National Cheng Kung University

Institute of International Business

Department of Statistic

TIME–SERIES COURSE

FINAL REPORT

Professor: Martin Lukasa

Students: Celia R66107043

Steve R68097066

INTRODUCTION

The exchange rate between the Chinese Yuan (CNY) and the US dollar (USD), as with other currency pairs, is the value of one currency against the other. Typically, the Yuan's exchange rate to the US dollar is quoted as USD/CNY to represent how many Yuan it takes to buy one US dollar. The relevance of this exchange rate is of utmost importance to each of the respective countries as China is the largest exporting country in the world and is also the largest importer to the United States. So, the USD/CNY exchange rate has a significant impact on each nation's economy by affecting the value of imports and exports. Hence, when the Yuan depreciates against the US dollar, Chinese goods exported to the US become cheaper. When the Yuan appreciates against the USD, it would have the opposite effect.

A cornerstone of China's economic policy is managing the yuan exchange rate to benefit its exports, which has been widely criticized by experts as the Yuan is consistently undervalued. China does not have a floating exchange rate determined by market forces, as is the case with most developed economies. Instead, it pegs its currency, the Yuan to the US dollar at a daily reference rate set by the People's Bank of China (PBOC) and allows the currency to fluctuate within a fixed band (set at 1%) on either side of the reference rate. In cases when the Yuan would appreciate significantly against the US dollar if it were allowed to float freely, China caps its rise by buying dollars and selling Yuan. The actual value of the Yuan is challenging to ascertain, and although various studies over the years suggest a wide range of undervaluation - from as low as 3% to as high as 50% - the general agreement is that the currency is substantially undervalued.

The exchange rate is a critical component in finance and political economics as well. Over the last decades, many scholars, financial managers, and investors have become more conscious of how the exchange rate moves and how to predict it. There are lots of models and methodologies are suggested to forecast the exchange rate. In the aspects of economics model, CAPM model, APT model, and Multi-factors model are widely used in previous research. On the other hand, statistical, machine learning, and deep learning methods are also commonly used to predict the exchange rate. Among these approaches, statistical time series analysis has played as a powerful method nowadays. In this context, the current work aims to perform a time-series analysis on the exchange rate between the US Dollar and the Chinese Yuan. The data used for analysis is the close price of the daily value of the exchange rate between the US dollar and the Chinese Yuan (USD/CNY) for a year from 12/02/2020 to 12/02/2021.

The study proceeds as follows. The following section presents methodologies and techniques which are used to forecast data. Section 2 discusses the characteristics of the data, the data analysis process, and findings. The last section concludes.

1. METHODOLOGY

This section describes the various methodologies and the motivation of using the ARIMA, ARCH, and GARCH models to analyze our series.

Time Series Analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend, or seasonal variation) that should be accounted for. One of the most common methods used is the Auto Regressive Integrated Moving Average (ARIMA) method. ARIMA models intend to describe the current behavior of variables in terms of linear relationships with their past values. It has an Integrated (I) component, which represents the amount of differencing to be performed on the series to make it stationary. The second component of the ARIMA consists of an ARMA model for the series rendered stationary through differentiation. The ARMA component is further decomposed into AR and MA components. Since the ARIMA procedure is carried out on stationary data. The notation Zt is used for the stationary data at time *t*, whereas *Yt* is the non-stationary data at that time. The ARIMA process considers linear models of the form

$$Zt = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + ... - \Theta_2\, e_{t-1} - \Theta_2\, e_{t-2} + e_t$$

Where et is zero-mean white noise with variance (et) = σ2e. Since the ARIMA modeling have the assumptions imposed on the errors: they are independent and identically distributed with zero mean and have a constant variance. In most real-life cases and especially in finance, it is fairly common to observe the violation of the variance assumption. In this project, we are dealing with financial data; hence when studying financial time series, researchers can regularly observe common characteristics (Cont, 2001). The main stylized facts for financial time series data are the absence of linear autocorrelation, heavy tails, asymmetry for gains or losses, agglomeration of volatilities, and leverage effect (Perlin et al., 2020). Therefore, the autoregressive conditional heteroskedasticity (ARCH) and the generalized autoregressive conditional heteroskedasticity (GARCH) model proposed by (Engel 1982) and (Bollerslev 1986) respectively would be a more appropriate method to model the USD/CNY exchange rate.

