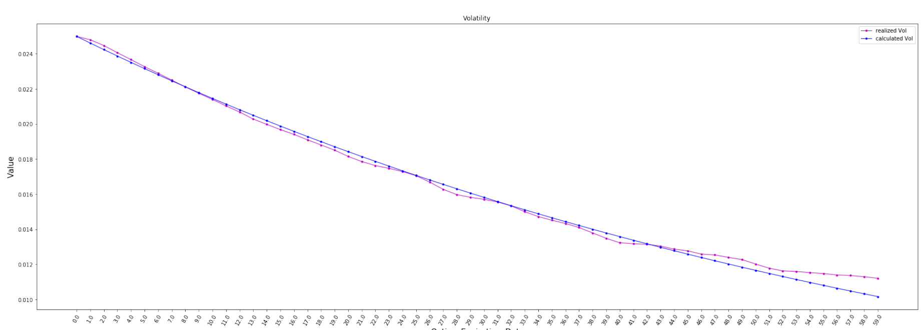

Figure 14: model 0 Realized and Calculated Volatility in Jul 2019

4.3.2 Model 1a

In model 1a, we first draw contour plot to get start values of Samuelson parameter B and long term volatility term σ_∞.

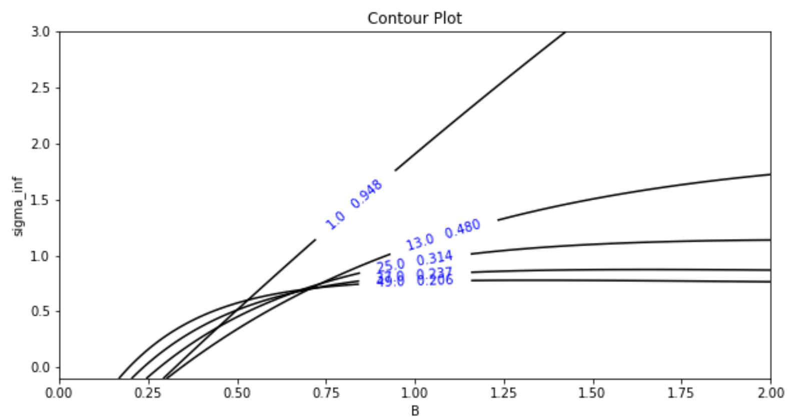

Figure 15: model 1a Contour plot in December 2019

After adding the long term volatility term, Samuelson parameter B does not become more stable, but RMSE decreases. The results are shown in Table 9. And figure 16 and figure 17 support that model 1a fits the ratios well.

Contract	B	σ_∞	RMSE
01/17/2019	0.2093	0.3524	3.36E-04
03/04/2019	0.3987	0.5381	7.33E-05
04/15/2019	0.8235	0.4578	7.68E-05
05/29/2019	0.1449	0.0000	2.20E-04
07/11/2019	0.2113	0.2635	1.71E-04
08/22/2019	0.2465	0.4823	2.96E-04
10/04/2019	0.6684	0.3971	3.78E-04
11/15/2019	0.1979	0.5462	2.08E-04
12/31/2019	0.4777	0.4772	1.05E-04
02/13/2019	0.4711	0.4369	2.17E-04
03/27/2020	0.9558	0.2938	1.92E-03
05/11/2020	4.4722	0.1778	7.27E-03
06/23/2020	0.8405	0.3916	4.22E-04

