Business Forecasting

A Forecasting Study of United Kingdom and Canada

Table of Contents

Evaluating Exchange Rate Models and PPP: A Forecasting Study of United Kingdom and Canada

1. Introduction

The following report aims to evaluate the validity of Purchasing Power Parity **(PPP)** using data from two countries. We selected the **United Kingdom** as the source country and **Canada** as the foreign country. The analysis covers **10 years of monthly data**, focusing on the nominal exchange rate, the real exchange rate, and the consumer price index (CPI) for both countries. Our main objective is to test the absolute and relative forms of PPP, model the real exchange rate, and forecast future values. In this analysis, we not only test the validity of PPP but also forecast the real exchange rate using an ARIMA model.

2. Data and Methodology

To do this, we collected data from the OECD **Data Explorer for the Consumer Price Index** (**CPI**) and from the **UK Office for National Statistics (ONS)** for the nominal exchange rate over the past 10 years **(from 2015 to 2025)**, using 2015 as the base year. The dataset includes the following variables:

- **Nominal Exchange Rate (E)**: The exchange rate at which one currency can be exchanged for another.
- **Real Exchange Rate (R)** : Adjusted for inflation, it reflects the purchasing power between both currencies.
- **Consumer Price Index (CPI) - Both countries**: Measures the change in the price levels of a basket of goods and services in each country.

All data manipulation, statistical testing, and forecasting were performed in R using the **forecast**, **tseries**, and **ggplot2** packages. For all hypothesis tests—including the Augmented Dickey-Fill test for the ADF, cointegration tests, and regression diagnostics—we used a 5% significance level as our general cutoff for rejecting or not rejecting the null hypothesis.

We used the following major formulas:
Absolute PPP was tested based on stationarity of the log real exchange rate:

$$log\left(R_t\right) = log\left(\frac{E_t \cdot P_t^*}{P_t}\right)$$

Where $E_t$is the nominal exchange rate, P_t is the domestic CPI, and P_t^* is the foreign CPI.
Relative PPP was tested using two regression approaches:

- **CPI ratio model**:

$$log\left(E_t\right) = \alpha + \beta. log\left(\frac{P_t}{P_t^*}\right) + \varepsilon_t$$

- **Separate CPI variables model**:

$$log\left(E_t\right) = \alpha + \beta_1. log\left(P_t\right) + \beta_2. log\left(P_t^*\right) + \varepsilon_t$$

Both models were checked for stationarity of the residuals to evaluate the presence of a long-run equilibrium relationship.

3. Data Exploratory Analysis

Before starting the analysing process, the summary statistics of the raw data were examined to understand the distribution and scale of each variable. As seen in the following summary table (**Figure 1),** the CPI home and CPI Foreign had ranges between 99.2 and 135.1, and 98.21 and 128.07, respectively, with average values of approximately 112.3 and 110.95. The Real Exchange Rate and Nominal Exchange Rate displayed similar trends, with nominal rates ranging from 1.489 to 2.054.

Figure 1. Summary Before Transformation

```
   CPI_Home        CPI_Foreign      Real_Exchange_Rate Nominal_Exchange_Rate
 Min.   : 99.2   Min.   : 98.21   Min.   :1.489      Min.   :1.509
 1st Qu.:103.5   1st Qu.:103.11   1st Qu.:1.645      1st Qu.:1.679
 Median :108.5   Median :107.93   Median :1.699      Median :1.714
 Mean   :112.3   Mean   :110.95   Mean   :1.716      Mean   :1.733
 3rd Qu.:121.2   3rd Qu.:120.65   3rd Qu.:1.736      3rd Qu.:1.757
 Max.   :135.1   Max.   :128.07   Max.   :2.054      Max.   :2.053
```

To stabilize the variance and make data suitable for time series modeling, the data were transformed using the natural logarithm using the following formula for this step:

Log transformation of variables: log(E), log(R), log(CPI_home), log(CPI_foreign)

After the log transformation, another summary statistics (**Figure 2**) was executed to see if the variables were CPI home and CPI Foreign Converted properly into a log scale, resulting in a range between 4.597 and 4.906, with the mean increasing slightly to 4.716. Similarly, the Real Exchange Rate and Nominal Exchange Rate also went through log transformation, showing a more compressed range (from 0.3980 to 0.7196 for the real exchange rate, and from 0.4116 to 0.7194 for the nominal exchange rate).

Figure 2. Summary After Transformation

```
 CPIHome_logts    CPIForeign_logts RealExchangeRate_logts Nominalexchange_logts
 Min.   :4.597    Min.   :4.587    Min.   :0.3980         Min.   :0.4116
 1st Qu.:4.640    1st Qu.:4.636    1st Qu.:0.4977         1st Qu.:0.5180
 Median :4.687    Median :4.681    Median :0.5299         Median :0.5387
 Mean   :4.716    Mean   :4.706    Mean   :0.5378         Mean   :0.5483
 3rd Qu.:4.797    3rd Qu.:4.793    3rd Qu.:0.5518         3rd Qu.:0.5638
 Max.   :4.906    Max.   :4.853    Max.   :0.7196         Max.   :0.7194
```

