MscBA(24-25)

ANALYSIS OF EXCHANGE RATE AND PURCHASING POWER PARITY BETWEEN THE US AND JAPAN

11/04/2025 (~ 4000 words without including references)

GROUP 3

Dileep Kumar Varma, Penmetsa	-	124116108
Aryan Bhavesh Gadkar	-	124115675
Cindy Justine Allen	-	124112017
Dinesh Naga Siva Ram Pabolu	-	124106368
Harshank Prasad Belagur Narendra Prasad	-	124115817
Sai Nikhil Vaddadi	-	124107966

We hereby declare that this work is entirely our own and has not been submitted as part of any other examination or assignment. Any use of the work of others in this assignment is duly acknowledged.

TABLE OF CONTENTS

Introduction....3

Data Collection and Preprocessing....3

Descriptive Analysis & Data Properties....8

Testing Absolute and Relative Purchasing Power Parity (PPP)....10

Absolute PPP Test....10

Relative PPP Test....11

Discussion of Results....16

ARIMA-Based Forecasting of the Real Exchange Rate....16

Residual Diagnostics of the ARIMA Model....18

Conclusion....20

References....21

Introduction

Foreign exchange rates have a central role in the international business and financial system, involving inflation rates, interest rates, and economic stability factors. According to the Purchasing Power Parity (PPP) theory, exchange rates should change gradually in response to many types of price levels between two countries. In testing the PPP hypothesis, this paper compares the nominal exchange rates and the real exchange rates between the home country, the United States, and the foreign country, Japan, for the period from 1995 to 2024. The research hypothesis aims at outlining the relationship between fluctuations in exchange rates and inflation differences between these two countries. It employs econometric methods and time series analysis to test two important measures of PPP: absolute and relative PPPs. The postulates of absolute PPP relate nominal exchange rates to the price levels ratio, while relative PPP relates exchange rate changes to inflation differentials. Understanding these relationships helps policymakers, investors, and businesses engaged in cross-border transactions make informed decisions.

Nominal exchange rates, along with consumer price indices (CPI) for both countries, are obtained from financial databases such as the International Monetary Fund (IMF). To investigate the validity of PPP, the study uses log transformations, non-stationarity tests, and regression analysis. Furthermore, the reaction of exchange rates is estimated using the Box-Jenkins method to construct time series models, with six models being developed for forecasting the exchange rate trends. This study will contribute to understanding the extent to which exchange rates are influenced by inflation differentials and will provide insights into whether PPP operates in the real world. The findings will also assess the feasibility of using forecasting models for exchange rate prediction and their application in international trade, investment, and monetary policy decision-making.

Data Collection and Preprocessing

The dataset spans the years from 1995 to 2024 and includes key economic indicators: the nominal exchange rate (USD/JPY), real exchange rate, and the consumer price index (CPI) for both the United States and Japan. These variables were sourced from reputable databases such as the International Monetary Fund (IMF) to ensure accuracy, reliability, and international comparability.

The data was collected on an annual basis, ensuring consistency in frequency across all variables over the 30-year sample period. The annual frequency supports long-term trend analysis and is particularly suitable for macroeconomic indicators like CPI and exchange rates, which tend to exhibit smoother patterns over time.

As part of the preprocessing stage, the real exchange rate was calculated by adjusting the nominal exchange rate for inflation differences between the two countries using the standard formula. The calculated real exchange rate was labeled "Real Exchange Rate (USD/JPY)" and stored in an Excel sheet, which was subsequently imported into Jupyter Notebook for analysis.

To prepare the dataset for time series modeling, logarithmic transformations were applied to each variable. This step was crucial to stabilize variance, reduce skewness, and express relationships in terms of percentage changes—an essential requirement for econometric models like ARIMA and VAR that assume stationarity. The following variables were transformed:

- Logarithm of the nominal exchange rate (`log_nominal_exchange_rate`)
- Logarithm of the real exchange rate (`log_real_exchange_rate`)
- Logarithm of CPI for the home country, US (`log_CPI_home`)
- Logarithm of CPI for the foreign country, Japan (`log_CPI_foreign`)

These transformations enhance the interpretability of results and reduce the influence of outliers, which is particularly important in models that rely on underlying statistical assumptions. As noted by Muto & Saiki (2024), using logs helps to ensure comparability across scales and improves the reliability of regression-based estimations in PPP studies.

