Final Project Report
Name: Sharon Zheng
Course: ECON 460 Economic Forecasting
Instructor: Bruce Hansen
Dec 3rd, 2020

Abstract:

In this project, I am going to forecast the trade balance between the U.S. and China for October 2020, for which the official data will be released on December 4th. The data for Trade Balance with China is not seasonally adjusted, and two seasonality adjustments are proposed and considered separately in the following sections. For the model with seasonality adjustment 1, the best option with the lowest AIC is picked in each step, and a Simple Best Model is thus constructed. For the model with seasonality adjustment 2, a Forecast Combination Model is proposed through the Granger-Ramanathan Combination.

In the construction of each model, multiple factors are considered, including the autocorrelation of the errors, the coefficient stability, the unbiasedness of the model, and the White Noise properties of the forecast error.

Finally, the two models are compared through the September realization. Due to the missing data of an indicator in the Simple Best Model, the final October forecast will be given through the Forecast Combination Model and an approximation of the Simple Best Model. And the period forecast will be made by the autoregression and seasonality factors of both models.

A. Dependent Variable Determination

The data of import from and export to China is retrieved from the Federal Reserve Economic Data (FRED), and the time series of trade balance and the absolute value of it is shown in Figure.1 and Figure.2, respectively.

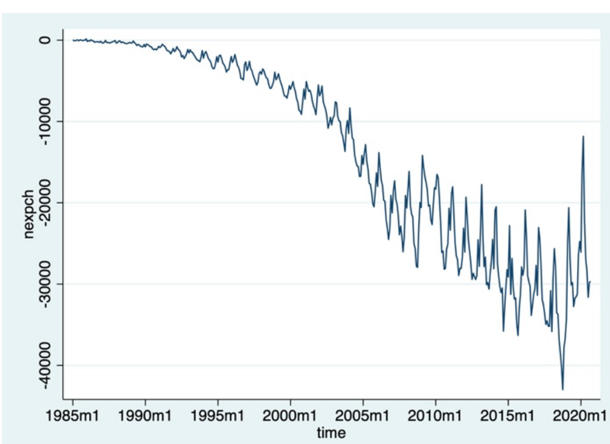


Figure.1 Trade Balance
Export - Import

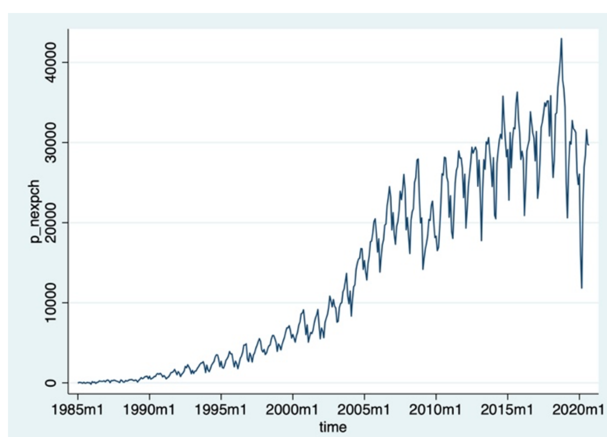


Figure.2 Absolute value of Trade Balance
|Export – Import|

Observing Figure.1, the trade balance is always negative, so taking the absolute value of it will not alter the integrity of the original data set. And in Figure.2, as the trend of the absolute value of trade balance does not show a linear pattern and is convex, the natural logarithm of the absolute value is introduced, and the natural logarithm-time graph is as in Figure.3.

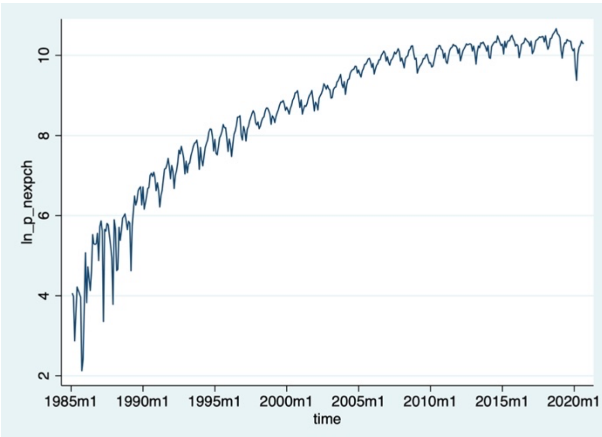


Figure.3 Natural Logarithm of Absolute Trade Balance
ln|Export – Import|

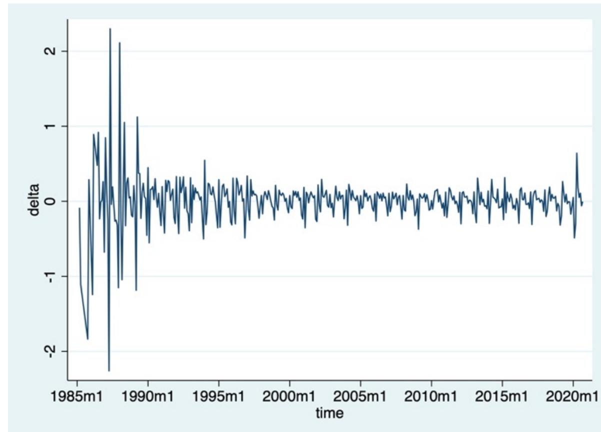


Figure.4 Differenced Natural Logarithm of Absolute Trade Balance
diff.ln|Export – Import|

After taking the logarithm, the data still does not show a linear pattern, rather, it turns from convex to concave.

The next step is to check the stationarity of data, and an Augmented Dickey-Fuller (ADF) test is performed on the logarithm of absolute value to check the existence of unit root. The test statistics is Z(t) = -2.845, which is larger than -3.130—the 10% critical value—and thus insignificant. As the unit root of the logarithm of the absolute value of trade balance is not rejected at a 90% confidence level, to get a stationary dependent variable, I take the differenced logarithm value for the next step and got Figure.4. And an ADF test on the differenced data gives me Z(t) = -11.883, which is smaller than the 1% critical value -3.447 and thus reject a unit root at a 99% confidence level.

Hence, after the mathematical manipulation and ADF test, I finally choose the differenced logarithm of trade balance's absolute value as my dependent variable and will denote it by ***y*** in the following sections.

B. Seasonality Adjustment

To check the consistency of the seasonality, regress ***y*** on 12-month dummy variables and test for consistency using rolling regression. The coefficients in the rolling process are shown in Figure.5 through Figure.10.