Table 9: model 1a Results Summary

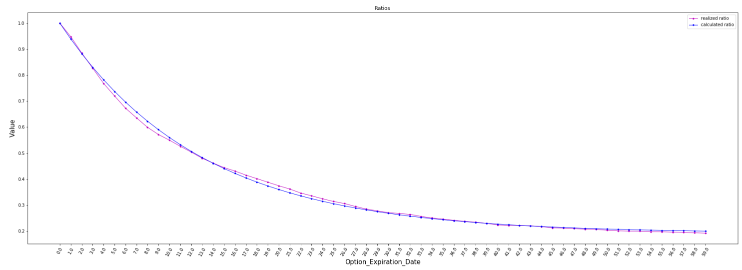

Figure 16: model 1a Realized and Calculated Ratios in December 2019

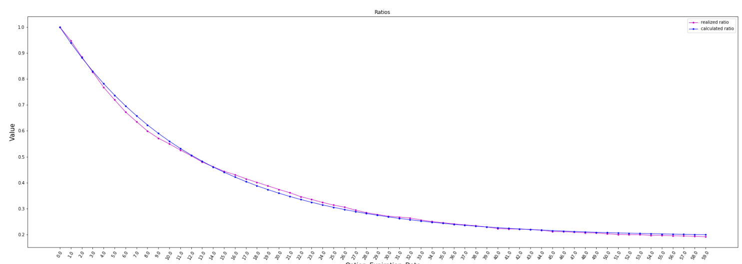

Figure 17: model 1a Realized and Calculated Volatility in December 2019

4.3.3 Model 1a (fixed B)

In model 1a, the Samuelson parameter B is too volatile, so we fix $B = 0.3$. Like we did before, do the calibration on long term volatility term σ_∞. The results are shown in Table 10. Also RMSEs does not increase a lot. After removing one parameter, the results does not perform too bad. And figure 18 and figure 19 shows that model 1a (fixed B) still can fit the ratios.

01/17/2019 0.7887	03/04/2019 0.6752	04/15/2019 0.4541	05/29/2019 0.8763	07/11/2019 0.7324
08/22/2019 0.8047	10/04/2019 0.4273	11/15/2019 0.9602	12/31/2019 0.5728	02/13/2020 0.5418
03/27/2020 0.2619	05/11/2020 0.0000	06/23/2020 0.3844		

Table 10: model 1a (fixed B) σ_∞

01/17/2019 2.53E-03	03/04/2019 9.15E-04	04/15/2019 8.54E-05	05/29/2019 1.53E-03	07/11/2019 2.44E-03
08/22/2019 1.90E-03	10/04/2019 6.78E-04	11/15/2019 9.41E-04	12/31/2019 7.50E-04	02/13/2020 8.22E-04
03/27/2020 1.92E-03	05/11/2020 2.12E-02	06/23/2020 3.85E-04		