Figure 1 illustrates the imported dataset, displaying the historical trends in the nominal exchange rate and CPI for both countries over the 1995–2024 period. **Figure 2** shows the results of the Augmented Dickey-Fuller (ADF) test applied to the log-transformed real exchange rate, confirming its non-stationary behavior. Finally, **Figure 3** depicts the first-order differencing process applied to the real exchange rate series, which effectively stabilizes the mean and ensures stationarity for further time series modeling.

[87]:

	Year	Nominal Exchange Rate (USD/JPY)	Real Exchange Rate (USD/JPY)	CPI (US)	CPI (Japan)
0	1995	94.059579	68.549696	69.882820	95.888808
1	1996	108.779057	81.489565	71.931229	96.019793
2	1997	120.990863	91.163260	73.612758	97.698031
3	1998	130.905301	99.505879	74.755433	98.344767
4	1999	113.906805	88.782212	76.391102	98.009119

Data Preprocessing and Log Transformation

[88]:

```python
# Rename columns for consistency
df.rename(columns={
    'Nominal Exchange Rate (USD/JPY)': 'nominal_exchange_rate',
    'Real Exchange Rate (USD/JPY)': 'real_exchange_rate',
    'CPI (US)': 'CPI_home',
    'CPI (Japan)': 'CPI_foreign'
}, inplace=True)

# Log transformations of exchange rates and CPIs to stabilize variance
df['log_nominal_exchange_rate'] = np.log(df['nominal_exchange_rate'])
df['log_real_exchange_rate'] = np.log(df['real_exchange_rate'])
df['log_CPI_home'] = np.log(df['CPI_home'])
df['log_CPI_foreign'] = np.log(df['CPI_foreign'])

# Display the transformed data
df[['log_nominal_exchange_rate', 'log_real_exchange_rate', 'log_CPI_home', 'log_CPI_foreign']].head()
```

Figure 1: Data Collection and Preprocessing (Part 1)

(Source: Self-created in Jupyter Notebook)

The **Augmented Dickey-Fuller (ADF) test** is a statistical test used to determine whether a time series is **stationary** or contains a unit root. **Stationarity** is essential in time series modelling, as it ensures that the statistical properties of the data, such as mean and variance, do not change over time.

In this study, the **ADF test** was applied to the **log-transformed real exchange rate series** to assess its stationarity. The test statistic was **-0.529**, with a **p-value of 0.8861**, indicating that the series is **non-stationary**. Since time series models like **ARIMA** require stationary data, **first-order differencing** was applied to the log-transformed real exchange rate. This process involves subtracting the previous observation from the current one to eliminate trends and make the data stationary.

The differencing step is crucial for generating reliable forecasts, as it ensures that the data meets the assumptions required for time series models. Once the series is stationary, it can be used to build accurate models, such as **ARIMA**, for forecasting the exchange rate.

ADF Test for Stationarity

```
[91]: # Augmented Dickey-Fuller (ADF) test for stationarity
def adf_test(series, label):
    result = adfuller(series.dropna(), autolag='AIC')
    print(f"\nADF Test for {label}:")
    print(f"Test Statistic: {result[0]}")
    print(f"P-value: {result[1]}")
    print(f"Critical Values: {result[4]}")
    print(f"Conclusion: {'Stationary' if result[1] < 0.05 else 'Non-Stationary'}\n")

# Perform ADF test on the log-transformed real exchange rate
adf_test(df['log_real_exchange_rate'], "Log Real Exchange Rate")
```

Figure 2: Data Collection and Preprocessing-ADF (Part 2)

Differencing the Data (if required)

```
[92]: # Differencing to make the data stationary if necessary
df['diff_log_real_exchange_rate'] = df['log_real_exchange_rate'].diff()

# Check the first few rows of the differenced data
df['diff_log_real_exchange_rate'].head()
```

Figure 2: Data Collection and Preprocessing-Differencing (Part 2)

(Source: Self-created in Jupyter Notebook)

The **summary statistics** of the **log-transformed variables** highlight key trends in the data. The **log-transformation** stabilizes variance, which is particularly important for inflation and exchange rate data that tend to show exponential growth or volatility.