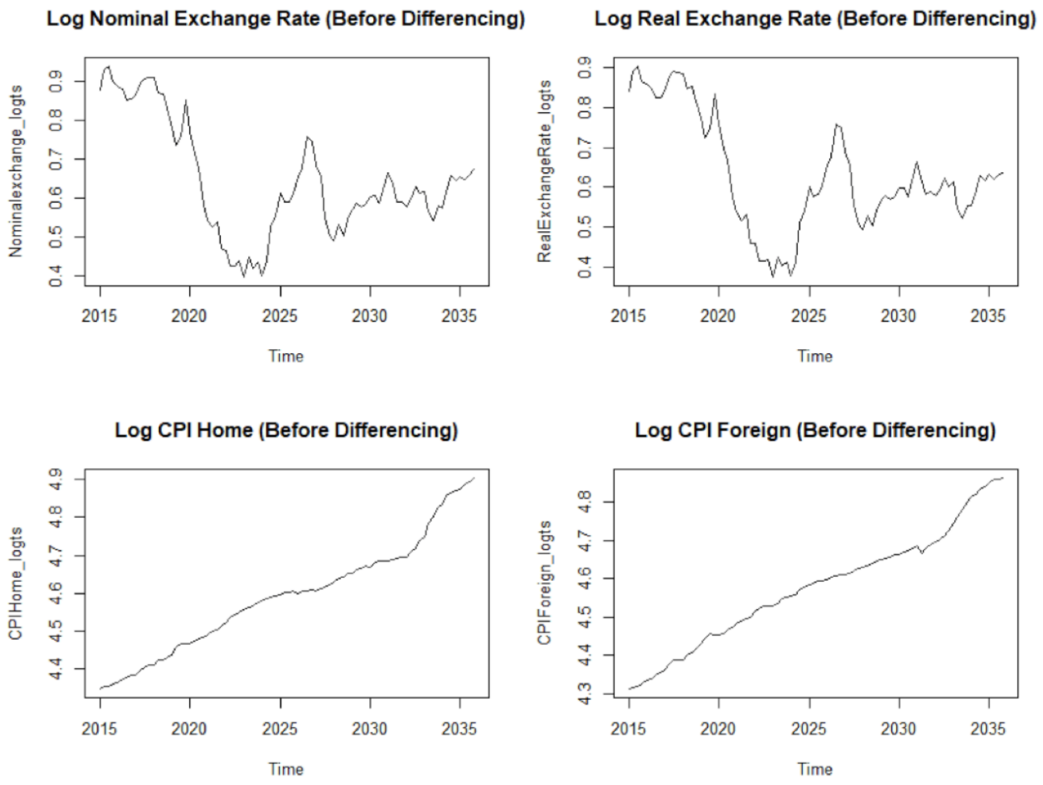
After the variables were transformed, their stationarity was analyzed using the **Augmented Dickey-Fuller (ADF) test**, which is used for detecting the presence of unit roots in time series data. The results (**Figure 3**) revealed that both nominal and real exchange rates, as well as the

CPI data, exhibited non-stationary **(p-value > 0.05)** behavior before differencing indicating that the variable had significant trends **(Figure 4)** that could lead to bias in the models.

Figure 3. ADF Test Results Summary

Variable	ADF Statistic	Lag Order	p-value	Stationary
Log CPI Home	-1.3769	4	0.8354	No
Log CPI Foreign	-2.0359	4	0.5616	No
Log Real Exchange Rate	-2.7148	4	0.2797	No
Log Nominal Exchange Rate	-2.4105	4	0.406	No

Figure 4. Visualizations of the series before differencing

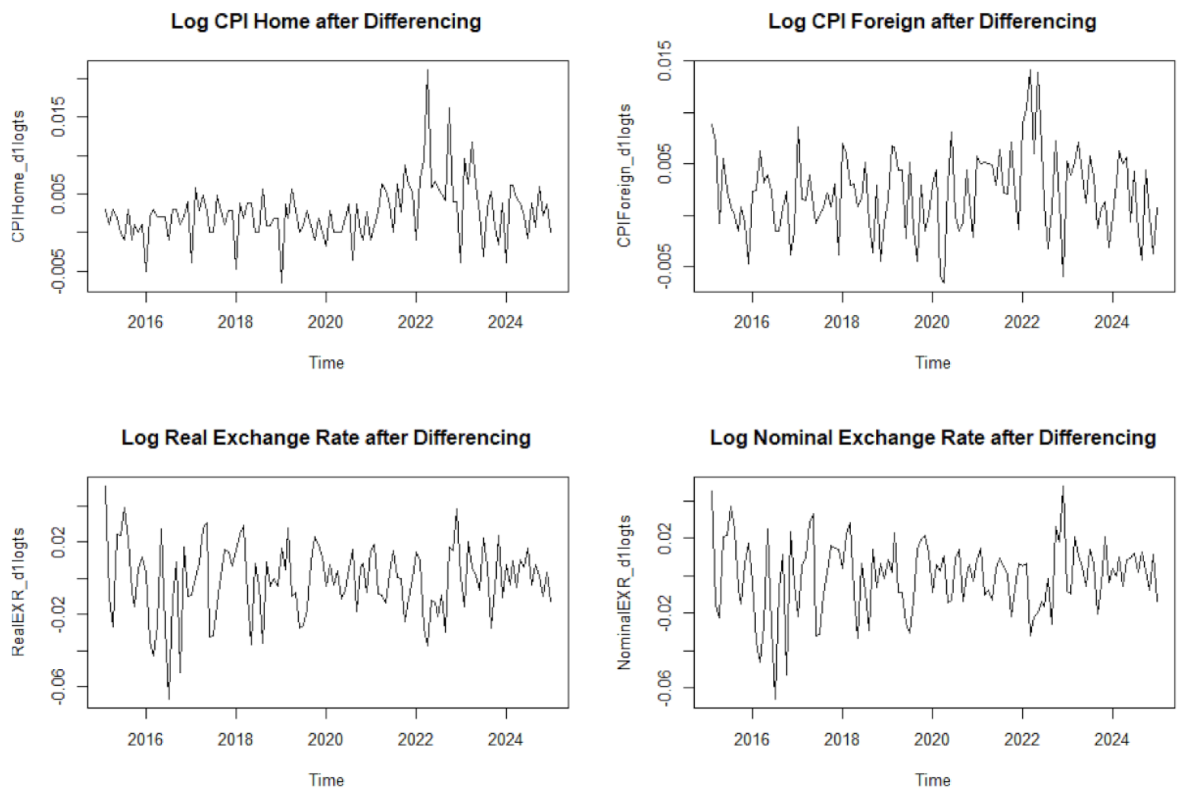


In that case, it was decided to perform first-order differencing on both series to remove these trends, a technique that subtracted each value in our times series from its preceding value to help to eliminate long-term trends. After differencing (**Figure 4**), the transformed time-series were tested once again to see if they were now stationary (**p-value < 0.05)** succeeding the expectations of this stage, making them suitable for time series modeling (**Figure 5**).

Figure 5. ADF Test Results After Differencing

Variable	ADF Statistic	Lag Order	p-value	Stationary
Log CPI Home (Differenced)	-3.4691	4	0.04805	Yes
Log CPI Foreign (Differenced)	-4.4842	4	0.01	Yes
Log Real Exchange Rate (Differenced)	-4.4724	4	0.01	Yes
Log Nominal Exchange Rate (Differenced)	-4.5393	4	0.01	Yes

Figure 6. Visualizations of the series after differencing



4. Testing the Validity of Purchasing Power Parity (PPP)

After confirming that the stationarity of the series was transformed through the ADF test, the following stage was to proceed to the test **Purchasing Power Parity (PPP)** hypothesis. The validity of the PPP is key in exchange rate modeling as it helps to determine whether exchange rates are aligned with relative price levels between countries. A **unit root test on the log-transformed real exchange rate** was conducted as the following step. The rationale behind is that, under absolute PPP, the real exchange rate should be stationary, meaning that any deviations from parity are temporary and revert to equilibrium over time.