2. DATA

The data set comprises 262 daily closing prices of the foreign exchange rate between the USD and CNY from December 12, 2020, to December 12, 2021. Figure 1 below shows the daily chart of the exchange rate for the one-year period. The Yuan depreciated to its highest peak on March 30, where the closing price was 6.572. This means that the Yuan was weaker against the dollar; this made exporting to the US cheaper. Similarly, the trough in the plot shows the Yuan has appreciated against the dollar making imports to the US more expensive, which could lead to less demand for Chinese goods. Unfortunately, the data set does not include the trade war that occurred in 2019, but it does include the COVID-19 pandemic. Hence, lots of remarkable patterns are played in the graph below (see Figure 1).

Figure 1. The plot of USD/CNY exchange rate

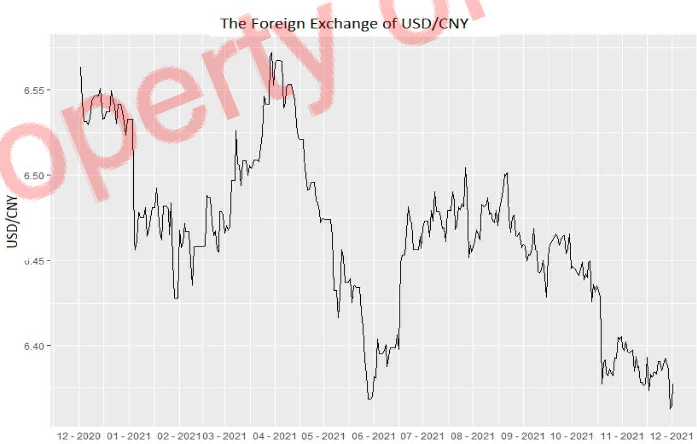


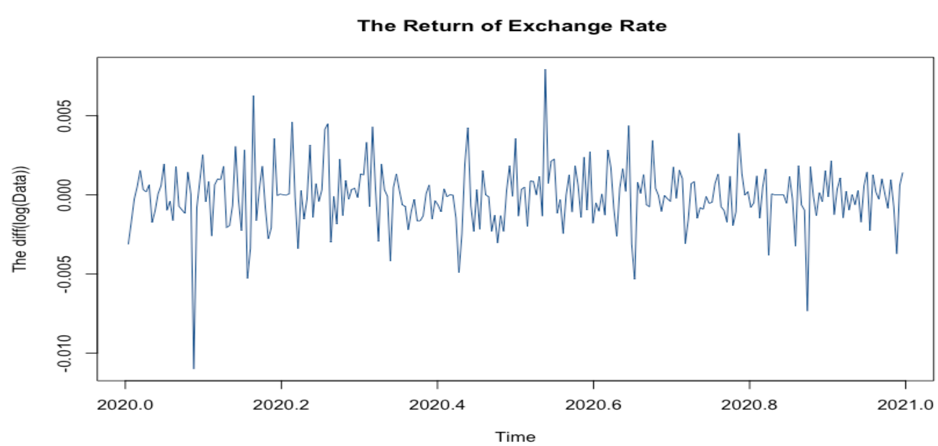
Figure 1 shows how the exchange rate fluctuates over time. At first glance, a downward trend is presented. At the beginning of the period, the exchange rate was quoted at 6.6; this quotation decreased to approximately 6.3 at the end. The data declines from December 2020 to February of 2021. Then, this data increases from 02/2021 to 04/2021; after that, it witnesses a significant fall. This fall lasts until the end of May of 2021 and reverses after that. The exchange rate has a tendency of stable movements from 06/2021 to 09/2021 and then plays a decreasing trend until the end of the period. According to this chart, the exchange rate volatility over the analyzed period is very larger. This dispersion is explained sentiment of investors and macroeconomic factors. Table 1 represents the probable factors that affected the fluctuations that occurred in the series.