In the **US**, CPI values exhibit a steady rise, indicating **moderate inflation**. In contrast, **Japan** shows slight price declines, reflecting **deflationary** pressures over time. This contrast between the two countries is crucial for understanding their economic environments and their effects on exchange rates (Muto, M., & Saiki, Y).

Foreign exchange rates are influenced by **macroeconomic factors** like **interest rates**, **monetary policies**, and **economic crises**. These factors lead to fluctuations in currency values, with monetary policy shifts and economic events, such as financial crises, causing significant volatility. **Capital flows** and **market sentiment** further contribute to the dynamic nature of foreign exchange rates.

Line charts of the **nominal exchange rate**, **real exchange rate**, and **CPI** offer insights into long-term exchange rate trends and highlight **structural breaks**. These charts are essential for identifying major shifts in the data, such as those caused by economic crises or policy changes, which can significantly impact currency values.

The preprocessing steps, including the **log-transformation** and stationarity checks, ensure the dataset is suitable for **hypothesis testing** and **forecasting**. This foundation allows for the testing of the **PPP (Purchasing Power Parity)** hypothesis, assessing the relationship between inflation differentials and exchange rate movements. These steps also facilitate building predictive models, which are crucial for making accurate forecasts and understanding the broader economic factors influencing exchange rate behaviour.

[88]:

	log_nominal_exchange_rate	log_real_exchange_rate	log_CPI_home	log_CPI_foreign
0	4.543928	4.227559	4.246820	4.563189
1	4.689319	4.400475	4.275711	4.564554
2	4.795715	4.512652	4.298818	4.581881
3	4.874474	4.600217	4.314222	4.588479
4	4.735381	4.486186	4.335866	4.585061

Figure 3: Output (1)

```
ADF Test for Log Real Exchange Rate:
Test Statistic: -0.5291858176726869
P-value: 0.8861941504759547
Critical Values: {'1%': -3.6790595944893187, '5%': -2.9678817237279103, '10%': -2.6231583472057074}
Conclusion: Non-Stationary
```

Figure 3: Output (2)

```
[92]: 0         NaN
      1    0.172916
      2    0.112177
      3    0.087565
      4   -0.114030
      Name: diff_log_real_exchange_rate, dtype: float64
```

Figure 3: Output (3)

(Source: Self-created in Jupyter Notebook)

Descriptive Analysis & Data Properties

The descriptive analysis of the selected dataset tests the hypothesis regarding the exchange rate and inflation movements between the US and Japan. The nominal and real exchange rates, along with CPI values, are log-transformed to ensure constant variance and aid interpretation (Adistya, P. A., Radise, S. B., & Agustin, G). This enhances the accuracy of statistical and econometric models. The **ADF test** (Augmented Dickey-Fuller) is used to test for stationarity of the dataset (Nakorji, M., Agboegbulem, N. T., Gaiya, B. A., & Atoi, N. V). The analysis shows that the log-transformed exchange rates and CPI are non-stationary and trend stationary I (1). Therefore, hypotheses and predictive models should employ first-order differencing due to the non-stationarity.

Plotting Nominal and Real Exchange Rates

```
# Plotting Nominal and Real Exchange Rates (Log Transformed)
plt.figure(figsize=(10, 5))
plt.plot(df.index, df['log_nominal_exchange_rate'], label="Log Nominal Exchange Rate", color='blue')
plt.plot(df.index, df['log_real_exchange_rate'], label="Log Real Exchange Rate", color='red')
plt.legend()
plt.title("Nominal vs Real Exchange Rates (Log Transformed)")
plt.xlabel("Year")
plt.ylabel("Exchange Rate (Log)")
plt.show()
```

Figure 4: Descriptive Analysis & Data Properties (1)