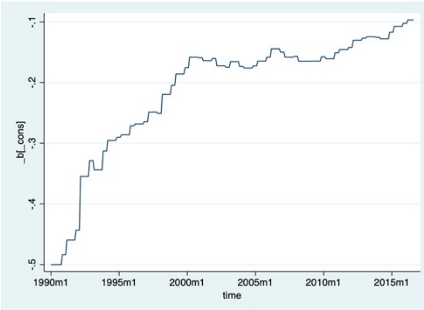


Figure.5 Rolling Regression—the constant

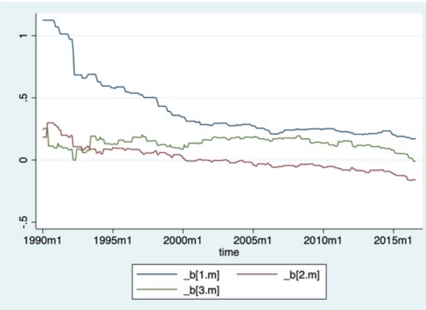


Figure.6 Rolling Regression—m1 to m3

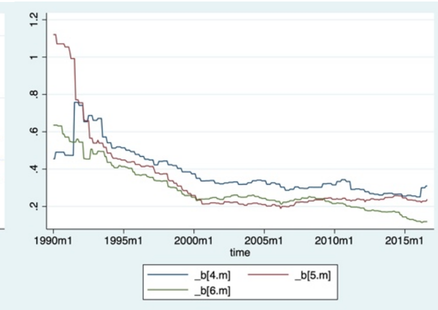


Figure.7 Rolling Regression—m4 to m6

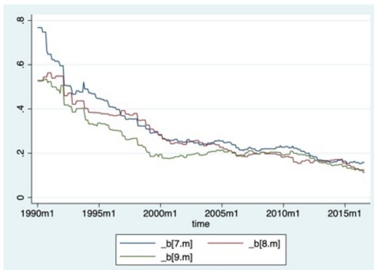


Figure.8 Rolling Regression—m7 to m9

Figure.9 Rolling Regression—m10 to m11

From the rolling regression result, it can be observed that almost all of the coefficients trend downward, but the range of changes differs among these regressors.

To deal with the inconsistency of coefficients on seasonal dummies, two breakpoints are introduced to adjust the seasonality. The first point is 1997m1, and the second is 2012m6, and the Chow test is conducted to check if these two breakpoints are significant from a statistics perspective. The results give F = 9.22 for the first point and F = 4.33 for the second point. Compared with the critical values of the QLR Statistic with 15% trimming, the test statistics show that the shifts at both points are significant at a 99% confidence level.

From the perspective of optimal model construction, the Akaike Information Criterion (AIC) is checked for two models (one with breakpoint only at 1997m1 and the other with both) as well. Adding two breakpoints strictly decrease the Akaike Information Criterion (AIC) of the model, from -576.51 to -645.28, and only including the first breakpoint will result in a slightly lower AIC estimation (-645.55). Figure.10 and Figure.11 compare the seasonality (fitted value in graphs) from 1995m6 to 2015m1 under these two adjustments, and both conditions will be considered in the following analysis.

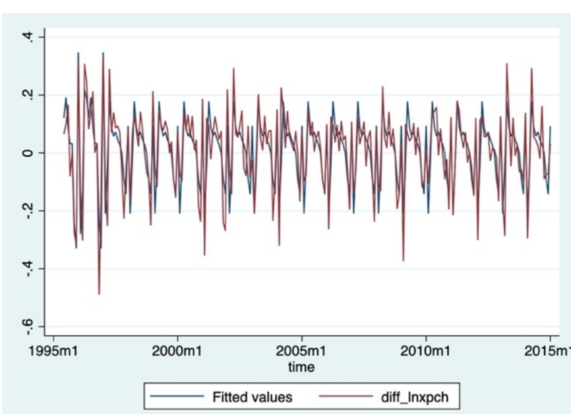


Figure.10 Seasonality adjustment at 1997m1

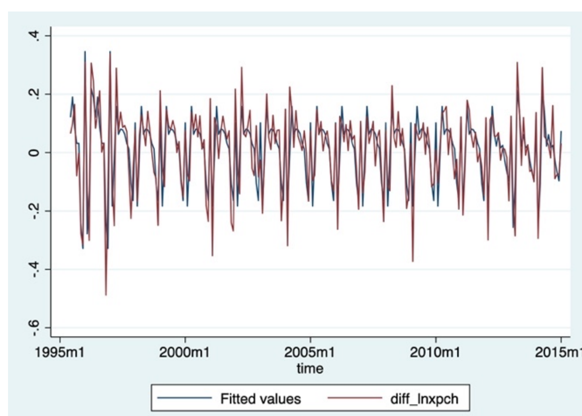


Figure.11 Seasonality adjustment at 1997m1 and 2012m6

C. Optimal Autoregression model of Seasonally Adjusted *y*

1. Data after Seasonality Adjustment at 1997m1 (seasonality adjustment 1)

The time series of the residuals after seasonality adjustment is shown in Figure.12. And to determine the optimal number of lags that should be involved in the autoregression, the autocorrelation of the seasonally adjusted data is checked first, and the result is shown in Figure.13.

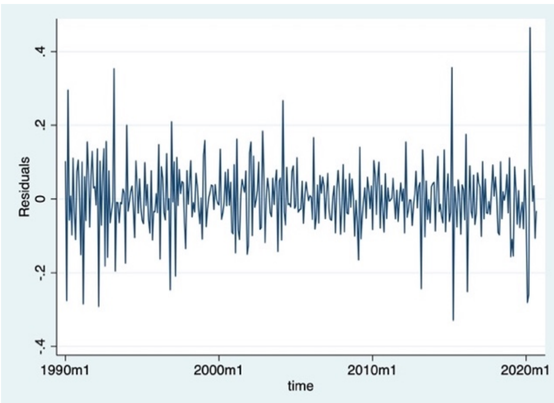


Figure.12 Residuals (seasonality adjustment 1)

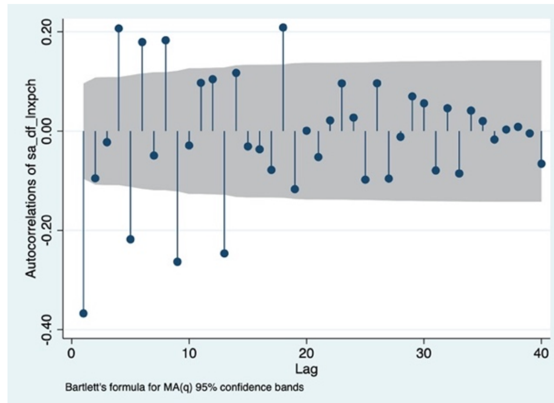


Figure.13 Autocorrelation of residuals (seasonality adjustment 1)

From Figure.13 we can see that the autocorrelation is either positive or negative, and there is no strong autocorrelation displayed in the data. As the last lag with statistically significant autocorrelation occurs slightly before the 20th lag, models with one through twenty lags are estimated and their AIC estimations are compared as in Table.14.

