Seoul National University Department of Economics

Modelling Volatility and Risk Spillover Between the Financial Markets of US and China Using GARCH Value-at-Risk Forecasting and Granger Causality

John Jongsun Suh

Abstract

In this paper, volatility forecasting and risk spillover evaluation is performed on a dataset consisting of the intraday returns from January 2007 to April 2020 of the S&P500, SSE Composite Index, and the Chinese Yuan to USD exchange rate. The chosen date range includes both the Great Recession of 2008 and the more recent Coronavirus Recession, offering a rare opportunity for comparative analysis of those two periods of elevated volatility against the overall backdrop of a strong, sustained bull market. For volatility forecasting, a Skewed Student's t ARIMA-GARCH model is used to estimate and forecast Value-at-Risk series', which are in turn used to test for the presence of Granger Causality between the US and Chinese financial markets. The GARCH-VaR method exhibits a high fit to historical data, and while a considerable degree of risk spillover is observed between the US and Chinese economies throughout the date range, its predictive power is shown to markedly diminish during the two Recession periods.

Keywords: VaR(Value at Risk), ARIMA-GARCH model, Risk management

Table of Contents

I. Introduction

Notable for the strong, consistent trend of global economic growth exhibited throughout, and the decade-long bull run especially in the US and Chinese markets, the 2010s have now ended up being sandwiched by two of the most severe downturns since the early 20th century: the Great Recession of 2007-2009, and the nascent Coronavirus Recession. Even prior to this most recent development, geopolitical shocks in the form of Brexit and the Trump administration-led US-China trade war did mark periods of elevated market volatility, but with the advent of a new age of social distancing, supply chain disruptions, shifting employment practices, skyrocketing fiscal deficits, and negative-approaching interest rates, understanding and anticipating the amplified twists and turns of market volatility has become an ever more important task.

Given that the world's major economies, having gained an acute appreciation for the fragility of their collective overreliance on China's erstwhile vast network of supply chains, are now in the process of frantically rectifying this failing, and moreover that the Trump administration has chosen this of all junctures to demonstrate a renewed interest in further prosecuting its economic conflict with China, the once inconceivable possibility of a fundamental decoupling between the first and second largest economies of the world has now become an alarmingly feasible probability.

All this would seem to provide ample motivation to reexamine the relationship between the economies of US and China, with a particular focus on evaluating just how closely interlinked and correlated the financial markets of the two have been, as well as whether and how the nature and distance of that association has changed over time.

To this end, a comparative analysis will follow of the representative indicators of each country’s financial performance: the S&P500 Index (hereinafter referred to as the “SPX”), the Shanghai Stock Exchange Composite Index (SSE). Additionally, the Chinese Yuan to USD exchange rate will also be examined, as its close oversight by the Chinese government makes it a source of information for implicit policy decisions by the same. Market close quote data of all three assets ranging from January 2007 to April 2020 will be analyzed, with the ARIMA-GARCH model being used to estimate their respective one-day-ahead parametric Value-at-Risk (VaR). Based on VaRs trained on data ending at December 2018, a forecast of the volatility for each asset will be performed for the date range of the Coronavirus Recession, and compared with the historical data from that same period to determine goodness-of-fit.

In addition, the presence and degree of risk spillover between the US and Chinese financial markets will be evaluated by performing bidirectional Granger Causality tests between the VaRs of the US and Chinese assets. In particular, separate tests will be performed on the full target date range, the Great Recession’s range, and the Coronavirus Recession’s range, from which it will be determined if and how the relationship between the US and Chinese markets changed during these intervals of elevated volatility compared to the overall trend.

II. Theoretical Framework

1. ARIMA(p,d,q) Model

The autoregressive integrated moving average or ARIMA model is a generalization of the autoregressive moving average (ARMA) model, both of which are used primarily on time series data to achieve a better fit, particularly in the presence of non-stationarity and autocorrelation. The AR part of the model regresses the time series data for the variable of interest on its own previous, or lagged, instances, with the "p" parameter denoting the extent of this lag, while the MA part formulates the regression error of the overall model as a linear combination of a set of residuals estimated at various contemporaneous and past time periods, again with the "q" parameter showing the number of lags, or order of the model. The remaining "I" component of the model, which stands for "integrated," describes a differencing process by which the data values are substituted by their difference with their immediately previous entries, with the "d" parameter denoting the number of iterations that this process is undertaken.

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1 - L)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right)\varepsilon_t$$

$$\frac{\phi(L)}{\theta(L)} X_t = \varepsilon_t$$

The above formulas describe the ARIMA and ARMA models, respectively. In this study, we used the ARIMA class in the python package "statsmodels" to fit our dataset. Financial data typically exhibits momentum and mean reversion effects, thus necessitating the application of the estimation of first moments in the distribution of the data.

2. GARCH(1,1) Model

The general autoregressive conditional heteroskedasticity model (GARCH) is a generalized version of the ARCH model, which is often employed when modelling financial time series that exhibit volatility clustering and heteroskedasticity of volatility against time.

The ARCH family of models explain the variance of the current residual term using the residual terms of the previous time periods. The error variance in a time series being well-fitted to an AR model indicates that the ARCH model will perform well, while if an ARMA model is assumed for the same, the GARCH model is more appropriately applied.

$$r_t = \mu + \epsilon_t$$

$$\epsilon_t | \psi_{t-1} \sim \mathcal{N}(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha_i \epsilon_{t-1}^2 + \beta_i \sigma_{t-1}^2$$

The above formulations describe the standard GARCH(1, 1) model, where r is the asset returns data, σ^2 are the ARMA-estimated volatility terms of ϵ, which are in turn the residuals of the mean process applied on the returns data.

A limitation of this model is that its assumption of normal errors may be ill-suited for modelling financial data, which, not least due to its susceptibility to black swan events, typically follow fat-tail distributions and are more sensitive to time-variation of the higher moments. Fortunately, Bayesian inference methods enable the assumption of Student’s t or even skewed Student’s t-distributed GARCH innovations, which were implemented using the python “arch” package in this study.