Date	***Event***	***USD/CNY***
January 4th, 2021	Robust exports growth in China, which has resulted in a larger current account surplus than previous year, has prompted the yuan to rise as more traders exchanged US dollars for the yuan	6.4616
March 30th 2021	US 10-year Treasury yield rises to 14-month highs	6.572
May 31 2021	The US Inflation rate jumps to 5% from 2.6% causing the USD to depreciate against all currency over these few months. US lifting Covid restrictions to the public.	6.370
June 2021	The Federal Bank announces that it will have three interest rates hikes pushing the interest rate to 2.1% by 2024.	6.4815
October/ November 2021	virtual meeting between U.S. President Joe Biden and his Chinese counterpart Xi Jinping – likely to remove some restrictions on tariffs on trade in coming months.	6.3830

The original data shows a sharp fluctuation in time-vary, which is one of the most considerable problems in time-series analysis. Consequently, the original data is transformed into other formats to reduce variance, keep time consistent, and then get a normal distribution. Taking return of original data is widely used in previous research as the most efficient method when handle with financial data. The return could be gained through transforming original data by log function and then taking the first difference of logged data. This study also uses return of exchange rate to generate forecasted model.

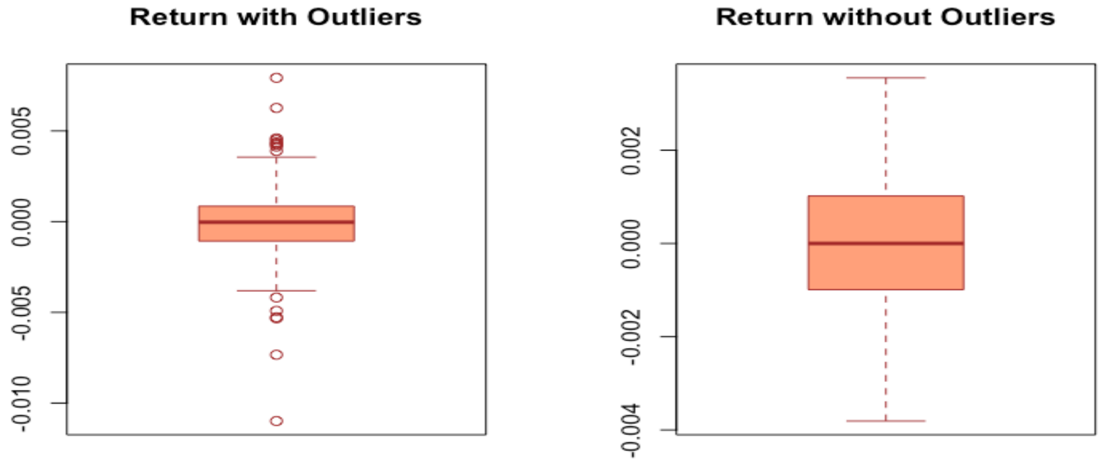
Figure 2 indicates patterns of the return of data over analysis time. At first glance, the data shows less fluctuation as opposed to original data and seems stationary.

Figure 2: The Graph of Return



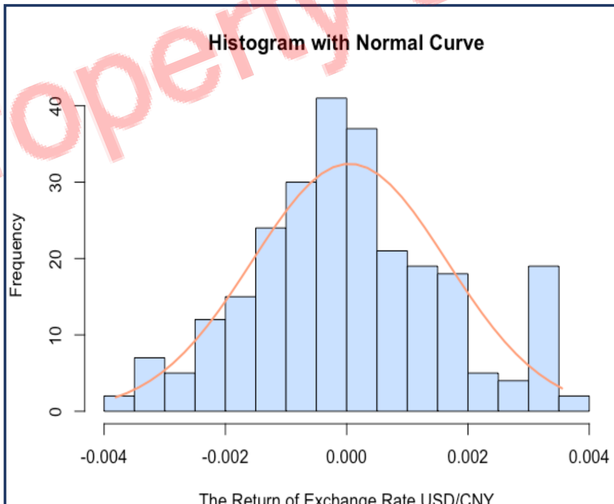
15 outliers were found at both higher and lower extreme in the return. The data also reported has no missing data and the total number of observations is 261 obs. The outliers could harm to accuracy of model, so we replace outliers at higher extreme by the 95th percentile value and outliers at lower extreme by the 5th percentile value. The result of replacing outliers process is showed at Figure 3.