Testing Absolute PPP: Unit Root Test on Log Real Exchange Rate

To test the validity of absolute Purchasing Power Parity, we conduct a unit root test on the log-transformed real exchange rate. The formula below indicates that under absolute PPP, the real exchange rate should be stationary, showing that any variation from parity is temporary and revert to equilinrium in the long term.

Figure 7: Formula for log of real exchange rate

$$r_s = s - (p - p_t)$$

Where:

- r_s is the log of the real exchange rate,
- s is the nominal exchange rate,
- p is the domestic price level,
- p_t is the foreign price level.

Using the Augmented Dickey-Fuller (ADF) test, the result of Dickey-Fuller statistics is -2,72 with a p-value of 0.28 (Figure 3). Since the p-value is greater than 0.05, we fail to reject the null hypothesis of a unit root. We then conclude that the real exchange rate is non-stationary. This result indicates that variations from absolute PPP are stable over time without reverting to a mean level. In essence, the exchange rate does not respond rapidly enough to equalize prices across countries, which goes against the expectations set by absolute PPP.

Testing Relative PPP: Regression Between Log Nominal Exchange Rate and Log CPI Ratio

While testing validity of absolute PPP assumes that real exchange rates remain unchanged, relative PPP emphasizes the relationship between changes in nominal exchange rate and changes in price ratio (between domestic and foreign consumer indexes). To assess this, we conduct an Ordinary Least Square (OLS) regression between the log-transformed nominal exchange rate and log-transformed consumer price index ratio (CPI). The ratio is calculated by dividing the domestic Consumer Price Index (CPI) by the foreign CPI.

Figure 8. Formula for log of nominal exchange rate (Ratio approach)

$$s_t = \alpha + \beta(p_t - p_t^*) + \varepsilon_t$$

Where:

- s_t is the **log of the nominal exchange rate**,
- $p_t - p_t^*$ is the **log relative price ratio** (domestic vs. foreign price level),
- α is the **intercept term** (captures systematic bias or baseline exchange rate),
- β is the **coefficient** on the price ratio (equals 1 under strict PPP),
- ε_t is the **error term**, representing **deviation from PPP** (stationary, possibly autocorrelated).

However, before running this regression, we have to ensure that both log of nominal exchange rate and the log of CPI ratio are in the same order of differencing by conducting ADF tests on each series individually. The results suggest that both are stationary after differencing. Before differencing, P-values of Augmented Dickey-Fuller test for Log of nominal exchange rate and log of CPI ratio are 0.9 and 0.4 respectively (see figure 9). This confirms that both variables are integrated of the same order, making the regression suitable in the context of testing for a long-run equilibrium.

Figure 9. ADF test results before OLS estimation

Variable	ADF Statistic	Lag Order	p-value	Stationary
Log Nominal Exchange Rate	-2.4105	4	0.406	No
Log Ratio (Pt/P*)	-1.1753	4	0.9072	No
Log CPI Home	-1.3769	4	0.8354	No
Log CPI Foreign	-2.0359	4	0.5616	No

After confirming that they are at the same order level, we conduct a unit root test on residuals from OLS regression using the ADF test. If relative PPP is valid, the residuals will be stationary, suggesting that deviations from the expected relationship are short-lived and tend to return to the mean. ADF tests on the residuals generate a DF statistic of -2,21. Comparing

Angle-Granger critical values, we have -2,21>-3.37 (5% significance level), indicating that is not sufficiently negative to reject the null hypothesis, indicating that the residuals are non-stationary and thus providing no evidence of cointegration between nominal exchange rate and relative price levels. In other words, relative PPP does not hold in the long run.

Alternative Approach: Testing Relative PPP by Regression Between Log Nominal Exchange Rate and separate parameters (β1,β2)

Given the failure of relative PPP using CPI ratio approach, we attempt an alternative approach by including separate parameters **β1** and **β2** instead of using ratio, which may help to explain the economic influences on exchange rate movements.

Figure 10. Formula for log of nominal exchange rate (**β1** and **β2** approach)

$$s_t = \alpha + \beta_1 p_t + \beta_2 p_t^* + \varepsilon_t$$

where:

- s_t is the **log of the nominal exchange rate**
- p_t is the **log of domestic prices**
- p_t^* is the **log of foreign prices**
- α is the intercept
- β_1 and β_2 are coefficients measuring the effect of domestic and foreign prices on the exchange rate
- ε_t is the **residual term**

Again, before conducting OLS calculation, we have to ensure that log of nominal exchange rate, and the log of foreign consumer index and log of domestic consumer index are integrated at the same order level by performing ADF tests on each of those time series (see figure 9). P-values of ADF tests are all above 0.05, indicating non-stationarity among three of them. After this, we could perform OLS regression, and extract the residuals to conduct an ADF test on it. The outcome once again demonstrates **non-stationarity with p-value DF statistic -2.7>-3.37 (**Angle-Granger critical value)**,** indicating that even with alternative explanatory variables, the exchange rate does not exhibit a persistent long-run relationship with these fundamental economic factors.

5. ARIMA Model selection Forecasting the Real Exchange Rate

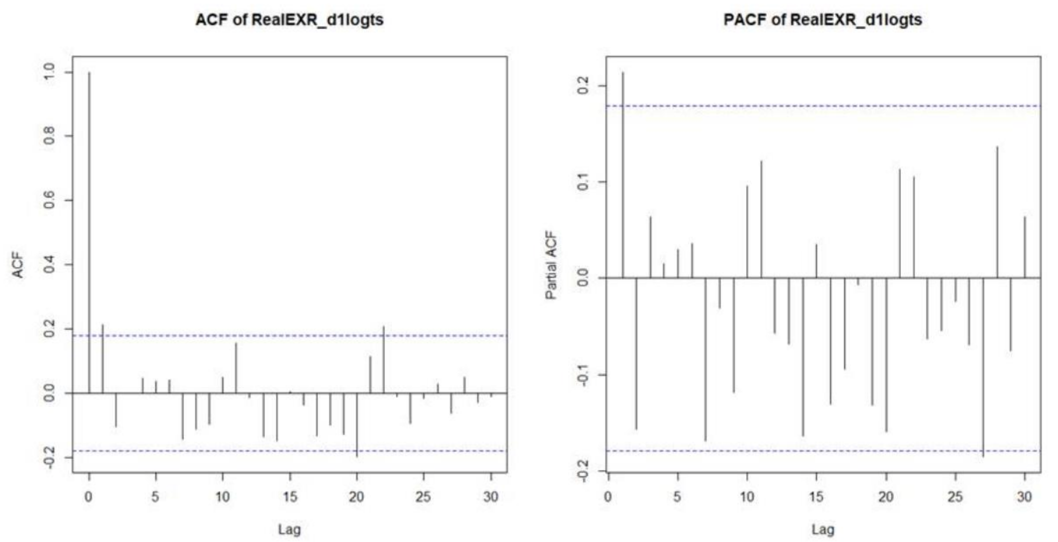
This section outlines the application of the Box-Jenkins methodology to model and forecast the real exchange rate using ARIMA models. The process follows five main stages: identification, estimation, diagnostic checking, model selection, and forecasting (Brooks, 2019, p. 269).