3. Parametric Value-at-Risk

Value-at-Risk (VaR) is a risk management statistic used to estimate and forecast the risk of maximum loss that a given asset portfolio could potentially incur over a specific time horizon, given everyday market conditions and under a set significance level. Originally developed by JP Morgan as a systematic method to distinguish extreme, so-called “black swan” events from everyday price movements in market analysis, the concept has since spawned a number of varieties.

$$VaR(\alpha) = \mu + \sigma \cdot N^{-1}(\alpha)$$

The most naïve among these is the above delta-normal approach, which assumes that all financial asset returns are normally distributed, computing the variance of a certain range of historical returns data. α here is the confidence level, which is normally set to 95% and 99%, and N^{-1} is the inverse pdf of the normal distribution, the corresponding left-tail quantile of which it generates given α.

$$VaR_{t+1|t}(\alpha) = \mu_{t+1|t} + \hat{\sigma}_{t+1|t} \cdot F^{-1}(\alpha)$$

While trivial to implement, the results of this method are naturally underwhelming, which lead us to the one-day-ahead parametric VaR approach used in this study. Given that financial asset returns exhibit time-varying volatility, it becomes necessary to use conditional variance when estimating VaR, which is contributed by the ARIMA-GARCH(1,1) model discussed above. Given the fat tails generally exhibited by financial time series data and the date ranges of high volatility that are contained in the dataset, a new assumption about the underlying distribution of the returns data seems necessary. Taking all this into consideration, the one-day-ahead VaRs using ARIMA-GARCH-estimated conditional volatility and assuming a skewed Student's t distribution for the returns is expressed as presented above.

4. Risk Spillover and Granger Causality

Spillover effect refers to a type of network effect where events, signals, or information affect others in a seemingly unrelated context. In general, the high level of globalization and mutual connection between the financial markets of today's world means that events or fluctuations in large economies are expected to have correspondingly large effects on other countries worldwide.

The Granger causality test, first proposed in 1969, provides a means to statistically determine the existence, degree, and time lag of spillover effects. More specifically, through a series of t-tests and F-tests on lagged values, the Granger causality test accepts or rejects the hypothesis that the information from one time series provides statistically significant information for forecasting the future values of the other. While Granger causality is distinct from true causality, the test is useful in determining whether events in one time series are temporally related to another, as well as the directionality of this relation – that is whether one consistently tends to precede the other.

$$y_t = a_0 + \sum_{i=1}^{m} a_i y_{t-i} + \sum_{j=p}^{q} b_j x_{t-j} + \varepsilon_t$$

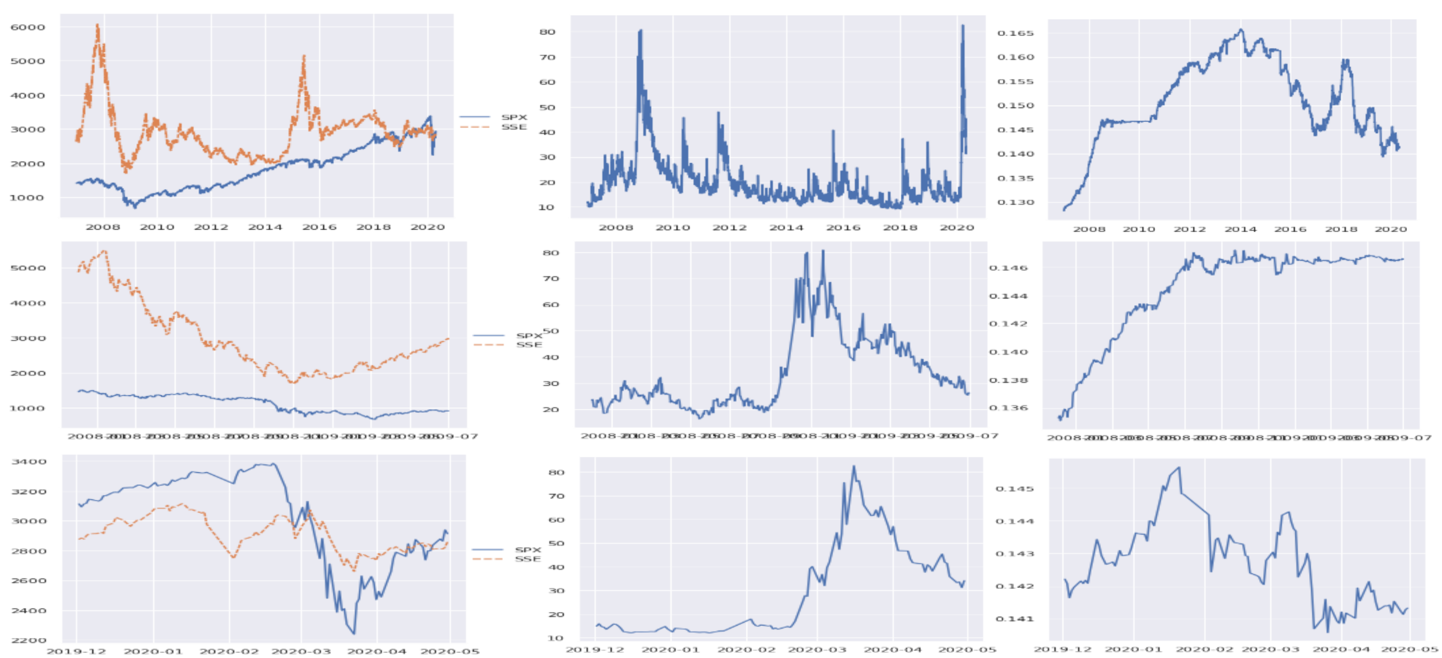
$$H_0: b_j = 0 \ \forall j \in [k, q], p \leq k < q$$

As shown above, to test whether, for two stationary time series y and x, x does not Granger-cause y, the univariate autoregression of y over a set of its lagged values is augmented by including the lagged values of x, and iterative F-tests are performed under the null hypothesis that no explanatory power is provided by adding x's beyond a certain lag level.