Table 11: model 1a (fixed B) RMSE

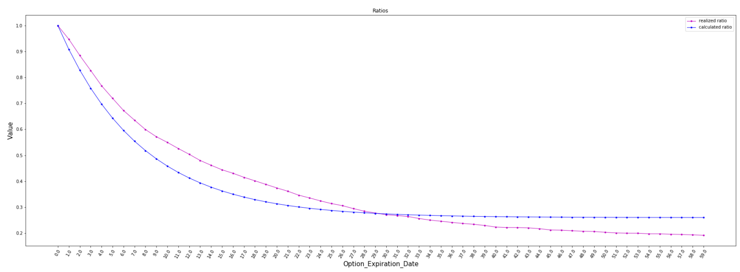

Figure 18: model 1a (fixed B) Realized and Calculated Ratios in December 2019

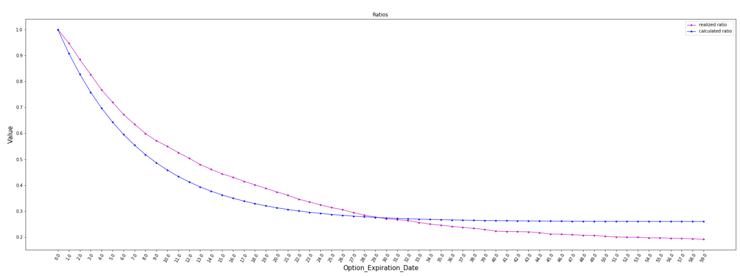

Figure 19: model 1a fixed B Realized and Calculated Volatility in December 2019

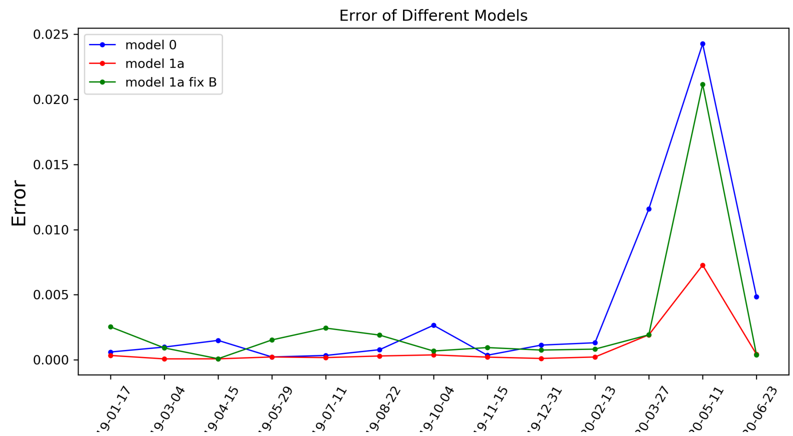

Figure 20: RMSE of three different models

From the results from plot 20. We can find that model 1a performs much better than other two models. But the RMSE is still larger compared to the historical volatility method. And if we want to fix B to a constant value, the RMSE will be much larger because B is much more volatile when dealing with nearby contract. If we want to keep B constant, the RMSE will definitely be large. And the data range of historical volatility method is 1 year, which is much larger than nearby contract method. Maybe this is the reason of smaller error.

4.4 Summary

4.4.1 Summary of RMSEs

According to the results we have before, we know there was a very volatile period in 2020 spring, both parameters and $RMSE_{volatilities}$ behave differently in that period. So when we summarize the results, the whole period is split into two parts, the normal period and the volatile period, we call it 1st period and 2nd period for simplicity.

In Table 12, this is how we split into two periods. Notice that, for implied volatility, we only have data from Jan 2019 to May 2020. For historical volatility, since we get it from historical data, and we calculate historical volatility on yearly based, it is only until May 2020 that can we get the historical volatility in most volatile period.

Data Source	1st period	2nd period
Implied volatility	Jan 2019 - Feb 2020	Mar 2020 - May 2020
Historical volatility	Mar 2019 - Apr 2020	May 2020 - Jun 2020
Nearby contracts	Jan 2019 - Feb 2020	Mar 2020 - Jun 2020

Table 12: Normal and Volatile period

And because results from implied volatility perform extremely bad, we only use model 0 to calibrate. Also since model 2 does not improve the fitting effect too much with one more parameter introduced, we do not introduce long term decays when we do the calibration with Historical volatility of nearby contracts. We call the model 1 with a fixed Samuelson parameter B as model 2 instead. That is why for nearby contracts, $RMSE_{volatilities}$ will increase when we change model 1 to model 2.

As we can see in Table 13 and Table 14, the results from historical volatility perform much better than others.

Data Source	Model 0	Model 1	Model 2
Implied volatility	0.0501		
Historical volatility	0.0007	0.0001	0.0001
Nearby contracts	0.001	0.0001	0.0014

Table 13: $RMSE_{volatilities}$ Summary in the 1st period

Data Source	Model 0	Model 1a	Model 2a
Implied volatility	0.2124		
Historical volatility	0.0051	0.0009	0.0034
Nearby contracts	0.0136	0.0032	0.0078

Table 14: $RMSE_{volatilities}$ Summary in the 2nd period

4.4.2 Comparison between different data sources

As we saw previously, fitting using implied vols does not bring particularly good results compared with other data sources. So, in this part, we only do the comparison between historical volatility of future data and historical volatility of nearby contract data.

For those two data sources, we both use Crude Oil WTI Futures historical price data to generate volatility. However, the processes are different. When we calculate the historical volatility of future data, the length of the moving window is one year, and we move the window forward or backward one month in each time. There will be overlapping parts every time we calculate. For historical volatility of nearby contract data, although the daily log-return formula and the way to get standard deviation are the same, when we switch into another date, we never take log-returns on different contracts, which means that there is no overlap in this method. And the length of the window is also different, 30 business days for nearby contracts.

For historical volatility of future data, the window is relatively longer than the other, which means overlaps bring smoother and more stable results. But it also includes too many changes, which may be more difficult to interpret. And for Historical volatility of nearby contract data, shorter window can capture changes on time, however, it also can can introduces additional noise.