Figure 4: Result of Replacing Outliers



When return distribution is plotted alongside the fitted normal distribution, it is immediately apparent that the data's distribution shape is flat, potentially asymmetrical, and has higher peak than fitted normal distribution, implying that the data is platykurtic and skewed (Table 1).
Further investigated, kurtosis is negative and less than 3 (-0.27 of value), however higher peak opposed to normal distribution implies higher risks derived from abnormal events. Hence, the exchange rate (USD/CNY) could be relatively risky; however, high risk expects high returns. The positive skewness suggests that the data shifts to the right-hand side of the mean. As a result, the null hypothesis of normality for differenced data is not rejected at the level of 5 through the Jarque-Bera statistic test. Hence, the data normally distributes.

Table 1: Descriptive Statistic of transformed data



Statistics	***Value***
Mean	4.64118566922243e-05
Standard Deviation	0.00161
Skewness	0.14305
Kurtosis	-0.27434
Prob. of Jarque-Bera	0.449681

Table 2 reports the result of the stationary test through the Augmented Dickey-Fuller Test. This result shows that the null hypothesis fails to be rejected in original data, whereas it is rejected in transformed data. Hence, the transformed data has no unit root; in other words, the transformed data is stationary.

Table 2: Unit root test results

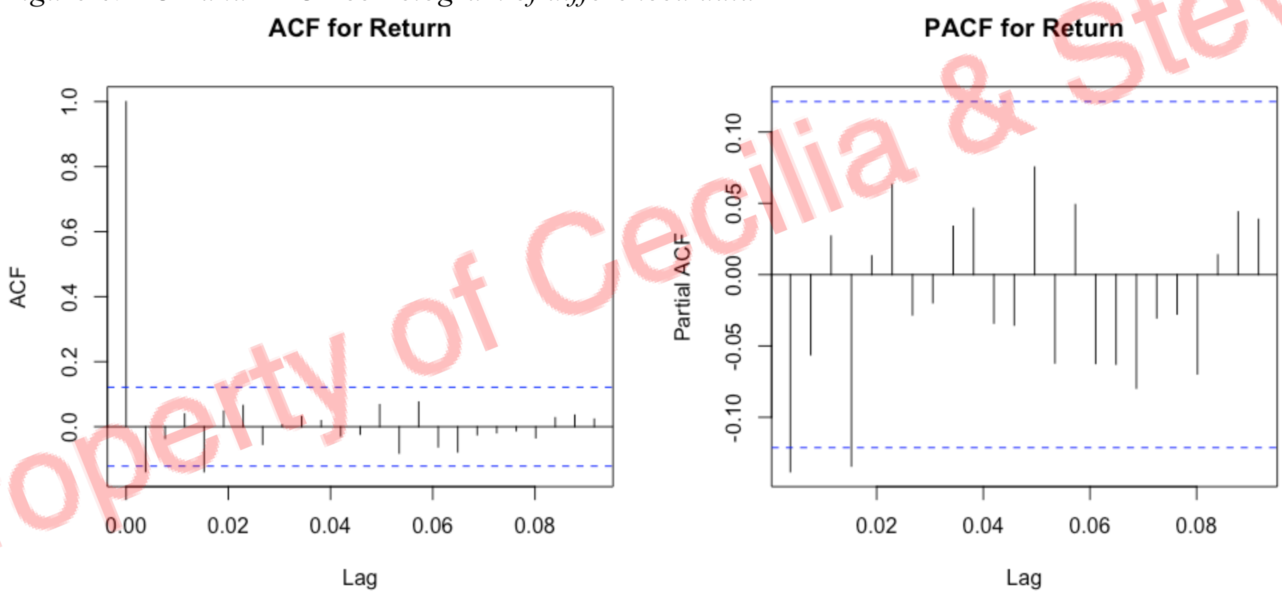
Augmented Dickey Fuller Test	
Ho: The data has a unit root *H_1: The data has not a unit root*	
Data	***Prob.***
Exchange Rate (USD/CNY)	0.3951
The Return of Exchange Rate (USD/CNY)	0.01

3. DATA ANALYSIS

3.1 ARIMA MODEL

The first step to identify ARIMA(p,d,q) model is examining ACF and PACF correlograms. These correlograms supports in finding correlated signs in data so that fit the order of AR and MA terms. Figure 4 shows the ACF and PACF plots for the return of USD/CNY exchange rate. The PACF correlogram is used to determine the order used in the AR models while the ACF is used to determine the order used in the MA model.