Model	*AIC*
AR (1)	-714.0313
AR (2)	-733.8533
AR (3)	-735.5756
AR (4)	-738.2226
AR (6)	-739.2948
AR (8)	-737.4656
AR (10)	-736.76
AR (12)	-738.6975
AR (14)	-740.0918
AR (16)	-742.9804
AR (18)	-740.6876
AR (20)	-737.9202

Table.14 AIC comparison for different AR models (seasonality adjustment 1)

The AR (16) model has the lowest AIC estimate of -742.98 and is picked for following analysis.

2. Data after Seasonality Adjustment at 1997m1&2012m6 (seasonality adjustment 2)
   Similarly, models with one through twenty lags are estimated and their AIC estimations are compared as in Table.15.

Model	*AIC*
AR (1)	-750.6192
AR (2)	-772.6182
AR (3)	-772.3469
AR (4)	-772.7382
AR (5)	-776.6488
AR (6)	-774.7043
AR (7)	-772.9519
AR (8)	-771.7185
AR (9)	-770.226
AR (10)	-770.6037
AR (11)	-772.2762
AR (12)	-771.8739

Table.15 AIC comparison for different AR models (seasonality adjustment 2)

The AR (5) model has the lowest AIC estimate of -776.65. Surprisingly, given the same sample size, even though the pure seasonality model with adjustment 2 has higher AIC than that with adjustment 1, after adding the autoregression terms to the model, comparing data in Table.14 and 15 we can see that, the seasonality adjustment 2 model behaves much better than the seasonality adjustment 1 model. Thus, in the following analysis, the models with AIC below -772.00—AR (2), AR (3), AR (4), AR (5), AR (6), AR (7) of seasonality adjustment 2 model—will be considered for forecast combination.

D. Granger Causality of indicators

For this target dependent variable ***y***, I picked nine possible indicators. Though each of them will be tested again to see the AIC in the following model construction process, I want to use the Granger Causality test to learn about the possible causal relationships between the potential indicators and ***y***. To keep consistency, the indicators are all:

- in the form of "ratio of the previous year (x/L.x)", and this manipulation can also avoid the unit root if there is any
- monthly (for the daily data, the monthly average is used)
- tested for seasonality, for those with correlated errors, Newey-West estimation is used
- seasonally adjusted if the seasonality pattern is significant

In the Granger causality test of each indicator, the Autoregressive Distributed Lag (ADL) model is used. $\boldsymbol{y}$ is regressed on its first 5 lags (AR (5) with seasonality adjustment 2 is the optimal combination as analyzed in the previous part) as well as the first twelve lags of the potential indicator, and the joint significance of the indicator's lags is estimated.

Denotation: Indicator/ FRED code	*F test statistics/P-value*
x_1:Spread between Baa Corporate Bond and 10-Year T-Bill/BAA10Y	2.13/0.0147(95%)
x_2: Spread between Baa Corporate Bond and Federal Funds/BAAFF	0.48/0.9283
x_3: Real Broad Effective Exchange Rate for China/RBCNBIS	1.57/0.1013 (approx. 90%)
x_4: China / U.S. Foreign Exchange Rate/DEXCHUS	2.60/0.0025(99%)
x_5: New Privately-Owned Housing Units Started/HOUST	1.54/0.1092(approx. 90%)
x_6: Manufacturers' New Orders: Consumer Goods/ACOGNO	2.61/0.2331
x_7: Producer Prices Index: Total Manufacturing/PIEAMP01USM661N	0.50/0.9172
x_8: Total Retail Trade in United States/USASARTMISMEI	2.16/0.0135(95%)
x_9: Spread between 10-Year T-Bill and 3-Month T-Bill/T10Y3M	2.71/0.0016(99%)

From the results above we can see that:

- The China / U.S. Foreign Exchange Rate and Spread between 10-Year T-Bill and 3-Month T-Bill has Granger causality on $\boldsymbol{y}$ at a 99% confidence level.
- The Spread between Baa Corporate Bond and 10-Year T-Bill and Total Retail Trade in the United States have Granger causality on $\boldsymbol{y}$ at a 95% confidence level.
- The Granger causality of the other four variables are rejected at a 90% confidence level, but the p-value for Real Broad Effective Exchange Rate for China and New Privately-Owned Housing Units Started is close to 0.1 thus should be paid attention to as well.

E. Single Best Model Analysis

If we insist on the AIC criterion and only pick the best outcome in seasonality adjustment (only one seasonality breakpoint at 1997m1), autoregression (AR (16)), and indicator lag test (the process is not discussed above), the following model will be constructed:

$$y_t = L(1/16).y + L(1/1).x_3 + L(1/1).x_8 + L(1/4).x_9 + e_t$$

$x_3 = \ Real\ Broad\ Effective\ Exchange\ Rate\ for\ China(ratio\ to\ the\ previous\ year),$
$x_8 = Total\ Retail\ Trade\ in\ United\ States(ratio\ to\ the\ previous\ year),$
$x_9 = \ Spread\ between\ 10 - Year\ Treasury\ Bill\ and\ 3 - Month\ Treasury\ Bill(ratio\ to\ the\ previous\ year)$
and Figure.16 represents the autocorrelation of the residuals:

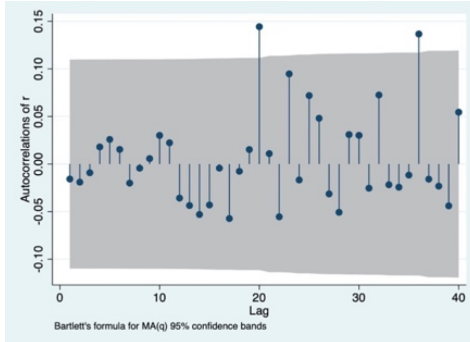


Figure.16 Autocorrelation of residuals for Single Best model

From the graph we can see that, the errors are not serially correlated, so the robust stand error is applicable in the following analysis.

1. Stability of coefficients on indicators & Adjustment

   Run a rolling test on the coefficients of indicators, and the result is shown in Figure.17 through Figure.21.