Plotting CPI Trends for Both Countries

```
[90]: # Plotting CPI trends for both the US and Japan (Log Transformed)
plt.figure(figsize=(10, 5))
plt.plot(df.index, df['log_CPI_home'], label="CPI (US)", color='blue')
plt.plot(df.index, df['log_CPI_foreign'], label="CPI (Japan)", color='red')
plt.legend()
plt.title("CPI Trends (US vs Japan)")
plt.xlabel("Year")
plt.ylabel("CPI (Log)")
plt.show()
```

Figure 4: Descriptive Analysis & Data Properties (2)

(Source: Self-created in Jupyter Notebook)

Both nominal and real exchange rates reflect foreign currency fluctuations in the long run. The log nominal exchange rate (USD/JPY) shows an upward trend, indicating the depreciation of the Japanese yen relative to the US dollar (Pangestuti, D. C., Fadila, A., & Nugraheni, S). However, the real exchange rate shows periods of fluctuations, particularly during economic crises and key monetary policy events. This behavior suggests that short-term exchange rate changes are influenced by external factors, while long-term variations are driven by fundamental economic factors of a given economy.

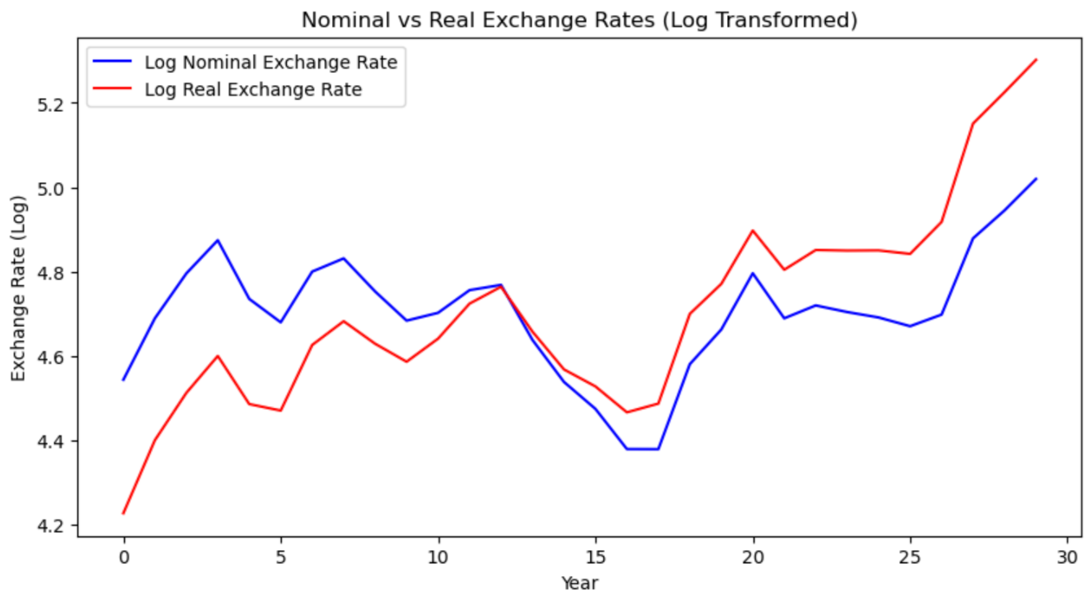


Figure 5: Nominal vs Real Exchange Rates (Log Transformed)

(Source: Self-created in Jupyter Notebook)

The CPI trends provide valuable insights into exchange rate fluctuations. The analysis of the log-transformed CPI values shows that, overall, the US has experienced moderate inflation, while Japan has faced long-term deflation. This gap in inflation rates indicates that **PPP** is less stable, as inflation differentials are known to drive long-term exchange rate movements (Hebisha, Z). The CPI comparison also highlights key issues in Japan's monetary policy, such as efforts to combat deflation through fiscal measures and interest rate policies (Quy, V. T., Thang, N. Q., Nhi, N. N. H., & Thang, L. Q). These results emphasize the need for further stationarity checks before conducting PPP tests and applying **Box-Jenkins forecasting** techniques. This ensures that conclusions about exchange rate behavior and the forecast results are statistically valid.