Step 1: Identification

The first step is to identify the appropriate model structure by determining the orders of autoregressive (AR), differencing (I), and moving average (MA) components, i.e., the ARIMA(p,d,q) specification (Levendis, 2023, pp. 105–108).

The autocorrelation function (ACF) of the differenced log real exchange rate series shows a sharp spike at lag 1, followed by a rapid drop-off, indicating an MA(1) process. The partial autocorrelation function (PACF) displays a strong spike at lag 1 and a smaller secondary spike at lag 2, with values trailing off afterward. This suggests the presence of autoregressive components of order one or two.

Figure 11. ACF and PACF plots for the original log real exchange rate series



Based on these observations, six ARIMA models were chosen: ARIMA(1,1,0) for the AR(1) structure, ARIMA(2,1,0) to include the AR(2) component, ARIMA(0,1,1) for the MA(1) process, ARIMA(0,1,2) to account for a potential second-order moving average component, ARIMA(1,1,1) to capture both AR and MA behavior, and ARIMA(2,1,1) to test whether two autoregressive terms with a moving average term would improve model performance.

These models were selected to provide a thorough yet parsimonious representation of the data's underlying dynamics.

Although the ACF and PACF were analyzed using the differenced series, the models were fitted to the raw log real exchange rate, with differencing handled internally by the ARIMA framework for consistency and proper lag alignment. Since the series is differenced (d = 1), no intercept was included. Differencing is expected to remove both stochastic and deterministic trends, and because no evidence of remaining trend or drift was found in the differenced series, including an intercept would risk unnecessary overfitting (Box et al., 2015, pp. 123, 160, 360–364).

Step 2: Estimation

The estimation stage involved fitting six ARIMA models to the log-transformed real exchange rate series using the Arima() function in R, applying maximum likelihood estimation (MLE). MLE estimates parameters by maximizing the probability of observing the data under the model, accounting for the internal structure of the time series, including moving average components (Brooks, 2019, pp. 269, 307).

The models estimated were ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(0,1,1), ARIMA(0,1,2), ARIMA(1,1,1), and ARIMA(2,1,1). For each, we recorded key statistics to assess model quality and complexity.

Coefficients, including AR and MA terms, represent the relationship between current values and past data or errors. Along with these, standard errors were noted, indicating the precision of the estimates. A smaller standard error implies higher confidence in the coefficient value (Brooks, 2019, pp. 110 - 111).

We also recorded the log-likelihood, a measure of model fit, and used penalized fit measures, AIC and BIC (the information criterions), to account for model complexity. Lower AIC and BIC values indicate more efficient models, with BIC penalizing complexity more heavily (Levendis, 2023, pp. 80 - 81; Brooks, 2019, p. 271).

Figure 12. Summary table of estimated coefficients and model fit statistics (AIC, BIC, standard errors) for all ARIMA models

Model	Coefficients	Std_Errors	LogLikelihood	AIC	BIC
ARIMA(1,1,0)	ar1 0.2271	0.0913	304.29	-604.57	-599.00
ARIMA(2,1,0)	ar1 0.2687, ar2 -0.1711	0.0929, 0.0925	305.97	-605.94	-597.58
ARIMA(0,1,1)	ma1 0.3104	0.0967	305.50	-607.00	-601.42
ARIMA(0,1,2)	ma1 0.2702, ma2 -0.0866	0.0942, 0.0863	305.99	-605.97	-597.61
ARIMA(1,1,1)	ar1 -0.1885, ma1 0.4721	0.2196, 0.1868	305.84	-605.68	-597.32
ARIMA(2,1,1)	ar1 0.0119, ar2 -0.1159, ma1 0.265	0.3765, 0.1366, 0.3699	306.17	-604.33	-593.18

Model 3 (ARIMA 0,1,1) had the lowest AIC (-607.00) and BIC (-601.42), with a single MA(1) coefficient of 0.3104 and standard error of 0.0967. It was parsimonious and well-fitting. Models 2 (ARIMA 2,1,0) and 4 (ARIMA 0,1,2) performed well but had slightly higher AIC and BIC values. Model 6 (ARIMA 2,1,1) had higher standard errors and AIC/BIC values, suggesting potential overfitting and instability.

Step 3: Diagnostic checking

Diagnostic checking ensures that the ARIMA models adequately capture the dynamics of the real exchange rate series. A well-specified model should leave residuals that are random, uncorrelated, and homoscedastic, meaning they behave like white noise. This confirms the model has captured all underlying structure in the data (Luo, 2024).

Two key statistical tests were performed: the t-test for coefficient significance and the Ljung-Box Q-test for residual autocorrelation.

The t-test tested whether each AR and MA coefficient was significantly different from zero. The null hypothesis (H_0) was that the coefficient equals zero, suggesting it does not significantly contribute to the model, while the alternative hypothesis (H_1) was that the coefficient differs significantly from zero (Brooks, 2019, p. 131). In Model 1 (ARIMA 1,1,0), the AR(1) coefficient was significant ($t = 2.49$), but in Model 5 (ARIMA 1,1,1), the AR(1) term was not significant ($t = -0.86$). Model 6 (ARIMA 2,1,1) had all three coefficients (AR(1), AR(2), and MA(1)) insignificant, indicating potential overfitting and unnecessary complexity in the model.