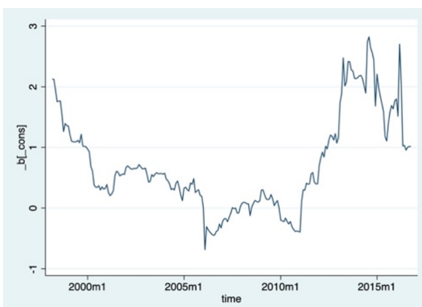


Figure.17 Constant

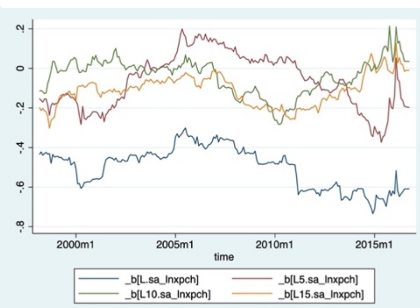


Figure.18 Coefficients on Autoregression

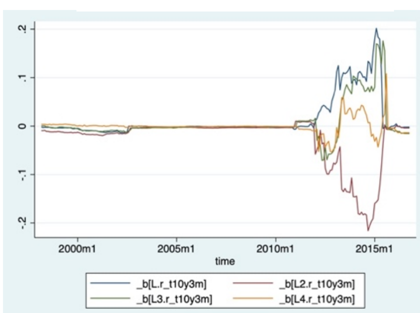


Figure.19 Coefficients on Spread between 10-Year T-Bill and 3-Month T-Bill

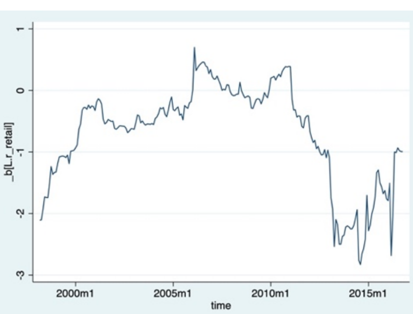


Figure.20 Coefficients on Total Retail Trade in United States

Figure.21 Coefficients on Real Broad Effective Exchange Rate for China

My observations are as follow:

a. Autoregression terms: coefficients fluctuate within a small range (around 0.2) and thus will be deemed as stable.
b. T-Bill spread: the coefficients of all four lags are approximately zero before 2010, and after 2010, there are bumps for coefficients on all of the four lags.
c. Total retail trade: before a time around 2013, the coefficient fluctuates around zero, but after that point, a plump happens and made the coefficient lower than -2.
d. Real broad effective exchange rate: before 2005m1, the coefficient fluctuates around some negative value, and after 2005m1, the coefficient starts to approach zero and then trend back to negative after a time around 2013.

Based on the observations above, a breakpoint at 2012m1 is tested using the Chow test for all of the three indicators and the AIC estimates are calculated as well. And the Chow test result shows that:

- The breakpoint for the T-Bill spread has an F-test statistic of 2.23, which fails to reject the non-existence of the breakpoint according to the QLR Statistic with 15% trimming.
- The breakpoint for the effective exchange rate has an F-test statistic of 2.18, which fails to reject the non-existence of the breakpoint according to the QLR Statistic with 15% trimming.
- The breakpoint for the total retail trade has an F-test statistic of 2.68, which fails to reject the non-existence of the breakpoint according to the QLR Statistic with 15% trimming.

Hence the Chow test result shows that the existence of this breakpoint is not valid in the statistics perspective. However, according to the AIC estimation, the breakpoint does decrease the AIC of the model:

Interacted indicator	*AIC*
None (original model)	-727.06
T-Bill spread (model A1)	-730.13
Effective exchange rate (model A2)	-728.29
Total retail trade (model A3)	-729.01
All three indicators (model A4)	-728.98

As is shown in the result above, involving the breakpoint 2012m1 for the T-Bill spread will give the optimal forecasting model among the four options. However, compared with the AIC of the original model, the lower AIC of model A2 A3 A4 confirms the assumption that some breakpoints for these indicators may make the model fit better, so I tested several more breakpoints. The 2005m1 breakpoint for the effective exchange rate and 2010m1 for total retail trade give me the best result among all tests, with the test statistics and AIC estimation shown as follows:

a. The breakpoint 2005m1 for the effective exchange rate has an F-test statistic of 3.45, which fails to reject the non-existence of the breakpoint according to the QLR Statistic with 15% trimming.
b. The breakpoint 2010m1 for the total retail trade has an F-test statistic of 2.66, which fails to reject the non-existence of the breakpoint according to the QLR Statistic with 15% trimming

Interacted indicator	*AIC*
None (original model)	-727.06
T-Bill spread (model A1)	-730.13
Effective exchange rate (model B2)	-729.92
Total retail trade (model B3)	-729.72
T-Bill spread & Effective exchange rate (model B4)	-730.40
T-Bill spread & Total retail trade (model B5)	-728.65
All three indicators (model B6)	-728.56

After the new adjustment, model B4 gives a lower AIC than the previous optimal model A1, thus become the analysis model for the following process.

2. Error check: unbiased and unforecastable
With the adjustments above, the following model is finally determined:

$$y_t = L(1/16).y + L(1/1).x_3 + L(1/1).x_8 + L(1/4).x_9 + d_1 * L(1/1).x_3 + d_2 * L(1/4).x_9 + e_t$$

$x_3 = \ Real\ Broad\ Effective\ Exchange\ Rate(ratio\ to\ the\ previous\ year)$,
$x_8 = Total\ Retail\ Trade\ in\ United\ States(ratio\ to\ the\ previous\ year)$,
$x_9 = \ Spread\ between\ 10 - Year\ Treasury\ Bill\ and\ 3 - Month\ Treasury\ Bill(ratio\ to\ the\ previous\ year)$,
$d_1 = time \geq\ 2005m1$,
$d_2 = time \geq 2012m1$

To evaluate the forecasts, I am going to check the unbiasedness of the model and see if the one-step-ahead error is unforecastable and thus White Noise.

a. Unbiasedness
As the fitted values we get from the model are the one-step-ahead forecasts for each period, the expectation of forecast errors is approximated by the mean value of the residuals:

- $\overline{residuals} = -9.84 * 10^{-11}$, quite close to zero

b. Unforecastable White Noise

Regress the following model:

$$y_{n+1} = \alpha + \beta y_{n+1|n} + e_{n+1}$$

and test for the joint $\alpha = 0$ *and* $\beta = 1$, the F-test statistic is 0 and the p-value is 1, failing to reject the null hypothesis. And if we convert ***y*** back to trade balance level, do the same regression and conduct the joint significance test, F-test statistic is 1.04 and the p-value is 0.36, failing to reject the joint insignificance of α *and* β as well. Hence, the forecast error is confirmed to be White Noise.