III. Empirical Analysis

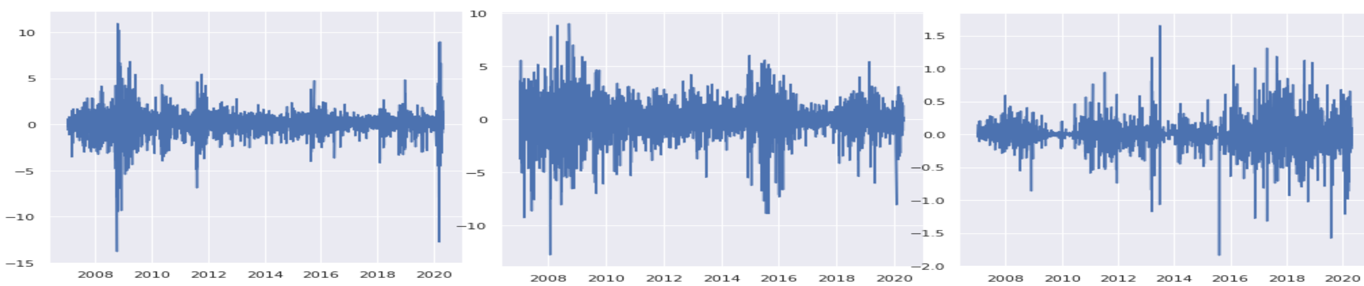
1. Dataset

The subject of analysis for this study is a dataset consists of the daily close price quotes of the three assets SPX, SSE, and CNY from January 2007 to April 2020. The three columns in the following set of plots show a comparison between, respectively, the price changes in the SPX and SSE, the VIX, and the CNY, while the three rows show the full time range of the dataset, the Great Recession period, and the Coronavirus Recession period, the latter two of which were defined as ranging from December 2007 to June 2009, and December 2019 to April 2020.



As these datasets are far from normally distributed as required for the time series analysis models to be employed in the study, the actual analysis was performed on the log percentage rate of daily returns derived from the price data using the following formula.

$$r_{t+1} = 100 \cdot \left(\ln \frac{P_{t+1}}{P_t}\right)$$



The above plots show the calculated log returns throughout the full time range. Not only does this process transform the data into having zero mean across time, which is advantageous for fitting into the ARIMA model, it also corresponds to a first-order differencing process that characterizes the "integrated" portion of the model, thus ensuring stationarity for the dataset.

2. ARIMA Model

The Box-Jenkins method uses an iterative three-stage approach to find the best fit of ARMA and ARIMA models to the past values of a time series. The three stages consist of model identification and selection, during which stationarity and seasonality is checked for, parameter estimation, for which we will attempt to find the model that minimizes the AIC, and finally statistical model checking to determine whether the residuals are independent and constant in mean and variance over time.

a. Stationarity: Augmented Dickey-Fuller Unit Root Test

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{k=1}^{p-1} \delta_k \Delta y_{t-k} + \varepsilon_t$$

$$DF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}, \qquad H_0: \gamma = 0, H_1: \gamma > 0$$

First, in order to confirm the stationarity and non-unit root presence of the log returns data, the augmented Dickey-Fuller unit root test (ADF), as described above, was performed using the python statstools package with the following results:

	ADF Statistic	**p-value**	**Lags Used**	**Critical Value**		**Stationarity**
				5%	1%	
SPX	-10.03840	1.519034e-17	29	-2.862474	-3.432463	TRUE
SSE	-20.690108	0.0	6	-2.862467	-3.432448	TRUE
CNY	-8.693642	3.978517e-14	24	-2.567266	-3.43246	TRUE

b. Parameter Search Under Akaike's Information Criterion

In order to determine the optimal parameters for the ARIMA model, the Akaike's information criterion (AIC) was used, as given by the following, where L is the likelihood of the model.

$$AIC = -2 \log L + 2(p + q + d)$$

The python "statsmodels" package was used to iteratively fit ARIMA models over the dataset while varying the p and q with a maximal order of 10. The d parameter was fixed to 0, as the differencing process already performed for calculating the log returns meant stationarity was secured, as confirmed by the ADF tests, and no further searches for higher d values would be necessary. As a result of this search process, the following three models were found to minimize the AIC value for each of the assets, and were chosen for further analysis: ARIMA(8,0,8) for SPX with an AIC of 10706.840412480784, ARIMA(9,0,7) for SSE, with an AIC of 12169.41590258502, and ARIMA(3,0,2) for CNY, with an AIC of -1011.603848904977.

c. Model Fitting

The fitted model parameters and their p-values and confidence intervals, as well as the log likelihood and standard deviation of each of the models are given in the output tables below. The results from the package also included the residuals from the fitted model, which will be used to fit the GARCH model.

Dep. Variable:	SPX	No. Observations:	3127
Model:	ARMA(8, 8)	Log Likelihood	-5335.420
Method:	css-mle	S.D. of innovations	1.333

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0298	0.021	1.430	0.153	-0.011	0.071
ar.L1.SPX	0.5922	0.449	1.319	0.187	-0.288	1.472
ar.L2.SPX	-0.0488	0.303	-0.161	0.872	-0.642	0.545
ar.L3.SPX	0.1350	0.232	0.582	0.560	-0.319	0.589
ar.L4.SPX	-0.5259	0.160	-3.283	0.001	-0.840	-0.212
ar.L5.SPX	-0.1749	0.183	-0.956	0.339	-0.533	0.184
ar.L6.SPX	0.2696	0.297	0.909	0.363	-0.312	0.851
ar.L7.SPX	0.3441	0.493	0.698	0.485	-0.623	1.311
ar.L8.SPX	-0.5385	0.187	-2.874	0.004	-0.906	-0.171
ma.L1.SPX	-0.7251	0.446	-1.625	0.104	-1.600	0.150
ma.L2.SPX	0.1472	0.367	0.401	0.688	-0.572	0.867
ma.L3.SPX	-0.1528	0.224	-0.682	0.495	-0.592	0.286
ma.L4.SPX	0.5144	0.197	2.611	0.009	0.128	0.901
ma.L5.SPX	0.1106	0.233	0.475	0.635	-0.346	0.567
ma.L6.SPX	-0.3111	0.329	-0.946	0.344	-0.956	0.333
ma.L7.SPX	-0.2670	0.548	-0.487	0.626	-1.341	0.807
ma.L8.SPX	0.5097	0.129	3.946	0.000	0.256	0.763