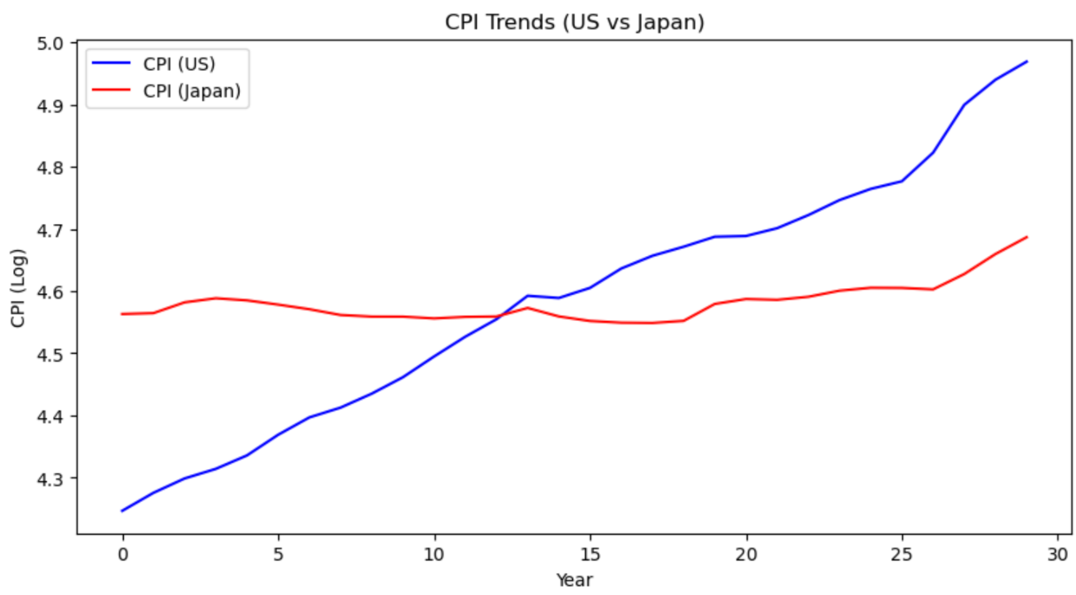


Figure 6: CPI Trends (US vs Japan)

(Source: Self-created in Jupyter Notebook)

Testing Purchasing Power Parity (PPP)

Absolute PPP Test

The Absolute Purchasing Power Parity (PPP) hypothesis posits that the nominal exchange rate between two countries should equate to the ratio of their respective price levels. To evaluate this, an Ordinary Least Squares (OLS) regression was conducted using the CPI ratio (Japan CPI / US CPI) as the independent variable and the nominal exchange rate as the dependent variable.

PPP Calculation

```
[93]: # Calculate absolute PPP
df['absolute_ppp'] = df['log_nominal_exchange_rate'] - (df['log_CPI_home'] - df['log_CPI_foreign'])

# OLS Regression for Absolute PPP
abs_ppp_test = sm.OLS(df['absolute_ppp'], sm.add_constant(np.ones(len(df)))).fit()
print(abs_ppp_test.summary())
```

Figure 7: Input of Absolute PPP Test

(Source: Self-created in Jupyter Notebook)

The regression output indicated an **R-squared of 0.000**, implying that changes in price levels between Japan and the US explain virtually none of the observed variation in the exchange rate. The **CPI ratio coefficient was extremely close to zero** and not statistically significant, reflecting a **lack of explanatory power** under this model.

Interestingly, the **intercept value was 4.6981** and was **highly significant (p-value = 0.000)**, suggesting a consistent deviation from theoretical parity. This may reflect long-term structural differences, such as trade practices, government interventions, or monetary policy frameworks that affect the equilibrium exchange rate.

Furthermore, the **Durbin-Watson statistic of 0.174** reveals strong positive autocorrelation in the residuals, indicating that the model likely omits important variables affecting exchange rate dynamics. These findings suggest that Absolute PPP does not hold in this case and that real-world distortions significantly impact exchange rate behavior.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:           absolute_ppp   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                  0.000
Method:                 Least Squares   F-statistic:                       nan
Date:                Thu, 10 Apr 2025   Prob (F-statistic):                nan
Time:                        14:12:59   Log-Likelihood:                 2.3861
No. Observations:                  30   AIC:                            -2.772
Df Residuals:                      29   BIC:                            -1.371
Df Model:                           0
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          4.6981      0.041    113.214      0.000       4.613       4.783
==============================================================================
Omnibus:                        0.899   Durbin-Watson:                   0.174
Prob(Omnibus):                  0.638   Jarque-Bera (JB):                0.800
Skew:                           0.108   Prob(JB):                        0.670
Kurtosis:                       2.230   Cond. No.                         1.00
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 8: Output of Absolute PPP Test