Figure 4: ACF and PACF correlogram of differenced data



From the figure, spikes die out significantly, spikes at lag 1 in both correlograms are out of interval, which indicates model probably has (1,1) of order. Additionally, In PACF correlogram, spike at lag 4 is also out of interval, indicating for possibility of lag 1.

We try several models. After each time estimate model, we conduct diagnostics check by testing white noise of residuals from each model. When the model fit is adequate, the model residuals should behave like white noise. We use Ljung-Box Test to check because it is supposed to be efficient and powerful test. The null hypothesis of Ljung-Box Test indicates that examined series is white noise, hence we choose models that fail to reject null hypothesis. In other words, we will choose models having Prob. Value of Ljung-Box Test greater than 0.05. The next step, we compare these models and select one having lowest AIC. After this process, model of ARIMA(0,0,1) is selected as our best model due to its lowest AIC. Figure 4 illustrates residuals from ARIMA(0,0,1) model and the result of this process is reported in Table 3 and Table 4.

Figure 4 Diagnostic Checking of Residual from ARIMA(0,0,1) Model

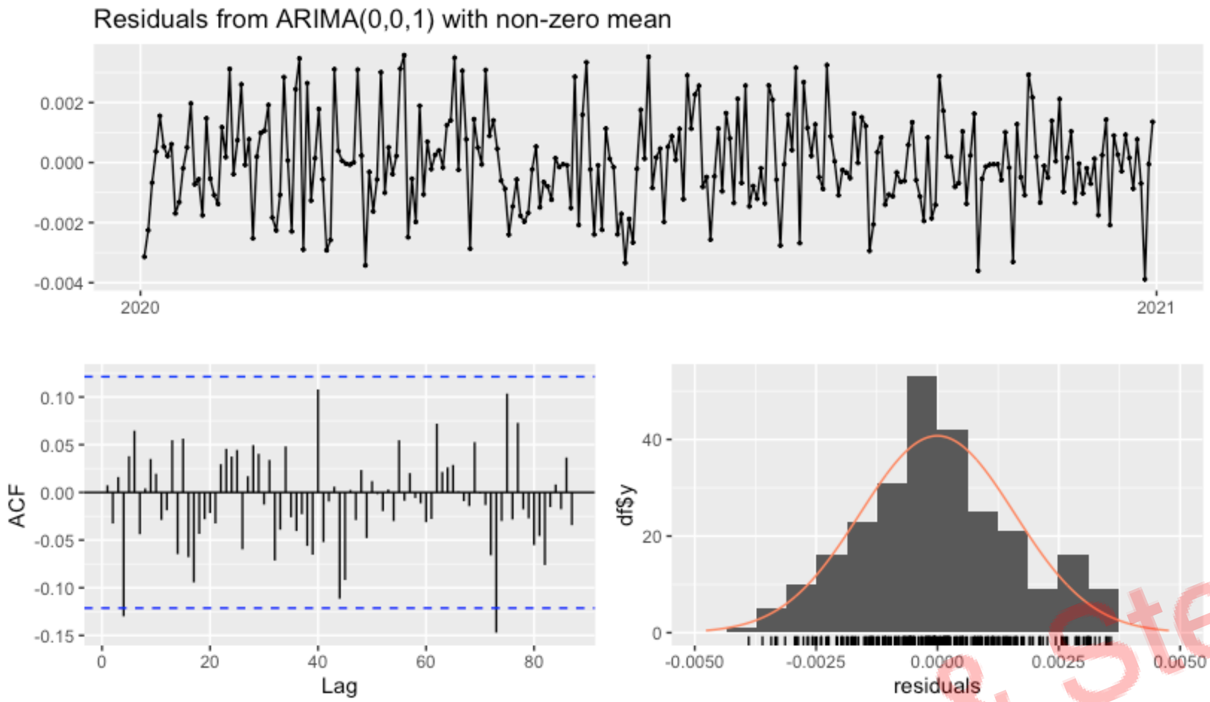


Table 3: The Results of Estimate Process and Diagnostic Checking

Model	***AIC Criteria***	***BIC***	***Prob. of Ljung-box Test***
Random Walk – ARIMA (0,0,0)	-2614.71	-2607.585	0.7256
ARIMA(1,0,0)	-2617.82	-2607.131	0.9338
ARIMA(2,0,0)	-2616.72	-2602.463	0.954
ARIMA(0,0,1)	**-2618.35**	**-2607.652**	0.9452
ARIMA(1,0,1)	-2616.85	-2602.504	0.9468
ARIMA(2,0,1)	-2614.94	-2597.116	0.9481
ARIMA(0,0,2)	-2616.76	-2602.504	0.9506
ARIMA(1,0,2)	-2614.95	-2597.123	0.9488
ARIMA(2,0,2)	-2614.48	-2593.093	0.9785
ARIMA(4,0,1)	-2615.76	-2590.804	0.988