Dep. Variable:	SSE	No. Observations:	3127
Model:	ARMA(9, 7)	Log Likelihood	-6066.708
Method:	css-mle	S.D. of innovations	1.684

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0287	0.032	0.885	0.376	-0.035	0.092
ar.L1.SSE	-0.2747	nan	nan	nan	nan	nan
ar.L2.SSE	-0.0892	nan	nan	nan	nan	nan
ar.L3.SSE	0.4280	nan	nan	nan	nan	nan
ar.L4.SSE	0.1717	nan	nan	nan	nan	nan
ar.L5.SSE	-0.5307	nan	nan	nan	nan	nan
ar.L6.SSE	-0.2641	0.066	-4.011	0.000	-0.393	-0.135
ar.L7.SSE	-0.6699	nan	nan	nan	nan	nan

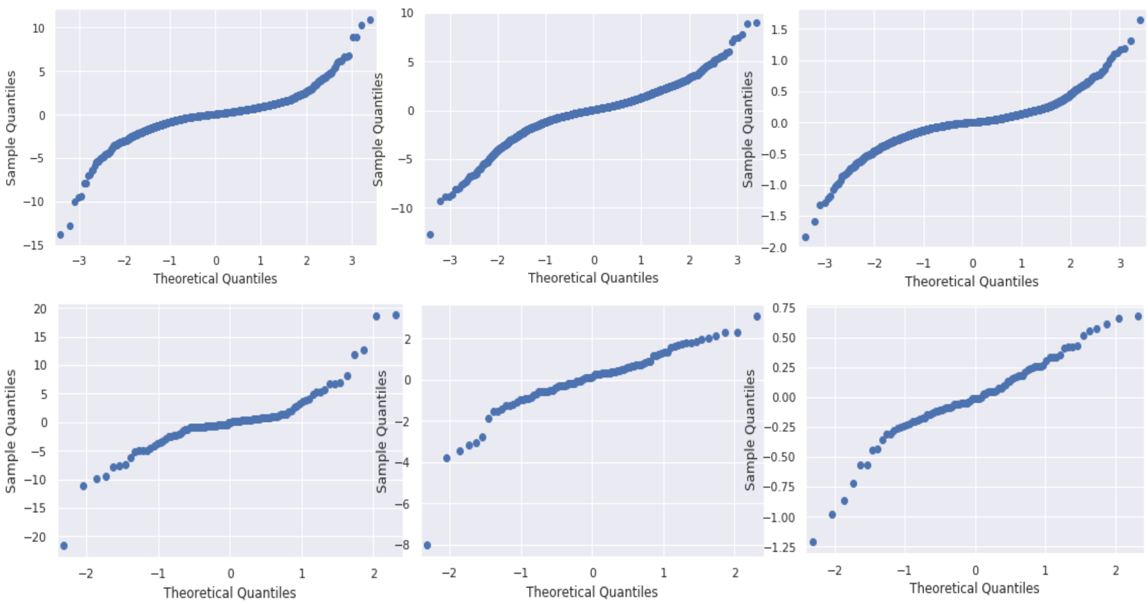
```
ar.L8.SSE      0.0598        nan        nan        nan         nan         nan
ar.L9.SSE      0.0762        nan        nan        nan         nan         nan
ma.L1.SSE      0.2732        nan        nan        nan         nan         nan
ma.L2.SSE      0.1061        nan        nan        nan         nan         nan
ma.L3.SSE     -0.3794        nan        nan        nan         nan         nan
ma.L4.SSE     -0.1136        nan        nan        nan         nan         nan
ma.L5.SSE      0.4993        nan        nan        nan         nan         nan
ma.L6.SSE      0.1966      0.062      3.167      0.002       0.075       0.318
ma.L7.SSE      0.6783        nan        nan        nan         nan         nan


==============================================================================
Dep. Variable:                    CNY   No. Observations:                 3127
Model:                     ARMA(3, 2)   Log Likelihood                 512.802
Method:                       css-mle   S.D. of innovations              0.205
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0007      0.005      0.142      0.887      -0.009       0.010
ar.L1.CNY      0.2920        nan        nan        nan         nan         nan
ar.L2.CNY      0.5506        nan        nan        nan         nan         nan
ar.L3.CNY      0.0794        nan        nan        nan         nan         nan
ma.L1.CNY     -0.4017        nan        nan        nan         nan         nan
ma.L2.CNY     -0.5011        nan        nan        nan         nan         nan
```

3. Diagnostic Tests

In order to use the residuals from the fitted ARIMA models, it is necessary to test for characteristics that may make them inappropriate for usage with the GARCH model, or otherwise require adjustment. Thus, the following three tests for performed to check for normality, heteroskedasticity, and autocorrelation.

a. Normality: Jarque-Bera Test, QQ-plots

Date Range	Asset	JB Statistic	p-value	Skewness	Kurtosis	Normality
Full	SPX	25033.86794234342	0.0	-0.9018581000369447	16.743498313691823	FALSE
	SSE	3450.372896140084	0.0	-0.6259213560761465	7.99147325351148	FALSE
	CNY	14746.49610160699	0.0	-0.33684424292698695	13.617287707579884	FALSE
Coronavirus Recession	SPX	27.128709813412982	1.28551016003307 54e-06	0.3901658145737026	5.498930039740938	FALSE
	SSE	221.23716001577432	9.09831788670082 1e-49	-1.567092481168136	9.744343898839142	FALSE
	CNY	21.235372958027604	2.44792073824114 05e-05	-0.6440325281834429	4.910427085777804	FALSE



The test statistic for the Jarque-Bera test is defined as follows:

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right)$$

$$n\text{: number of observations}, S\text{: sample skewness}, K\text{: sample kurtosis}$$

As can be seen from the results of the test in the above table, as well as the QQ plots that follow, the residuals are decidedly not normally distributed, with the high kurtosis values suggesting fat tails, and some skewness being observed as well. This indicates that the GARCH model will need to assume a non-normal distribution for its in put data.