Figure 13. t-values of AR and MA coefficients in selected ARIMA models

Model	Coefficient	Estimate	Std..Error	t.value	t.critical	Decision
ARIMA(1,1,0)	ar1	0.2271	0.0913	2.4859	1.9801	Reject H0, accept H1
ARIMA(2,1,0)	ar1	0.2687	0.0929	2.8918	1.9803	Reject H0, accept H1
ARIMA(2,1,0)	ar2	-0.1711	0.0925	-1.8495	1.9803	Do Not Reject H0
ARIMA(0,1,1)	ma1	0.3104	0.0967	3.2086	1.9801	Reject H0, accept H1
ARIMA(0,1,2)	ma1	0.2702	0.0942	2.8675	1.9803	Reject H0, accept H1
ARIMA(0,1,2)	ma2	-0.0866	0.0863	-1.0037	1.9803	Do Not Reject H0
ARIMA(1,1,1)	ar1	-0.1885	0.2196	-0.8581	1.9803	Do Not Reject H0
ARIMA(1,1,1)	ma1	0.4721	0.1868	2.5279	1.9803	Reject H0, accept H1
ARIMA(2,1,1)	ar1	0.0119	0.3765	0.0315	1.9804	Do Not Reject H0
ARIMA(2,1,1)	ar2	-0.1159	0.1366	-0.8485	1.9804	Do Not Reject H0
ARIMA(2,1,1)	ma1	0.2650	0.3699	0.7163	1.9804	Do Not Reject H0

In addition to the t-test, the Ljung-Box Q-test was performed to check for residual autocorrelation at lags 4, 8, and 12. The null hypothesis (H_0) of the Q-test was that the residuals are uncorrelated (white noise), while the alternative hypothesis (H_1) was that autocorrelation exists (Box, 2015, pp. 583–590). All models passed the Q-test with p-values well above the 0.05 threshold, indicating no autocorrelation in the residuals. For example, Model 3 had p-values of 0.819, 0.674, and 0.594 at lags 4, 8, and 12.

Figure 14. Ljung-Box Q-test p-values at lags 4, 8, and 12 for all ARIMA models

Model	Q_Statistic_4	P_Value_4	Q_Critical_4	Decision_4	Q_Statistic_8	P_Value_8	Q_Critical_8	Decision_8	Q_Statistic_12	P_Value_12	Q_Critical_12	Decision_12
ARIMA(1,1,0)	3.7611939	0.4392883	9.487729	Do Not Reject H0	7.677209	0.4656193	15.50731	Do Not Reject H0	12.699210	0.3912771	21.02607	Do Not Reject H0
ARIMA(2,1,0)	0.6909727	0.9524375	9.487729	Do Not Reject H0	4.899895	0.7682199	15.50731	Do Not Reject H0	8.483405	0.7463057	21.02607	Do Not Reject H0
ARIMA(0,1,1)	1.5429211	0.8190123	9.487729	Do Not Reject H0	5.761254	0.6739548	15.50731	Do Not Reject H0	10.247121	0.5942907	21.02607	Do Not Reject H0
ARIMA(0,1,2)	0.4589986	0.9773678	9.487729	Do Not Reject H0	5.047020	0.7525390	15.50731	Do Not Reject H0	9.247191	0.6816846	21.02607	Do Not Reject H0
ARIMA(1,1,1)	0.7312529	0.9474196	9.487729	Do Not Reject H0	5.238788	0.7317759	15.50731	Do Not Reject H0	9.633570	0.6480723	21.02607	Do Not Reject H0
ARIMA(2,1,1)	0.2774231	0.9912246	9.487729	Do Not Reject H0	4.911594	0.7669819	15.50731	Do Not Reject H0	8.819406	0.7182733	21.02607	Do Not Reject H0

Residual plots supported these findings, showing no significant spikes in the ACF and standardized residuals randomly scattered with no patterns or outliers. The plot of the Ljung-

Box p-values remained above the significance threshold, confirming the residuals behaved as expected.

Figure 15. Residual ACF plots for the selected ARIMA models

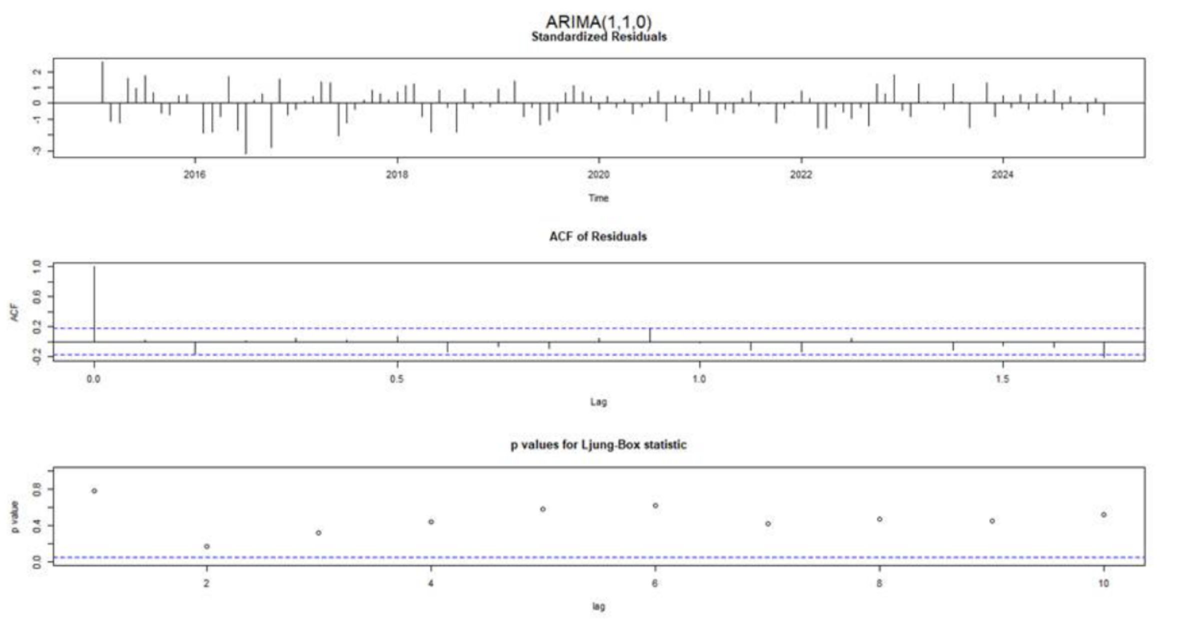
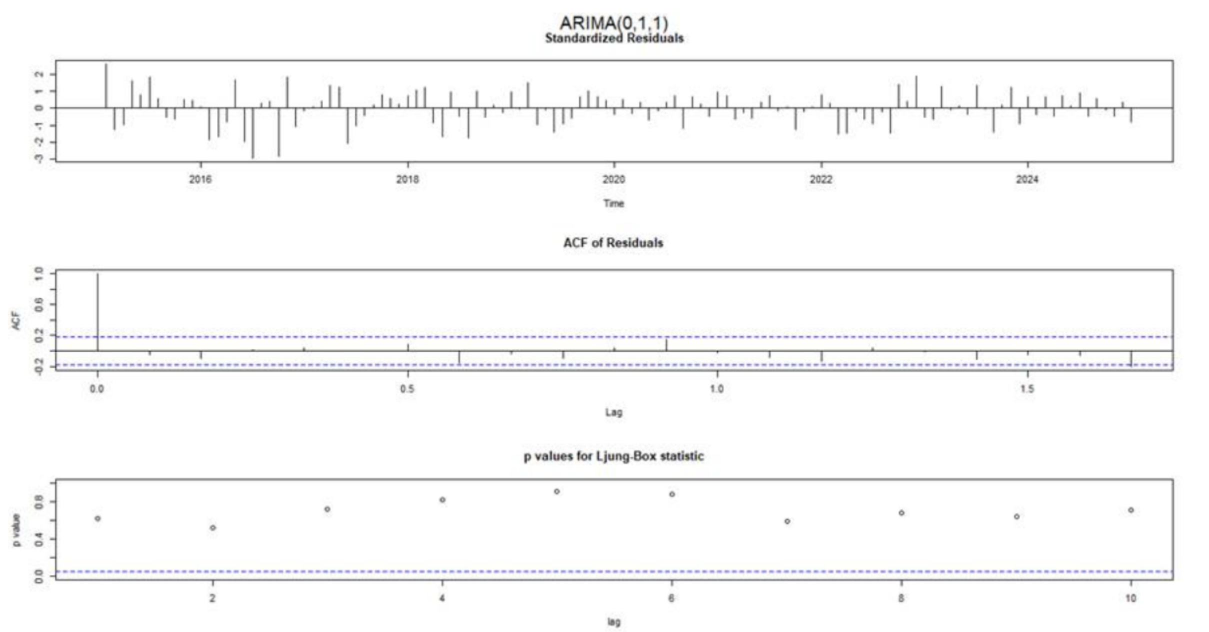