As it is shown in Table 15, we compare the results of those two different data sources. They are both calibrated by model 1a, Samuelson one decay variance model. And since we move the window with different step lengths, we just match the first date of those two situations. We also avoid the most volatile period. It is easy to find that historical volatility generate more stable results.

n^{th}	**Historical volatility of future data**		**Historical volatility of nearby contract data**	
Date	$\boldsymbol{B}$	$\boldsymbol{\sigma_\infty}$	$\boldsymbol{B}$	$\boldsymbol{\sigma_\infty}$
1	0.3081	0.6705	0.2093	0.3524
2	0.3069	0.6568	0.3987	0.5381
3	0.3181	0.6646	0.8235	0.4578
4	0.3339	0.6680	0.1449	0.0000
5	0.3028	0.6581	0.2113	0.2635
6	0.2795	0.5848	0.2465	0.4823
7	0.2815	0.5893	0.6684	0.3971
8	0.2749	0.5551	0.1979	0.5462
9	0.3270	0.5372	0.4777	0.4772
10	0.3208	0.5218	0.4711	0.4369

Table 15: Comparison of two data sources

5 Application: Mapped Volatility

Set today $t_1 = July302020$, the pricing date of a swaption on WTI future contracts, the expiration of which is $t_2 = Dec152020$. We first get implied volatilities for all 12 months in 2021.

Our goal is to figure out which volatilities for WTI futures we should use in the "swaption evaluation". A swaption is an option granting its owner the right but not the obligation to enter into an underlying swap. We need to take into account that the volatility decreases faster to the expiration of the contract. We will use the result we get from model 1a with historical volatility,the latest data (June 20), in which $B = 1.038$, $\sigma_\infty = 0.215$.

Then we can calculate the mapped volatility through the following steps:

1. For each contract with expiration T_i, calculate the parameter $sigma_0^i$, so that the calculated model variance from t_1 until expiration of the option on the contract $\sigma^2_{i,implied} \times (\tau_i - t_1)$

$$\sigma^2_{0,i} = \frac{2B\sigma^2_{imp,i} \cdot (\tau_i - t_0)}{g(T_i, \tau_i)}$$

where

$$g(T_i, \tau_i) = [e^{-2B(T_i - \tau_i)} - e^{-2B(T_i - t_0)}] \\ + 4\sigma_\infty[e^{-B(T_i - \tau_i)} - e^{-B(T_i - t_0)}] + 2B\sigma^2_\infty(\tau_i - t_0)$$

2. Given B, σ_∞ and σ_0^i, calculate variance between t_1 and t_2, $Var_i[t_1, t_2]$. From the other hand, we can write this variance as $\sigma_i^{*2} \times (t_2 - t_1)$, where $sigma_i^*$ is the "mapped" volatility to the option expiration t_2. This mapped volatility will be used in Monte Carlo simulations to value swaption.

$$\sigma_i^{*2} = \frac{1}{t_* - t_0} \cdot \sigma^2_{0,i}\{\frac{1}{2B}[e^{-2B(T_i - t_*)} - e^{-2B(T_i - t_0)}] + \\ \frac{2\sigma_\infty}{B}[e^{-B(T_i - t_*)} - e^{-B(T_i - t_0)}] + \sigma^2_\infty(t_* - t_0)\}$$

Contract	Expiration	Implied Volatility	Mapped Volatility
CLG21	01/20/2021	0.4175	0.399045
CLH21	02/22/2021	0.4220	0.367149
CLJ21	03/22/2021	0.4199	0.381174
CLK21	04/20/2021	0.4241	0.353148
CLM21	05/20/2021	0.4259	0.337937
CLN21	06/22/2021	0.4276	0.324197
CLQ21	07/20/2021	0.4290	0.312335
CLU21	08/20/2021	0.4305	0.299001
CLV21	09/21/2021	0.4317	0.288044
CLX21	10/20/2021	0.4330	0.278205
CLZ21	11/19/2021	0.4343	0.267931
CLF21	12/20/2021	0.4352	0.259898

Table 16: Mapped Volatilities

6 Conclusions

We try different forms of Samuelson parameterization: a simple exponential decay form, a exponential decay form with long term volatility term, two exponential decays form with long term volatility term. And use those models on different volatilities: implied volatility, historical volatility of future data, historical volatility of nearby contract data. It turns out that, introducing long term volatility term will significantly improve fitting result. And it works relatively well in historical volatility. Using two exponential decays forms will not enhance too much by including three parameters. Finally, we use one exponential decay with long term volatility term to re-calculate volatility of swaption.

7 Further Research

The next step of our research is to generalize the calibration procedures to seasonal commodities as Natural Gas , Power and Agricultural products.