F. Forecast Combination model analysis

1. Pick combination components & determine weight

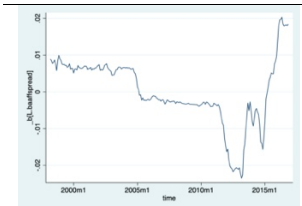
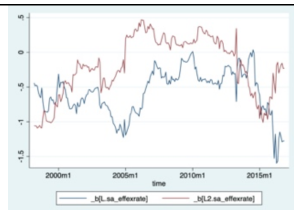
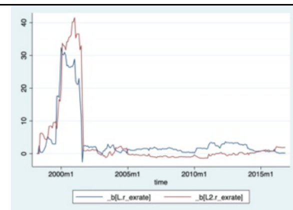
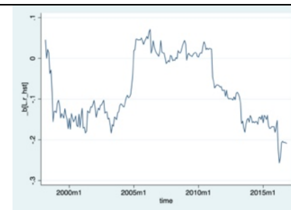
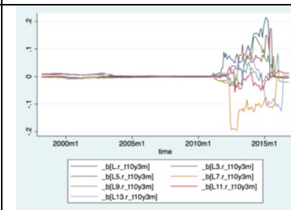
Though the test in the previous section gives the evidence of each potential indicators' predictive causality relationship with y, further steps need to be taken to decide whether they will contribute to the model construction. For the AR (p) model (p = 2, 3, 4, 5, 6, 7), estimate the following model of n = 1 through 14 for each potential indicator

$$y = L(1/p).y + L(1/n).x + e_t$$

and compare the AIC estimation within each indicator's estimations. The models that strictly improve the original model by giving a lower AIC than the simple autoregression model are recorded in the following table (in each cell is the number of lags for the indicator).

y	x_2 *(baaff)*	x_3 *(effexrate)*	x_4 *(exrate)*	x_5 *(hst)*	x_9 *(t10y3m)*
AR (2)		1,2	1,2,4,6,8,12,13	1,2	1 through 14
AR (3)	1	1,2	1,2,4,6,8,12,13	1,2,6,8	1 through 14
AR (4)	1	1,2	1,2,4,6,8,12,13,14	1,2,6,8	1 through 14
AR (5)		1,2	1,2,12	1,6,8	1 through 14
AR (6)		1,2	1,2,12,13,14	1	2,3,4,5,6,10,12,13,14
AR (7)		1,2	1,2,4,6,12,13,14	1	2 through 14

To make sure the combination test can capture the changes of indicators' effects on ***y*** over time, the coefficient stability test is firstly conducted for the five indicators. The stability of coefficients on x_2 is tested with an AR (4) model of y, while other indicators are tested in the AR (7) model of y. It is shown in the graphs that the effects of the exchange rate and T-bill spread for 10-year and 3-month maturity on $\boldsymbol{y}$ is significant in some specific period, while approximately zero in other periods.

x_2 *(baaff)*	x_3 *(effexrate)*	x_4 *(exrate)*	x_5 *(hst)*	x_9 *(t10y3m)*
				

I have tried several breakpoints for the exchange rate, but they neither pass the Chow test nor decrease the AIC, so in the following process the exchange rate will be tested without adjustment. For the T-Bill spread, a breakpoint at 2012m1 passes the Chow test (with p-value = 0.023) and decreases the AIC by 0.8, and hence will be included in the following analysis.

I pick 1995m4 to 2014m12 as the in-sample period and regress the actual value on out-of-sample forecasts for the period after 2015m1. With each number in the cell representing a model, the Granger-Ramanathan Combination test is conducted through the following steps:

- After combination test 1: df = 46 (other indicators are omitted), indicators remain to be tested:

y	x_2 *(baaff)*	x_3 *(effexrate)*	x_4 *(exrate)*	x_5 *(hst)*	x_9 *(t10y3m)*
AR (2)		2	2,6,8,13	1,2	3,5,6,7,8,12,14
AR (3)	1		4,13	2,8	1
AR (4)			1,8,13,14	6,8	10
AR (5)				6,8	2,12
AR (6)		2	2,13		6,13
AR (7)		1	4,6,12,13	1	2,5,9,14

- After 6 other consecutive combination tests, the following models have positive coefficients:
  - AR (2) with 2 lags on x3 (effective exchange rate) has the weight of 0.0446
  - AR (3) with 1 lag on x9 (10-year & 3-month T-Bill spread) has the weight of 0.1290
  - AR (6) with 2 lags on x4 (exchange rate) has the weight of 0.8264
  - RMSE = 0.1010, lower than the separate models

2. Error check: unbiased and unforecastable

   With the adjustments above, the following model is finally determined:

$$y_t = 0.0446 * y_1 + 0.1290 * y_2 + 0.8264 * y_3$$
$$y_1 = L(1/2).y + L(1/2).x_2 + e_t$$
$$y_2 = L(1/3).y + L.x_9 + d_3 * L.x_9 + e_t$$
$$y_3 = L(1/6).y + L(1/2).x_3 + e_t$$

$x_2 = Spread\ between\ BAA\ and\ Federal\ Funds(ratio\ to\ the\ previous\ year),$

$x_3 = Real\ Broad\ Effective\ Exchange\ Rate\ for\ China(ratio\ to\ the\ previous\ year),$

$x_9 = Spread\ between\ 10-Year\ Treasury\ Bill\ and\ 3-Month\ Treasury\ Bill(ratio\ to\ the\ previous\ year)$

   To evaluate the forecasts, the unbiasedness of the model and the unforecastable one-step-ahead error are going to be checked.

   a. Unbiasedness

      As the fitted values we get from the model are the one-step-ahead forecasts for each period, the expectation of forecast errors is approximated by the mean value of the residuals:

      - $\overline{residuals} = -0.00018$, quite close to zero

   b. Unforecastable White Noise

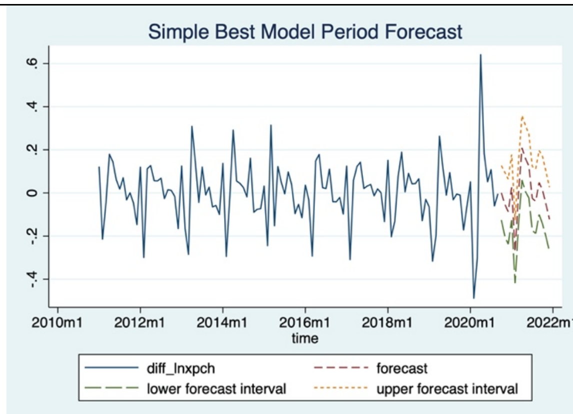
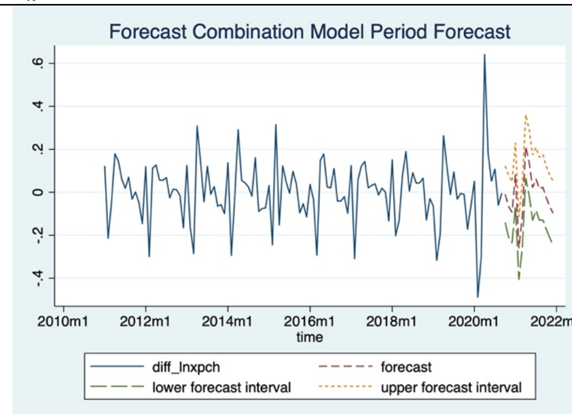
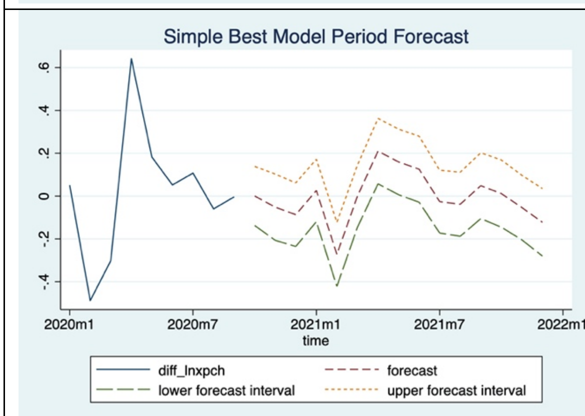
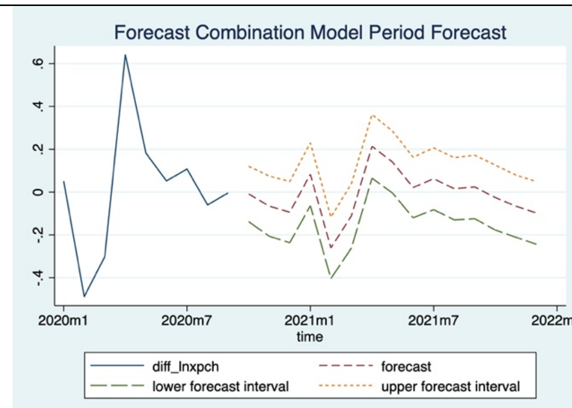
      Regress the following model:

$$y_{n+1} = \alpha + \beta y_{n+1|n} + e_{n+1}$$

   and test for the joint $\alpha = 0\ and\ \beta = 1$, the F-test statistic is 0 and the p-value is 1, failing to reject the null hypothesis. And if we convert ***y*** back to trade balance level, do the same regression and conduct the joint significance test, F-test statistic is 0 and the p-value is 0.9984, failing to reject the joint insignificance of $\alpha\ and\ \beta$ as well. Hence, the forecast error is confirmed to be White Noise.

G. Period forecast based on Autoregression and Seasonality

Because the future values of the indicators are unknown and unpredictable based on the current models, the forecast for next 12 month will be roughly constructed based on the autoregression and seasonality.

<table>
<tr><td></td><td>The Best Simple Model without Indicators</td><td>The Forecast Combination Model without Indicators</td></tr>
<tr><td>The model</td><td>$$y_{t+1|t} = \sum D_{it} + d_{s1} * \sum D_{it} + L(1/16).y_t + e_t$$
$d_{s1} = time \leq\ 1997m1,$
$D_{it}: Month\ dummies$</td><td>$$y_{t+1|t} = \sum D_{it} + d_{s1} * \sum D_{it} + d_{s2} * \sum D_{it} + L(1/6).y + e_t$$
$d_{s1} = time \leq\ 1997m1,$
$d_{s2} = time \geq 2012m6,$
$D_{it}: Month\ dummies$</td></tr>
<tr><td rowspan="2">The graph of period forecast</td><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Conclusion</td><td colspan="2">The 12-step-ahead forecasts of both models show a similar trend, including two crests at around 2021m1, 2021m5, and one trough at 2021m2. Two main differences are noted:
1. After the crest at 2021m5, the downward trend of the forecast from the Best Simple Model is flatter than that of the other model.
2. The fluctuation at 2021m7 seems more violent in the Best Simple Model.</td></tr>
</table>

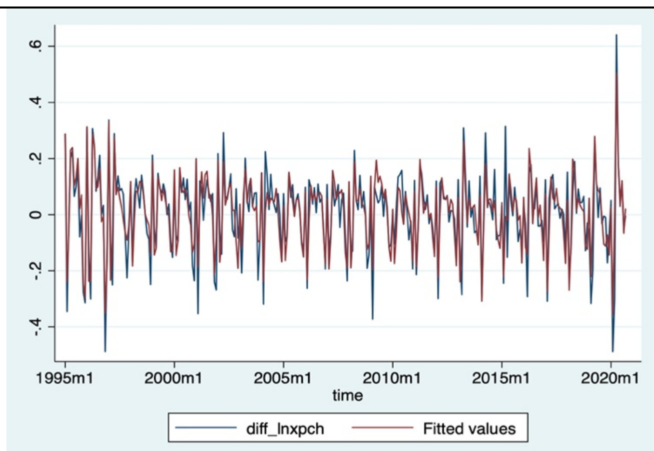
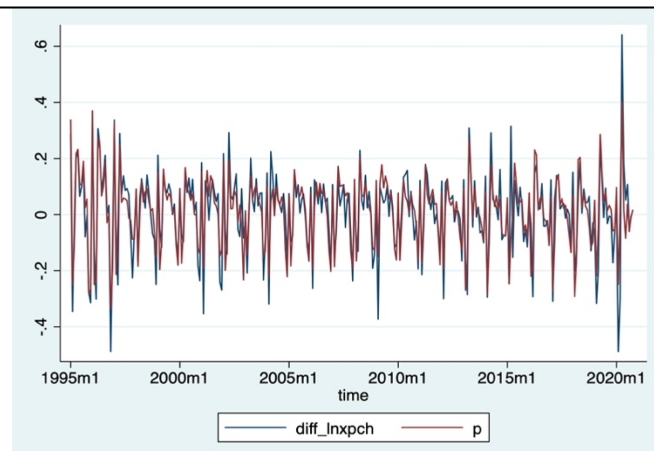
H. Compare the two models by September forecast (released)