The bottom three QQ plots are data from the Coronavirus Recession time range, included for comparison. All three assets exhibit significant deviation from normality, with particularly heavy left tails, which matches the volatile crashes observed in both the US and Chinese markets during that time period.

b. Heteroskedasticity: Breusch-Pagan Test

Date Range	Asset	Lagrange Multiplier Statistic	p-value	F-test Statistic	p-value	Heteroskedasticity
Full	SPX	129.61027026137305	3.459273572114013e-29	135.1715130058786	1.2801709900797024e-30	TRUE
	SSE	178.03367030488133	6.77870056065016e-42	188.72146750844604	9.627076375599673e-42	TRUE
	CNY	19.89158701031016	1.0754371179053111e-05	20.01253021435381	7.967903112359148e-06	TRUE
Coronavirus Recession	SPX	0.09209195598987752	0.04259238886976515	0.09121098590682027	0.7633104386422986	FALSE
	SSE	22.20915611276444	1.6878494861937183e-08	28.592569478360307	6.141032702629358e-07	TRUE
	CNY	14.400975113055871	0.0002352858133910397	16.766041477037337	8.882669192386184e-05	TRUE

The Breusch-Pagan test uses the following Lagrange multiplier test statistic to check for the presence of heteroskedasticity in the ARIAM-fitted residuals data. As can be seen from the data, all three assets demonstrated heteroskedasticity in the full time range, thus justifying the usage of a further GARCH model for analysis.

$$\varepsilon_t^2 = \gamma_1 + \sum_{i=2}^{p} \gamma_i z_{it} + \eta_t$$

$$LM = nR^2 \sim \chi_{p-1}^2$$

c. Autocorrelation: Box-Ljung Test

Date Range	Asset	Q*	χ^2	p-value	Degrees of Freedom	Autocorrelation
Full	SPX	3279.148450953904	8865.157180262046	0.0	3110	TRUE
	SSE	3229.0022840631095	12463.235224356216	0.0	3110	TRUE
	CNY	3249.546030219083	3097.5631005111145	0.6136894526472225	3121	FALSE
Coronavirus Recession	SPX	164.2271799173051	966.0926255038767	2.3010562829172895e-148	87	TRUE
	SSE	141.38584986018915	257.51672519658615	2.9165535011330294e-20	81	TRUE
	CNY	166.03647630197528	108.10740411429796	0.082262206539686	89	FALSE

Finally, the Box-Ljung test was performed up to 90 lags to check for the presence of autocorrelation in the ARIMA-fitted residuals. Unfortunately, the null hypothesis of independent distribution was rejected for both the SPX and SSE dataset, thus necessitating further fitting with the GARCH model. The Box-Ljung test uses the following Q* test statistic and null hypothesis:

$$Q^* = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n=k}$$

$$\hat{\rho}_k: sample\ autocorrelation\ at\ lag\ k,$$

$$h: number\ of\ lags\ (90\ in\ this\ study)$$

$$H_0: Q^* > \chi^2_{1-\alpha,h}$$

4. GARCH(1,1) Model Fitting

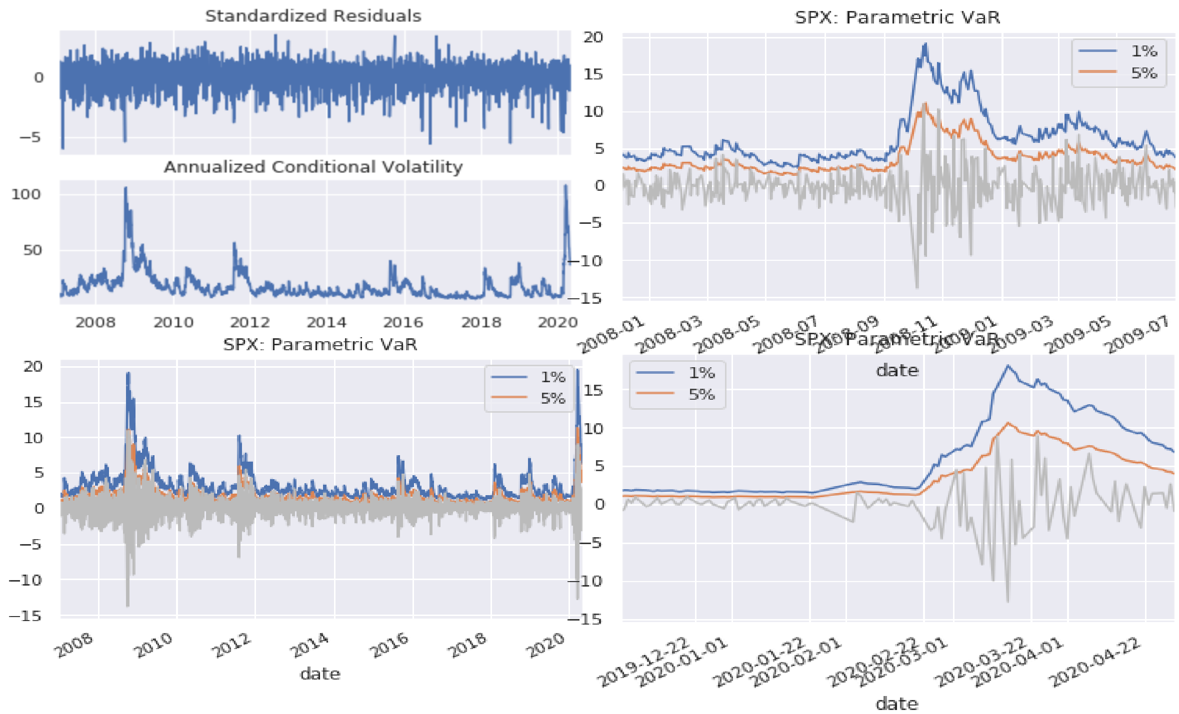
Finally, using the python "arch" package, the results from the fitted ARIMA models were fed into the GARCH(1,1) model, assuming a skewed Student's t distribution due to the mixed results from the diagnostic tests. The volatility models and distribution for the three assets are given as follows:

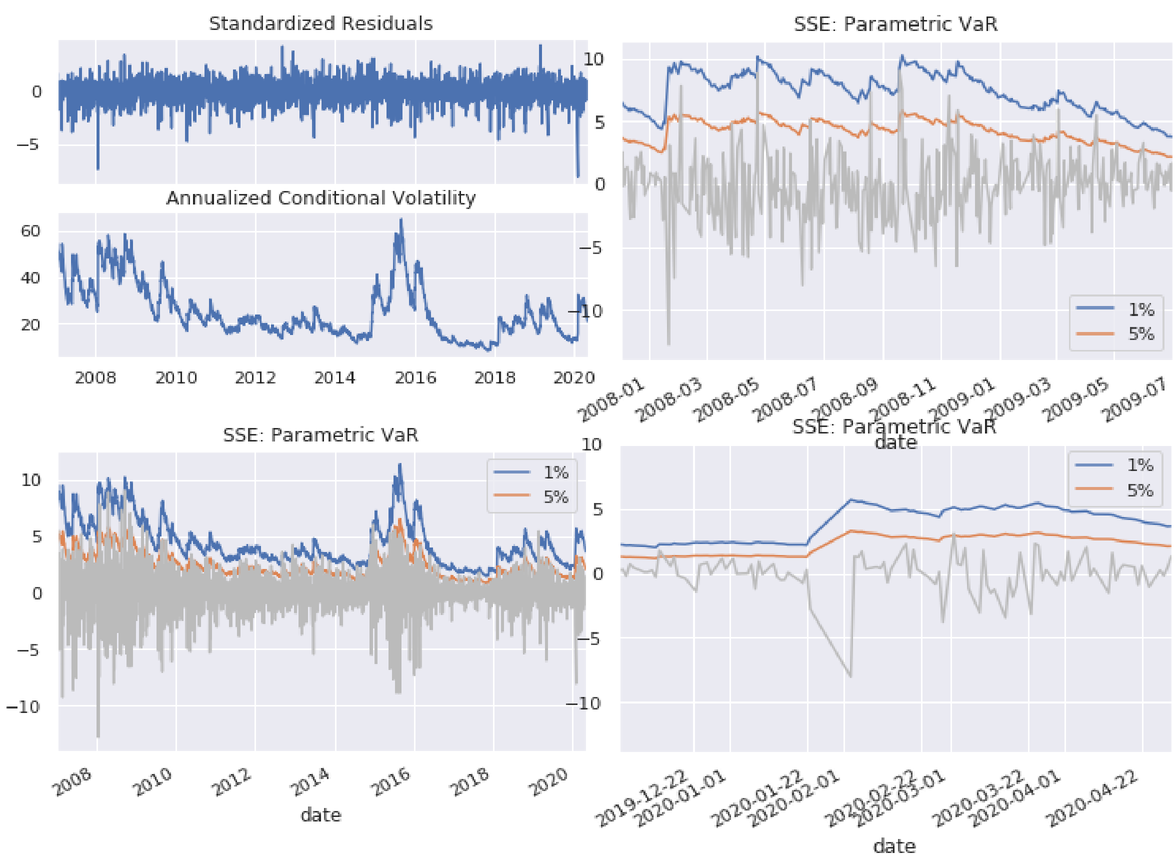
```
================================================================================
                           Dep. Variable: SPX
Vol Model:                                GARCH   Log-Likelihood:               -4215.94
Distribution:      Standardized Skew Student's t   AIC:                           8441.89
Method:                      Maximum Likelihood   BIC:                           8472.13

                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega          0.0170  4.689e-03      3.620  2.944e-04 [7.784e-03,2.616e-02]
alpha[1]       0.1381  1.653e-02      8.350  6.840e-17     [  0.106,  0.170]
beta[1]        0.8619  1.517e-02     56.833      0.000     [  0.832,  0.892]
                              Distribution
========================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
------------------------------------------------------------------------
nu             5.1420      0.456     11.269  1.871e-29 [  4.248,  6.036]
lambda        -0.1593  1.982e-02     -8.037  9.209e-16 [ -0.198, -0.120]
========================================================================
```