Table 4: Estimated Values of Model

	Constant	AR(1)	AR(2)	MA(1)	MA(2)
ARIMA(1,0,0)	0.00004726 **(0.65616)**	-0.14021833 **(0.02312)***			
ARIMA(2,0,0)	4.7837e-05	-1.4916e-01	-5.9128e-02		

	(0.63959)	**(0.01675)***	**(0.34318)**		
ARIMA(0,0,1)	4.7604e-05 **(0.64584)**			-1.5459e-01 **(0.0161)***	
ARIMA(1,0,1)	4.9170e-05 **(0.6182)**	3.2537e-01 **(0.3958)**		-4.7292e-01 **(0.1874)**	
ARIMA(2,0,1)	4.9776e-05 **(0.6098)**	4.7363e-01 **(0.3189)**	3.5529e-02 **(0.7482)**	-6.2403e-01 **(0.1834)**	
ARIMA(0,0,2)	0.00004834 **(0.63044)**		-0.14880684 **(0.01778)***	-0.04552925 **(0.51760)**	
ARIMA(1,0,2)	4.9483e-05 **(0.6127)**	5.3747e-01 **(0.4213)**		-6.8809e-01 **(0.3060)**	4.3781e-02 **(0.7576)**
ARIMA(2,0,2)	4.8287e-05 **(0.6327)**	-7.9575e-01 **(0.2185)**	1.0100e-01 **(0.8798)**	6.5430e-01 **(0.3014)**	-2.7354e-01 **(0.6769)**

Figure 4 visually shows that the residual is possibly white noise, however, its fluctuation seems asymmetric. Consequently, a GARCH model could be suggested, and we will discuss this thing at the next part.

The equation for the ARIMA(0,0,1) is given below:

$\hat{y}_t = u_t - 1.5459e\text{-}01 * u_{t-1}$

We have that: $E[Y_t] = 0$

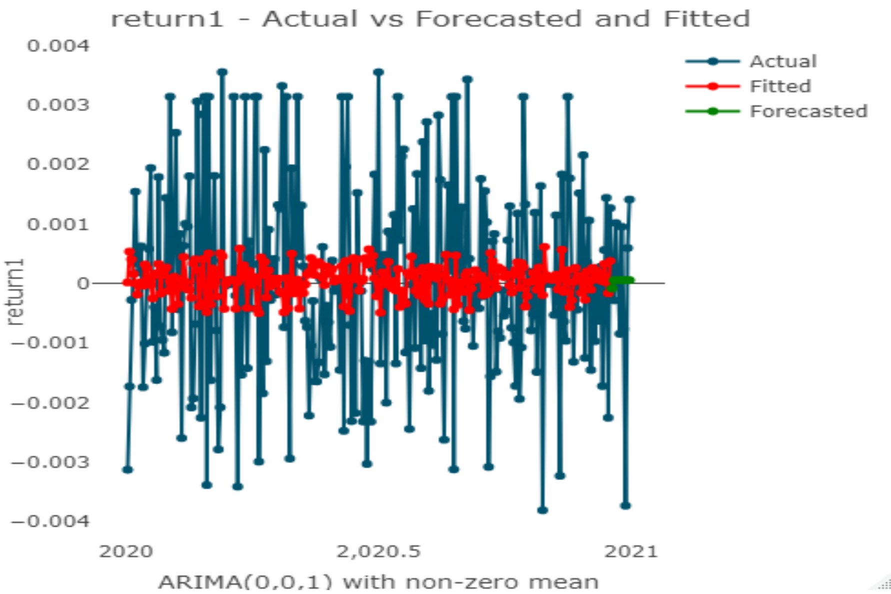
$Var[Y_t] = \sigma_u^2(1 + \Theta_1^2) = 0.000002516 *(1 + (-0.1546)^2) = 0.7148$