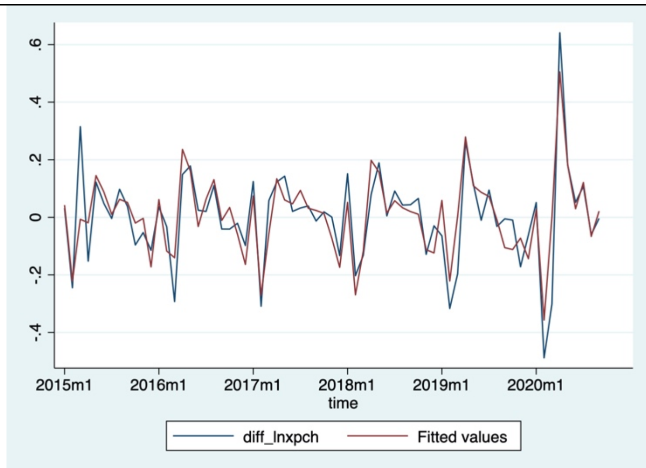
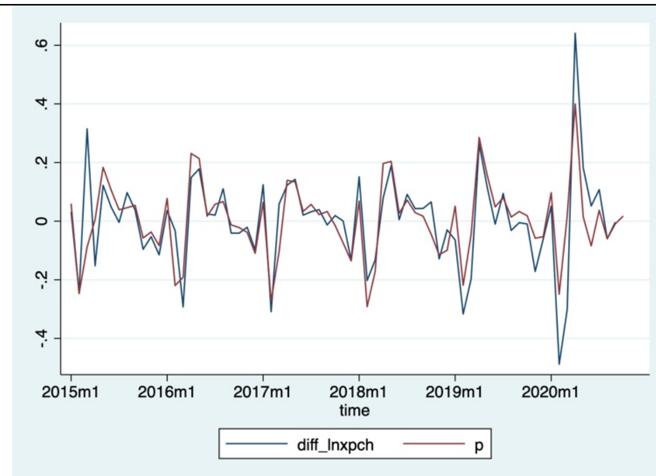
In the Simple Best Model, the indicator Total Retail Trade is only updated to August 2020, thus an October forecast with desired precision is inaccessible through this model. To compare the forecast of these two models, the September realization, which has been released in November, will be checked.

For the calculation process, because the fitted value $\hat{y}$ is the differenced natural logarithm of the absolute value of trade balance, the function of $\hat{y}$ for the trade balance next year is:

$$\widehat{NX_{n+1|n}} = NX_n * \exp\left(\widehat{y_{n+1|n}}\right)$$

Add seasonality adjustment 1 and 2 to the Simple Best Model and Forecast Combination Model accordingly, and the summary for each model is as follows:

	Simple Best Model w/ seasonality adjustment 1 (September Evaluation)	*Forecast Combination Model w/ seasonality adjustment 2 (September Evaluation)*
The model	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + L(1/16).y_t + L(1/1).x_3 + L(1/1).x_8 + L(1/4).x_9 + d_3 * L(1/1).x_3 + d_9 * L(1/4).x_9 + d_3 + d_9 + e_t$ $x_3 =$ *Real Broad Effective Exchange Rate (ratio to the previous year),* $x_8 =$ *Total Retail Trade in United States (ratio to the previous year),* $x_9 =$ *Spread between 10 − Year Treasury Bill and 3 − Month Treasury Bill (ratio to the previous year),* $d_{s1} = time \leq 1997m1,$ $d_3 = time \geq 2005m1,$ $d_9 = time \geq 2012m1,$ D_{it}: *Month dummies*	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + d_{s2} * \sum D_{it} + 0.0446 * y_1 + 0.1290 * y_2 + 0.8264 * y_3$ $y_1 = L(1/2).y + L(1/2).x_2 + e_t$ $y_2 = L(1/3).y + L.x_9 + d_9 * L.x_9 + d_9 + e_t$ $y_3 = L(1/6).y + L(1/2).x_3 + e_t$ $x_2 =$ *Spread between BAA and Federal Funds (ratio to the previous year* $x_3 =$ *Real Broad Effective Exchange Rate for China (ratio to the previous year),* $x_9 =$ *Spread between 10 − Year Treasury Bill and 3 − Month Treasury Bill (ratio to the previous year),* $d_{s1} = time \leq 1997m1,$ $d_{s2} = time \geq 2012m6,$ $d_9 = time \geq 2012m1,$ D_{it}: *Month dummies*
Graph for the complete sample		

Graph for time after 2015m1		
RMSE of the sample before 2020m9	**0.0685**	**0.0723**
*Point forecast for **y** (differenced natural logarithm)*	**0.0215182**	**-.0095017**
Point forecast for September trade balance (Millions of Dollars)	**-30428.13**	**-29498.75**
95% conf. interval	For ***y***: **(-0.1175524, 0.1605888)** For September trade balance (Millions of Dollars): **(-34968.17, -26477.55)**	For ***y***: **(-0.1284352, 0.1094318)** For September trade balance (Millions of Dollars): **(-33224.29, -26190.96)**
September released data (Millions of Dollars)	**-29671.58** (fall in the 95% forecast interval of both models)	
Forecast error (Millions of Dollars &Percentage)	**756.55 (-2.55%)**	**-172.83 (0.58%)**
Conclusion	For the September forecast, the Forecast Combination model is doing a better job than the Simple Best forecast model. But overall, for the whole range of the sample, the combination model has a higher RMSE than the simple best model. And from the graph, we can see that, though these two models both show to be "conservative" at making predictions for the extreme points, the simple model fits better at these points than the combination model.	

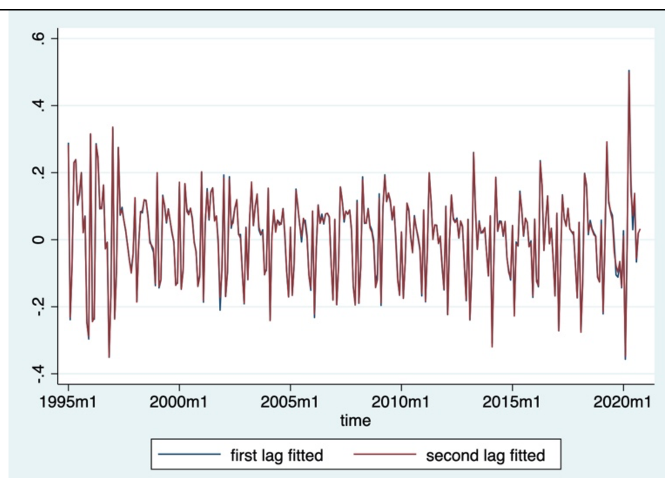
I. Forecast for the October trade balance (unreleased till Dec 4th)