Figure 16. Residual ACF plots for the selected ARIMA models



These diagnostic checks confirm that all six models are statistically adequate in terms of residual behavior. However, the presence of insignificant coefficients in several models raises concerns about overfitting and the inclusion of unnecessary terms, which is addressed in the next step through model selection.

Step 4: Model Selection

Model 3 (ARIMA(0,1,1)) emerges as the best model for forecasting the real exchange rate based on several key criteria. Model 1 (ARIMA(1,1,0)), Model 2 (ARIMA(2,1,0)), and Model 3 all have significant AR(1) and MA(1) coefficients, indicating meaningful contributions to explaining the data. In contrast, Model 5 (ARIMA(1,1,1)) has an insignificant AR(1) coefficient (t-value = -0.86), and Model 6 (ARIMA(2,1,1)) has multiple non-significant coefficients, suggesting overfitting and lack of additional value from extra parameters.

Model 3 has the lowest AIC and BIC, indicating the best fit with the fewest parameters. Despite more parameters, Model 6 has higher AIC and BIC values, suggesting overfitting. Model 1 and Model 5 perform well but don't surpass Model 3 in terms of fit and simplicity.

The Ljung-Box Q-test, assessing autocorrelation in residuals, shows all models pass with p-values above 0.05, indicating no autocorrelation. Model 3 performs particularly well, with high p-values across all lags, supporting its reliability.

Finally, Model 3 has the lowest Sum of Squared Residuals (SSR), meaning it produces the smallest errors between observed and predicted values.

Figure 17. Comparison of ARIMA models based on AIC, BIC, and Sum of Squared Residuals (SSR)

Model	α1 (AR1)	α2 (AR2)	β1 (MA1)	β2 (MA2)	SSR	AIC	BIC	Q4	Q8	Q12
ARIMA(1,1,0)	0.2271 (t=2.49)				0.04406	-604.57	-599	3.76 (p=0.4393)	7.68 (p=0.4656)	12.70 (p=0.3913)
ARIMA(2,1,0)	0.2687 (t=2.89)	-0.1711 (t=-1.85)			0.04282	-605.94	-597.58	0.69 (p=0.9524)	4.90 (p=0.7682)	8.48 (p=0.7463)
ARIMA(0,1,1)			0.3104 (t=3.21)		0.04316	-607	-601.42	1.54 (p=0.8190)	5.76 (p=0.6740)	10.25 (p=0.5943)
ARIMA(0,1,2)			0.2702 (t=2.87)	-0.0866 (t=-1.00)	0.04281	-605.97	-597.61	0.46 (p=0.9774)	5.05 (p=0.7525)	9.25 (p=0.6817)
ARIMA(1,1,1)	-0.1885 (t=-0.86)		0.4721 (t=2.53)		0.04291	-605.68	-597.32	0.73 (p=0.9474)	5.24 (p=0.7318)	9.63 (p=0.6481)
ARIMA(2,1,1)	0.0119 (t=0.03)	-0.1159 (t=-0.85)	0.2650 (t=0.72)		0.04267	-604.33	-593.18	0.28 (p=0.9912)	4.91 (p=0.7670)	8.82 (p=0.7183)

In summary, Model 3 (ARIMA(0,1,1)) is the most suitable choice, excelling in AIC, BIC, SSR, and simplicity compared to Models 1 and 5. Model 6's complexity and weaker performance make it less viable, solidifying Model 3 as the optimal option for forecasting the real exchange rate.

Step 5: Forecasting

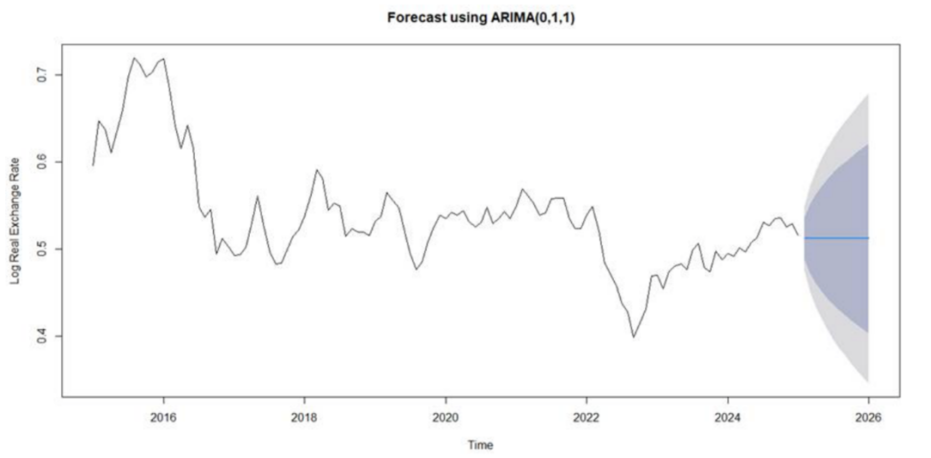
The ARIMA(0,1,1) model was chosen as the best option for forecasting the real exchange rate from February 2025 to January 2026. The point forecast remains constant at 0.5119, but confidence intervals widen over time, showing increased uncertainty. For instance, the 95% interval expands from 0.4746–0.5492 in February 2025 to 0.3454–0.6784 by January 2026, illustrating growing variability.