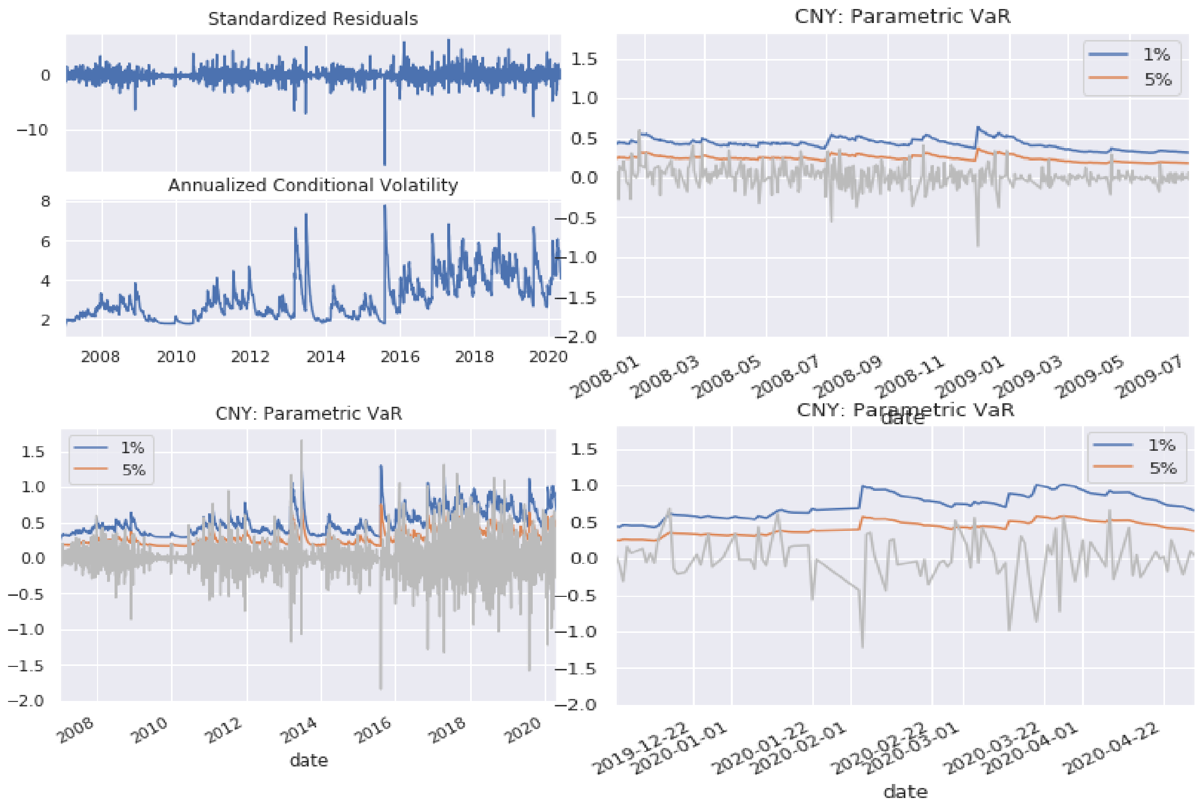
```
==========================================================================================
                                   Dep. Variable: SSE
Vol Model:                                    GARCH   Log-Likelihood:               -5375.57
Distribution:      Standardized Skew Student's t   AIC:                           10761.1
Method:                          Maximum Likelihood   BIC:                           10791.4

                                 Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      7.2603e-03  3.312e-03      2.192  2.835e-02 [7.696e-04,1.375e-02]
alpha[1]       0.0504  7.989e-03      6.306  2.866e-10 [3.472e-02,6.603e-02]
beta[1]        0.9496  7.563e-03    125.570      0.000     [  0.935,  0.964]
                                   Distribution
============================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
----------------------------------------------------------------------------
nu             4.4693      0.360     12.405  2.465e-35    [  3.763,  5.175]
lambda        -0.0792  1.801e-02     -4.397  1.098e-05 [ -0.115,-4.389e-02]
============================================================================

==========================================================================================
                                   Dep. Variable: CNY
Vol Model:                                    GARCH   Log-Likelihood:                1444.80
Distribution:      Standardized Skew Student's t   AIC:                           -2879.60
Method:                          Maximum Likelihood   BIC:                           -2849.36
                                 Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      8.4580e-04  8.424e-05     10.040  1.015e-23 [6.807e-04,1.011e-03]
alpha[1]       0.0500  6.697e-03      7.466  8.265e-14 [3.687e-02,6.313e-02]
beta[1]        0.9300  5.664e-03    164.201      0.000     [  0.919,  0.941]
                                   Distribution
==============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
------------------------------------------------------------------------------
nu             4.1952      0.165     25.494 2.306e-143      [  3.873,  4.518]
lambda         0.0000  1.462e-02      0.000      1.000 [-2.866e-02,2.866e-02]
==============================================================================
```

5. Parametric Value-At-Risk Volatility Forecast and Evaluation

Next, the conditional mean and conditional residual variance series data from the fitted ARIMA-GARCH model was used to generate one-day-ahead rolling-window VaR forecasts. The following set of plots show the standardized residuals and annualized conditional volatility from the GARCH model, as well as the analytically-estimated parametric VaR for each asset over the full time range, and those of the Great Recession and Coronavirus Recession compared to the historical returns data over the same. The Coronavirus Recession in particular was estimated only using training data ranging until the end of 2017.


Standardized Residuals
Annualized Conditional Volatility
SPX: Parametric VaR
1%
5%
date
SPX: Parametric VaR
1%
5%
date
SPX: Parametric VaR
1%
5%
date


Standardized Residuals
Annualized Conditional Volatility
SSE: Parametric VaR
1%
5%
date
SSE: Parametric VaR
1%
5%
date
SSE: Parametric VaR
1%
5%
date



While the plots alone show that the estimated VaR are a pretty good fit to the historical data there is a simple method of roughly testing their level of performance. The failure ratio, or HIT is defined as follows:

$$HIT = \frac{1}{N}\sum_{i=1}^{N} 1\big(r_i < VaR_i(\alpha)\big)$$

Here, 1 is an indicator function that takes the value of 1 when any given historical return does not exceed its corresponding VaR. A model is considered to be well specified if this ratio is equal to or better than the confidence level. That is, the measure simply aggregates the ratio of instances in which the actual data exceeds the maximal loss value estimated by the VaR. The results of this test applied to our dataset are as follows:

Date Range	Asset	HIT		Well-specified	
		5%	**1%**	**5%**	**1%**
Great Recession	SPX	0.9704301075268817	1.0	TRUE	TRUE
	SSE	0.967741935483871	1.0	TRUE	TRUE
	CNY	0.9623655913978495	0.9973118279569892	TRUE	TRUE
US-Chinese Trade War	SPX	0.9870848708487084	1.0	TRUE	TRUE
	SSE	0.966789667896679	0.9981549815498155	TRUE	TRUE
	CNY	0.9446494464944649	1.0	FALSE	TRUE
Coronavirus Recession	SPX	0.8645833333333334	0.9583333333333334	FALSE	FALSE
	SSE	0.9791666666666666	1.0	TRUE	TRUE
	CNY	0.9166666666666666	0.9895833333333334	FALSE	FALSE

While the performance of the VaR forecasts was generally good as expected, the HIT ratios for the Coronavirus Recession period tended to fall short of the mark. In comparison, a forecast for the two years of 2018 to 2019 during which high market volatility was observed due to the ongoing US-Chinese trade war, did yield satisfactory results, as did estimations for the Great Recession period, which makes it probable that not only the historically high levels of volatility, but also the brevity of the time range led to the comparatively lower performance for the more recent recession.