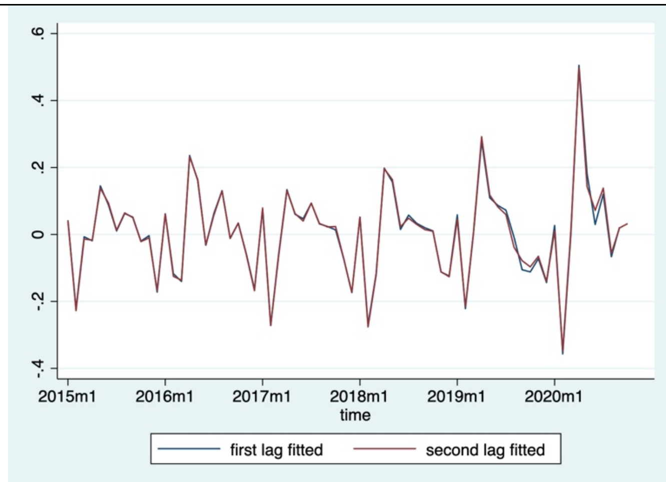
The official data for the October trade balance between the U.S. and China will be released on Dec 4th. From the comparison above, to be more rigorous, firstly the Granger-Ramanathan

Combination test should also be conducted for the models with seasonality adjustment 1, and secondly, the two combination models of both seasonality adjustment 1 and 2 should go through a Granger-Ramanathan Combination again so that the weight of each model can be determined. A new combination model of lower RMSE than either of them thus can be constructed. But due to the limitation on time, I am going to skip that step and go directly forecast with the Forecast Combination model (with seasonality adjustment 2) proposed above and an approximation to the Simple Best model (with seasonality adjustment 1).

1. The approximation to the Simple Best Model

   Among all indicators, only the data for Total Retail of September is not available, so a new model with the second lag of this indicator is estimated and compared with the original model:

	Original Model	*Approximated Model*
The model	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + L(1/16).y_t + L(1/1).x_3 + L(1/1).x_8 + L(1/4).x_9 + d_3 * L(1/1).x_3 + d_9 * L(1/4).x_9 + d_3 + d_9 + e_t$	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + L(1/16).y_t + L(1/1).x_3 + L(2/2).x_8 + L(1/4).x_9 + d_3 * L(1/1).x_3 + d_9 * L(1/4).x_9 + d_3 + d_9 + e_t$
	$x_3 =$ *Real Broad Effective Exchange Rate(ratio to the previous year)*, $x_8 =$ *Total Retail Trade in United States(ratio to the previous year)*, $x_9 =$ *Spread between 10 − Year Treasury Bill and 3 − Month Treasury Bill(ratio to the previous year)*, $d_{s1} = time \leq 1997m1$, $d_3 = time \geq 2005m1$, $d_9 = time \geq 2012m1$, D_{it}: *Month dummies*	
Graphical Comparison on the whole sample range		

Graphical Comparison for time after 2015m1		
AIC Comparison	-700.0551	-696.4751
RMSE Comparison	0.0685	0.0689
Error Check	Model: Unbiased Forecast Error: White Noise	Model: Unbiased Forecast Error: White Noise

In the graphs and statistics, though the approximated model has higher AIC and RMSE, the difference is acceptable, and thus the approximated model will be used in the following forecast for the October data.

2. The forecast

	Approximated Simple Best Model (October Evaluation)	*Forecast Combination Model (October Evaluation)*
The model	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + L(1/16).y_t + L(1/1).x_3 + L(2/2).x_8 + L(1/4).x_9 + d_3 * L(1/1).x_3 + d_9 * L(1/4).x_9 + d_3 + d_9 + e_t$ $x_3 = Real\ Broad\ Effective\ Exchange\ Rate(ratio\ to\ the\ previous\ year),$ $x_8 = Total\ Retail\ Trade\ in\ United\ States(ratio\ to\ the\ previous\ year),$ $x_9 = Spread\ between\ 10-Year\ Treasury\ Bill\ and\ 3-Month\ Treasury\ Bill(ratio\ to\ the\ previous\ year),$ $d_{s1} = time \leq 1997m1,$ $d_3 = time \geq 2005m1,$ $d_9 = time \geq 2012m1,$ $D_{it}: Month\ dummies$	$y_{t+1\|t} = \sum D_{it} + d_{s1} * \sum D_{it} + d_{s2} * \sum D_{it} + 0.0446 * y_1 + 0.1290 * y_2 + 0.8264 * y_3$ $y_1 = L(1/2).y + L(1/2).x_2 + e_t$ $y_2 = L(1/3).y + L.x_9 + d_9 * L.x_9 + e_t$ $y_3 = L(1/6).y + L(1/2).x_3 + e_t$ $x_2 = Spread\ between\ BAA\ and\ Federal\ Funds(ratio\ to\ the\ previous\ year),$ $x_3 = Real\ Broad\ Effective\ Exchange\ Rate\ for\ China\ (ratio\ to\ the\ previous\ year),$ $x_9 = Spread\ between\ 10-Year\ Treasury\ Bill\ and\ 3-Month\ Treasury\ Bill(ratio\ to\ the\ previous\ year)$ $d_{s1} = time \leq 1997m1,$ $d_{s2} = time \geq 2012m6,$ $d_9 = time \geq 2012m1,$ $D_{it}: Month\ dummies$
RMSE of the sample before 2020m10	**0.0689**	**0.0722**

*Point forecast for **y** (differenced natural logarithm)*	**0.0323822**	**0.017654**
Point forecast for October trade balance (Millions of Dollars)	**-30648.14**	**-30200.05**
95% conf. interval	For ***y***: **(-0.1130254, 0.1777898)** For October trade balance (Millions of Dollars): **(-35444.90, -26500.52)**	For ***y***: **(-0.101115, 0.136423)** For October trade balance (Millions of Dollars): **(-34008.57, -26818.04)**
October released data (Millions of Dollars)	**-30105.01** (fall in the 95% forecast interval of both models)	
Forecast error (Millions of Dollars &Percentage)	**548.13 (-1.82%)**	**95.04 (-0.32%)**
Summary	The released data falls in the 95% confidence intervals of forecasts from both models, and the two point forecasts deviate in the same direction from the realized value. What's more, to compare these two models, the forecast error and its percentage of the realized value are computed above. From the results we can see that, though the Forecast Combination Model has a slightly higher RMSE than the approximated Simple Best Model, it does a better job again for this one-step-ahead forecast. For a decision-maker who is uncertain about the precision of these two models, taking the average of these two forecasts gives an acceptable result (a forecast error of 321.59 and an error percentage of -1.07%). With the forecast evaluation of September and October released data, however, the Forecast Combination Model is overperforming the Simple Best Model for the recent periods, and the decision-maker can either focus more on this model or conduct a Granger-Ramanathan Combination process again on these two models and generate a new combined model with higher reliability.	