Figure 18. Forecast plot from ARIMA(0,1,1) model for February 2025 to January 2026

Month	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Feb 2025	0.5119	0.4875	0.5363	0.4746	0.5492
Mar 2025	0.5119	0.4717	0.5521	0.4504	0.5734
Apr 2025	0.5119	0.4605	0.5633	0.4333	0.5905
May 2025	0.5119	0.4514	0.5724	0.4193	0.6045
Jun 2025	0.5119	0.4434	0.5804	0.4072	0.6166
Jul 2025	0.5119	0.4363	0.5875	0.3963	0.6275
Aug 2025	0.5119	0.4299	0.5940	0.3864	0.6374
Sept 2025	0.5119	0.4238	0.6000	0.3772	0.6466
Oct 2025	0.5119	0.4182	0.6056	0.3686	0.6552
Nov 2025	0.5119	0.4129	0.6109	0.3605	0.6633
Dec 2025	0.5119	0.4079	0.6159	0.3528	0.6710
Jan 2026	0.5119	0.4031	0.6207	0.3454	0.6784

The forecast plot visually represents the point forecasts along with the 80% and 95% confidence intervals. The shaded areas in the plot provide a clear view of the uncertainty around the predictions, with the wider 95% interval indicating greater uncertainty about future values.

Figure 19. Zoomed-in view of the forecast intervals



Residual analysis confirmed the model's adequacy. The Ljung-Box test, with a Q-statistic of 31.081 and a p-value of 0.1208, showed no significant autocorrelation in the residuals, which resemble white noise. Residual plots revealed no discernible patterns, while the autocorrelation function (ACF) showed all spikes within bounds. A histogram indicated normal distribution, further validating the model's reliability.

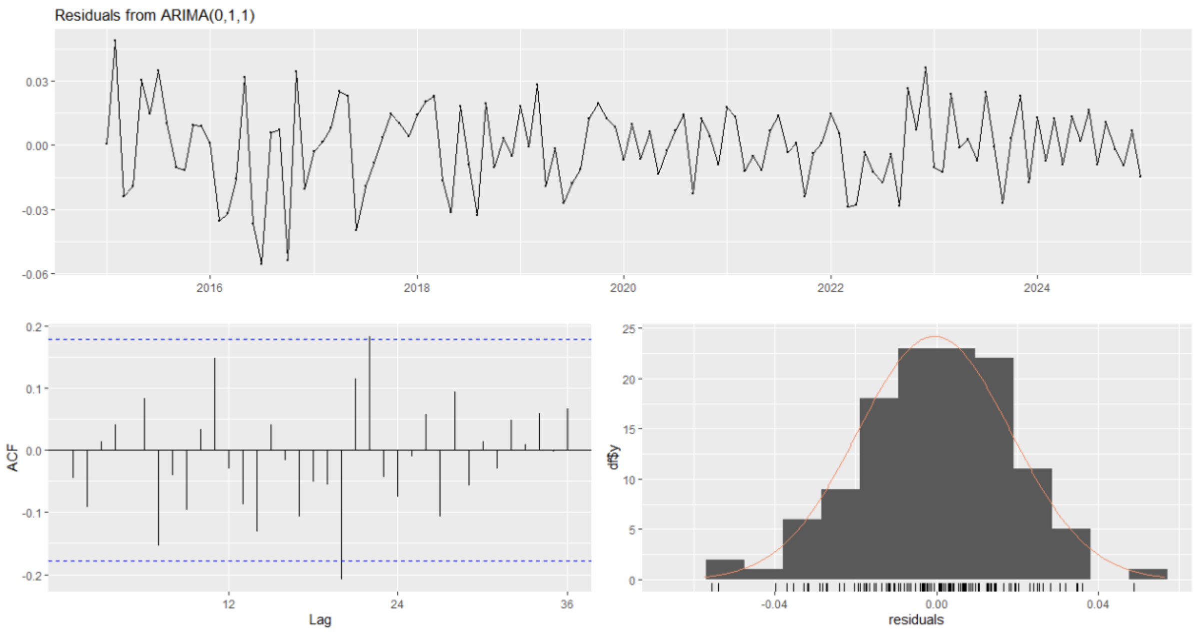
Figure 20. Ljung-Box test for selected model ARIMA(0,1,1)

```
	Ljung-Box test

data:  Residuals from ARIMA(0,1,1)
Q* = 31.081, df = 23, p-value = 0.1208

Model df: 1.   Total lags used: 24
```

Figure 21. Residual Diagnostic Plot for ARIMA(0,1,1)



Although metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were not computed due to unavailable data, the residual diagnostics provide strong evidence of the model's reliability (Brooks, 2019, p. 285). The residuals are centered around zero and lack significant correlations, further validating the model's suitability for forecasting the future real exchange rate.

6. Results and Discussion

6.1 Summary of Key Findings

6.1.1 Results and Validity Analysis of Purchasing Power Parity (PPP)

The results of Absolute purchasing power parity (Absolute PPP) show that the null hypothesis cannot be rejected (ADF statistic is − 2.72, p value is 0.28 (>0.05)). This leads to the conclusion that the real exchange rate is non-stationary and the divergence from the absolute PPP is persistent rather than a transitory fluctuation. This suggests that absolute PPP does not hold and that exchange rates fail to adjust quickly enough to equalize cross-country price levels.

For the Relative purchasing power parity (PPP) test, the ADF statistic of the resulting residual is − 2.21, which is greater than the critical value of − 3.37 (5% significance level), which also shows that the null hypothesis of non-stationarity cannot be rejected. In this regard, the CPI is further split and regressed, which further verifies the failure of relative PPP.