6. Granger Causality Tests

Finally, bidirectional Granger Causality tests were performed on the VaR series for the three date ranges and between the US and Chinese assets, the results of which are given in the following table. The test was performed over 10 lags, and Strong Granger causality was said to be shown when 90% or more of these lags observed precedence in the data at the set significance level, while Weak Granger causality was concluded as present when at least half of the lags exhibited the same. A delayed effect was said to be had when if either the first lag or more than 3 of the first 5 lags did not exhibit Granger causality.

DATE RANGE	EXPLANATORY VARIABLE	RESPONSE VARIABLE	SIGNIFICANCE LEVEL	GRANGER CAUSALITY		
				Strong	Weak	Delayed
FULL	SPX	SSE	5%	TRUE		TRUE
			1%	TRUE		TRUE
	SSE	SPX	5%	TRUE		
			1%		TRUE	
	SPX	CNY	5%	TRUE		TRUE
			1%	TRUE		TRUE
	CNY	SPX	5%		FALSE	
			1%		FALSE	
US-CHINESE TRADE WAR	SPX	SSE	5%	TRUE		TRUE
			1%		TRUE	TRUE
	SSE	SPX	5%		FALSE	
			1%		FALSE	
	SPX	CNY	5%	TRUE		TRUE
			1%	TRUE		TRUE
	CNY	SPX	5%		FALSE	
			1%		FALSE	
GREAT RECESSION	SPX	SSE	5%		FALSE	
			1%		FALSE	
	SSE	SPX	5%		FALSE	
			1%		FALSE	
	SPX	CNY	5%		FALSE	
			1%		FALSE	
	CNY	SPX	5%		FALSE	
			1%		FALSE	
CORONAVIRUS RECESSION	SPX	SSE	5%		FALSE	
			1%		FALSE	
	SSE	SPX	5%		FALSE	
			1%		FALSE	
	SPX	CNY	5%	TRUE		TRUE
			1%		TRUE	TRUE
	CNY	SPX	5%		FALSE	
			1%		FALSE	

From the results, it is apparent that overall, strong risk spillover effects exist between the US and China, especially from the former to the latter, but during shorter timespans, the transfer of volatility from Chinese markets to the US becomes much weaker, with this effect being especially stronger in periods of recession. The ability of the CNY exchange rates to affect the US economy was especially weak, with no Granger causality being observed over any of the time periods observed.

IV. Conclusion

The conclusion that risk spillover tended to decrease during the two recession periods of the last decade was rather surprising, but further study seems required to determine whether these results can be generalized to all such periods of high volatility, especially in the globalized world of today. One possible interpretation is that the scope of this study simply failed to capture the time scale at which macroeconomic effects were transferred between the two countries during those periods. The course that these recessions took in these countries, while certainly interrelated, differed in time range by an order of months. The Shanghai stock market was crashing from early 2008 while the peak of the US financial crisis came in late 2008 to 2009, for example, and likewise with the Coronavirus Recession, which began in December 2019 for China, but only began to severely affect the US markets in February 2020.

A further interesting subject of study would be to examine the implied volatility gauges for the SPX and SSE, respectively the VIX and iVIX, and look at their relationship with realized volatility, and their predictive power, as well as the existence of risk spillover between those two indices. The annualized conditional volatility graphs generated from the ARIMA-GARCH-fitted residuals closely tracked the corresponding implied volatility data, and comparing the forwardlooking ability of the market to predict events compared to statistical methodology may yield valuable insights in risk management.

In this time of great uncertainty, when the novel Coronavirus threatens to greatly accelerate the alarming global trend towards isolationism and protectionism, the US and China seem locked on a collision course that will lead to further market fluctuations, at the very least, and the juggernaut fiscal response of the world's governments towards this unprecedented event threatens to either greatly exacerbate or reverse the already prevailing trend of extreme low interest rates and deflation, furthering our understanding of the relationships between the world's largest markets and how any future shock events or volatile outcomes may affect the global economy seems a vital task. While the many flaws and limitations of the methodology of this study may have kept it from unveiling any previously unknown insights, it is a comforting thought that every day more data and analysis continues to be generated to improve our understanding of our uncertain fates, whatever such may entail.

V. References

Box, G; Jenkins, G. (1970), “Time Series Analysis: Forecasting and Control”, San Francisco: Holden-Day.

Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity”, Journal of Econometrics, April, 31:3, pp. 307–27.

Granger, C. W. J. (1969), “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods,” Econometrica 37, 424-438.

Granger, C.W.J. (1980), “Testing for Causality: A Personal View,” Journal of Economic Dynamics and Control 2, 329-352.

Hamilton, J.D. (1994), “Time Series Analysis”, Taylor & Francis US.

Hansen, B. (1994), “Autoregressive Conditional Density Estimation,” International Economic Review 35, 705-730.

Lee, S. and B. Hansen (1994), “Asymptotic Theory for the GARCH(1,1) Quasi-maximum Likelihood Estimator,” Econometric Theory.

Morgan, J.P. (1996), “Risk Metrics–Technical Document”, 4rd Edition, Morgan Guaranty Trust Company: New York.