$Cov(Y_t, Y_{t-1}) = \Theta_1 \sigma^2 = 0.000002516*(-0.1546) = --3.889736e\text{-}7$

$$\rho_1 = \frac{\theta_1}{1+\theta_1^2} = \frac{-0.1546}{1+(-0.1546)\hat{}2} = -0.15098$$

Base on ARIMA(0,0,1) model, first, we conduct in-sample forecast to compare the difference between forecasted data and actual data (Figure 5). Then we have a judgment about the accuracy of forecasted model. The next steps, we proceed to forecasting the exchange rate up to 10 steps ahead (see Figure 6) Figure 5: The forecast of 10 steps ahead exchange rate by ARIMA(0,0,1) model

Figure 5. In-Sample Forecast

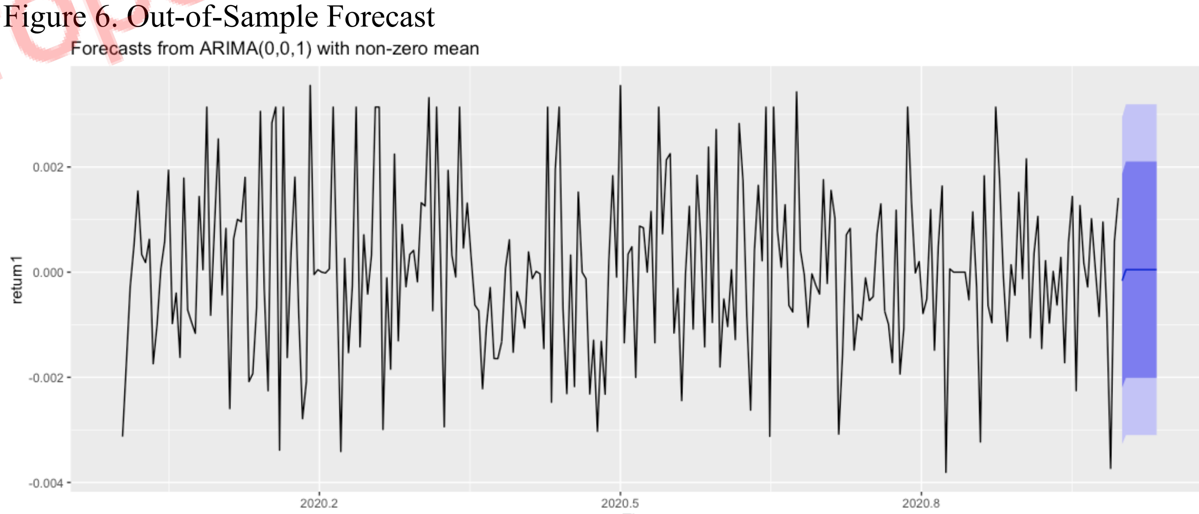


RMSE	0.001593
MAPE	Inf
MAE	0.001251
MASE	0.650244

Table 5: In-Sample Forecast

Steps Ahead	Forecasted Value	Actual Value
H + 1	-8,26E-05	1.722343e-04
H + 2	5,54E-05	-2.818533e-04
H + 3	5,54E-05	1.017429e-03
H + 4	5,54E-05	6.257724e-05
H + 5	5,54E-05	-8.451234e-04
H + 6	5,54E-05	9.546241e-04
H + 7	5,54E-05	-7.667572e-04
H + 8	5,54E-05	-3.732635e-03
H + 9	5,54E-05	5.969024e-04
H + 10	5,54E-05	1.412297e-03

Figure 6. Out-of-Sample Forecast



The final fit:

$\hat{y}_t = u_t - 1.5459e\text{-}01 * u_{t-1}$

The one-step-ahead forecast, $\hat{y}(1)$, is written as:

$\hat{y}_t(1) = 1.5459e\text{-}01 * [u_t] = 1.5459e\text{-}01 * u_t$

Similarly, two-step-ahead forecast $\hat{y}(1)$, is

$\hat{y}_t(2) = 1.5459e\text{-}01 * [u_{t+1}] = 1.5459e\text{-}01 * u_t(1)$

……

$\hat{y}_t(10) = 1.5459e\text{-}01 * [u_{t+9}] = 1.5459e\text{-}01 * u_t(9)$

Steps Ahead	**Forecasting**
H + 1	-1.62131E-04
H + 2	4.760441E-05
H + 3	4.760441E-05
H + 4	4.760441E-05
H + 5	4.760441E-05
H + 6	4.760441E-05
H + 7	4.760441E-05
H + 8	4.760441E-05
H + 9	4.760441E-05
H + 10	4.760441E-05

3.2 STRUCTURAL BREAK–ARIMA MODEL

The depreciation of USD in May 2021 led to the decline in the USD/CNY exchange rate. The exchange rate rapidly decreased from around 6.6 to below 6.3. This change suggested a structural break ARIMA model with a change point at 31-May 2021. There are some challenges when we face the structural break model. The original data clearly shows a sudden change; however, the return series is used to build the model instead of the original data, and this data does not clearly show any change point. Hence, we test for structural change detection involving an unknown breakpoint through the supF test. On the contrary, the result of this test shows the return has no structural change. (see Appendix 1)

3. 2 GARCH MODEL

Base on patterns appearing in residual from ARIMA model, a GARCH model could be suggested. First, we conduct the Lagrange Multiplier (LM) Test to test for conditional heteroscedasticity. We collect residuals from ARIMA model to test Arch effect. The result is showed in the table 4

Table 4. ARCH Effect Test Result

ARCH – LM Statistics	**Prob. Chi-square**
12.938	0.3735
Ho: There are no ARCH effect on residuals	

Unlike what we expect at previous part, the residuals show no ARCH effect. Hence, we conclude that univariate conditional volatility models - GARCH model is not apply for this empirical data. We use only ARIMA to forecast the exchange rate.

4. CONCLUSION

This study performs a comparison of various ARIMA models. When we conducted diagnostic checking, we found that the residuals chart indicates volatility, which suggests the GARCH model. We clarify this concern by using the LM test; however, contrary to our judgment, the residual shows no ARCH effect. Consequently, the ARIMA model is used as the main technique to forecast the exchange rate. On the process of the fitting model, we found that the return has no structural change and AR term. The model of ARIMA(0,0,1) is selected as the best model due to the lowest AIC.

We forecast ten steps ahead of the exchange rate. The forecast from the ARIMA model shows stable movement, which is our most concern. 2021 is a tough and challenging year under the impact of COVID-19. This pandemic caused lots of economic damages; hence, under the situation of lasting pandemic due to new virus variants, the financial market will continue to fluctuate significantly fluctuates. The exchange rate is also strongly affected by pandemics since the pandemic could restrict international trade between two countries. Similarly, when the trade balance of a country is in deficit, it leads to a change in the exchange rate. To conclude, the ARIMA model could not be the best approach to forecast the USD/CNY exchange rate. The future work should find other techniques and consider the combination of technical approaches and economic theories. Then, multivariate statistical models could be the choice.

REFERENCE

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the econometric society, 987-1007.

Cont, R. (2001) Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. Quantitative Finance, 1, 223-236. http://dx.doi.org/10.1080/713665670

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31(3), 307-327.

APPENDIX A

```
################################################################
##############Check Structural Break########################
library(strucchange)
plot.ts(return1)
model1 <- Fstats(return1~1, from = 0.0000001); model1
sctest(model1)
strucchange::breakpoints(return1~1)
```

```
> sctest(model1)

        supF test

data:  model1
sup.F = 4.854, p-value = 0.489

> strucchange::breakpoints(return1~1)

         Optimal 1-segment partition:

Call:
breakpoints.formula(formula = return1 ~ 1)

Breakpoints at observation number:
NA

Corresponding to breakdates:
NA
```

P value of 0.489 shows that we fail to reject null hypothesis of no structural change

APPENDIX B

```
################################################################################
##############################Check ARCH Effect###############################
resid <- arima3$residuals
resid1 <- resid^2
ArchTest(resid1)
```

```
> ArchTest(resid1)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  resid1
Chi-squared = 12.938, df = 12, p-value = 0.3735
```

APPENDIX C

Source Code could be found at https://github.com/thaohuongg/Applying-ARIMA-model-in-Forecast-Exchange-Rate.git