6.1.2 Time Series Modeling Results and Applicability Analysis of Forecasting Model

Based on the Box-Jenkins methodology and through a systematic modeling process, we constructed six models. The following is the performance comparison of the modeling results:

Figure 22. Performance Comparison of Six Models

MODEL TYPE	*SIGNIFICANT PARAMETERS*	*AIC*	*BIC*	*RESIDUAL TEST*	*KEY ADVANTAGES*	*MAIN LIMITATIONS*
ARIMA (0,1,1)	MA1(0.3104, t=3.21)	-607.00	-601.42	Passed	Simplest structure, smallest errors	Only captures 1st-order MA effect
ARIMA (1,1,0)	AR1(0.2271, t=2.49)	-604.57	-599.00	Passed	Clear autoregressive property	Ignores moving average component
ARIMA (2,1,0)	AR1(0.2687, t=2.89) AR2(-0.1711, t=-1.85)	-605.94	-597.58	Passed	Captures 2nd-order autocorrelation	AR2 coefficient insignificant
ARIMA (0,1,2)	MA1(0.2702, t=2.87) MA2(-0.0866, t=-1.00)	-605.97	-597.61	Passed	Flexible MA structure	MA2 coefficient insignificant
ARIMA (1,1,1)	MA1(0.4721, t=2.53) AR1(-0.1885, t=-0.86)	-605.68	-597.32	Passed	Combines AR and MA features	AR1 coefficient insignificant
ARIMA (2,1,1)	All coefficients insignificant	-604.33	-593.18	Passed	Theoretically complete	Clearly overfit, low practicality

Finally, the optimal prediction model ARIMA (0,1,1) model is determined according to the statistical significance (parameter t test value needs to be significant), information criterion, model simplicity, and judgment criterion of white noise residual diagnosis.

ARIMA(0,1,1) (optimal model) : AIC/BIC is the lowest (-607.00/-601.42), the sum of squared residuals is the lowest (SSR=0.04316), and the goodness of fit is the best; the one-parameter model is easy to estimate and interpret and is suitable for real-time forecasting; the residuals passed all diagnostic tests (Ljung-Box $p>0.5$) and showed no autocorrelation or heteroscedasticity, indicating strong robustness. However, MA (1) only captures the error shocks of the previous period, which is insufficient in response to long-term trends or structural changes and may also underestimate the risk of market breaks (constant point forecast (0.5119)).

This model is suitable for short-term (1-3 months) exchange rate fluctuation prediction, the trend tracking and capture of the market stable period and the need to quickly calculate real-time forecasts. It is not suitable for the prediction of extreme events such as policy shocks or financial crises (Cheung et al., 2005).

6.2 Implications for Banks/Policymakers

6.2.1 Implications from PPP Test Results

For exchange rate system, it is not appropriate to rely solely on PPP as the basis of exchange rate anchoring, but to establish an exchange rate evaluation framework including multiple factors (Rogoff, 1996) ; since CPI differentials are not adjusted across borders through exchange rates and the price transmission mechanism is blocked, monetary policy Monetary policy should pay more attention to domestic inflation expectations (Cavallo et al., 2014); for the risk hedging strategy, the risk of PPP failure can be compensated by extending the hedging contract term, and the dynamic hedging ratio can be adopted; in the asset valuation of international investment decision, the traditional PPP valuation model should be revised to increase the weight of local market factors (Bartram et al., 2010).

6.2.2 Implications from ARIMA Model Prediction

In terms of trading strategy, the 1–3-month prediction interval of ARIMA (0,1,1) can be used to set the trading threshold and optimize the timing of opening/closing positions by combining with the lag effect of MA (1); adjust foreign currency provisions according to forecast fluctuations, match the currency maturity structure of assets and liabilities, and implement liquidity management; in terms of foreign exchange intervention, we can make a progressive intervention strategy by referring to the forecast interval, and establish an automatic response mechanism of "forecast error trigger" (Dominguez, 1998).

6.3 Limitations of the Study

The PPP test does not cover the price difference of non-tradables, the heterogeneity of consumption preferences, institutional trade barriers and other factors.

The boundary of ARIMA model is not suitable for extreme scenarios of currency crisis, the initial stage of digital currency shock and other scenarios, and subsequent measures such as the establishment of model failure warning indicators can be taken for optimization.

7. Conclusion

The ADF tests confirmed that the original nominal and real exchange rate series, domestic and foreign CPI data, were non-stationary but became stationary after first differencing, showing integration of order one (I(1)). Tests for absolute and relative PPP showed that deviations from parity are persistent, suggesting that real-world exchange rates reflect additional factors—such as policy interventions or market sentiment—beyond relative prices alone. ARIMA(0,1,1) was proved as the best model, balancing parsimony with robust residual diagnostics and strong fit criteria (AIC/BIC). Although it is effective for short-term forecasts, this may not capture structural shifts or policy shocks, if without further refinement or macroeconomic indicators. Overall, although PPP provides theoretical perspectives, actual exchange rates can diverge for market frictions, and a carefully selected ARIMA model can offer practical short-run predictive power.

References

- Bartram, S. M., Brown, G. W., & Minton, B. A. (2010). *Resolving the exposure puzzle: The many facets of exchange rate exposure*. Journal of Financial Economics, *95*(2), 148–173. https://doi.org/10.1016/j.jfineco.2009.09.002
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Box, G. E. P. (2015). *Time series analysis: Forecasting and control* (1st ed.). John Wiley & Sons.
- Brooks, C. (2019). *Introductory econometrics for finance* (2nd ed.). Cambridge University Press.
- Cavallo, A., Neiman, B., & Rigobon, R. (2014). C*urrency Unions, Product Introductions, and the Real Exchange Rate.* The Quarterly Journal of Economics, *129*(2), 529–595. https://doi.org/10.1093/qje/qju008
- Cheung, Y.-W., Chinn, M. D., & Pascual, A. G. (2005). *Empirical exchange rate models of the nineties: Are any fit to survive?* Journal of International Money and Finance, *24*(7), 1150–1175. https://doi.org/10.1016/j.jimonfin.2005.08.002
- Dominguez, K. M. (1998). C*entral bank intervention and exchange rate volatility*. Journal of International Money and Finance, *17*(1), 161–190. https://doi.org/10.1016/s0261-5606(97)98055-4
- Levendis, J. D. (2023). *Time series econometrics: Learning through replication* (2nd ed.). Springer.
- Luo, L. (2024). *Statistical model validation through white noise hypothesis testing in regression analysis and ARIMA models.* Theoretical and Natural Science, 50(1), 99–104. https://doi.org/10.54254/2753-8818/50/20240672
- *Monetary policy and central banking*. (n.d.). https://www.imf.org/en/About/Factsheets/Sheets/2023/monetary-policy-and-central-banking
- *Times Series: Cointegration*. (n.d.). https://pure.au.dk/ws/portalfiles/portal/82215843/rp14_38.pdf
- Rogoff, K. (1996)**.** *The purchasing power parity puzzle.* Journal of Economic Literature, 34(2), 647–668.