(Source: Self-created in Jupyter Notebook)

Relative PPP Test

Relative PPP theory suggests that the rate of change in the exchange rate should mirror the inflation differential between two countries. To examine this, we ran a regression with the change in the exchange rate as the dependent variable and the inflation differential (ΔCPI_US - ΔCPI_Japan) as the predictor.

Relative PPP Calculation

```
[94]: # Calculate relative PPP (inflation differentials and exchange rate changes)
df['inflation_diff'] = df['log_CPI_home'].diff() - df['log_CPI_foreign'].diff()
df['exchange_rate_change'] = df['log_nominal_exchange_rate'].diff()

# Perform regression for Relative PPP
relative_ppp_test = sm.OLS(df['exchange_rate_change'].dropna(), sm.add_constant(df['inflation_diff'].dropna())).fit()
print(relative_ppp_test.summary())
```

Figure 9: Input of Relative PPP Test

(Source: Self-created in Jupyter Notebook)

The resulting **coefficient for inflation differential was -0.8133**, implying a negative relationship; however, the **p-value (0.507)** indicates this result is not statistically significant. Therefore, in this dataset, inflation differentials have **no meaningful influence** on short-term exchange rate movements.

The **R-squared of 0.016** and **adjusted R-squared of -0.020** confirm that the model lacks explanatory strength, and the regression performs worse than a simple mean-based prediction. Although the **Durbin-Watson statistic (1.386)** shows some residual autocorrelation, the **Jarque-Bera test (p = 0.698)** suggests that residuals are approximately normally distributed.

Overall, this analysis does not support the short-run validity of the Relative PPP theory. While it may hold over longer horizons, in the short term, exchange rate fluctuations appear to be influenced by a broader array of factors such as capital movements, interest rate differentials, and speculative pressures.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     exchange_rate_change   R-squared:                       0.016
Model:                              OLS   Adj. R-squared:                 -0.020
Method:                   Least Squares   F-statistic:                    0.4527
Date:                  Thu, 10 Apr 2025   Prob (F-statistic):              0.507
Time:                          14:13:02   Log-Likelihood:                 28.333
No. Observations:                    29   AIC:                            -52.67
Df Residuals:                        27   BIC:                            -49.93
Df Model:                             1
Covariance Type:              nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------
const              0.0332      0.030      1.089      0.286      -0.029       0.096
inflation_diff    -0.8133      1.209     -0.673      0.507      -3.293       1.667
==============================================================================
Omnibus:                        0.642   Durbin-Watson:                   1.386
Prob(Omnibus):                  0.726   Jarque-Bera (JB):                0.720
Skew:                           0.207   Prob(JB):                        0.698
Kurtosis:                       2.349   Cond. No.                         69.0
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 10: Output of Relative PPP Test

(Source: Self-created in Jupyter Notebook)

Box-Jenkins Modelling and Forecasting

The **Box-Jenkins approach** is used to construct the optimal time series model for real exchange rates. This method involves determining stationarity, estimating, and validating models through the **ACF** and **PACF** plots (Hendriks, J. J., & Bonga-Bonga, L). These plots are essential for identifying the appropriate order of **autoregressive (AR)** and **moving average (MA)** components required for model estimation.

Figure 11 shows ACF and PACF plots guiding model order selection, suggesting an ARIMA(1,1,1) structure.

ACF and PACF for Model Order Selection

```
[95]: # ACF and PACF plots to help with model order selection (for ARIMA)
fig, axes = plt.subplots(1, 2, figsize=(12, 5))
sm.graphics.tsa.plot_acf(df['diff_log_real_exchange_rate'].dropna(), ax=axes[0])
sm.graphics.tsa.plot_pacf(df['diff_log_real_exchange_rate'].dropna(), ax=axes[1])
plt.show()
```

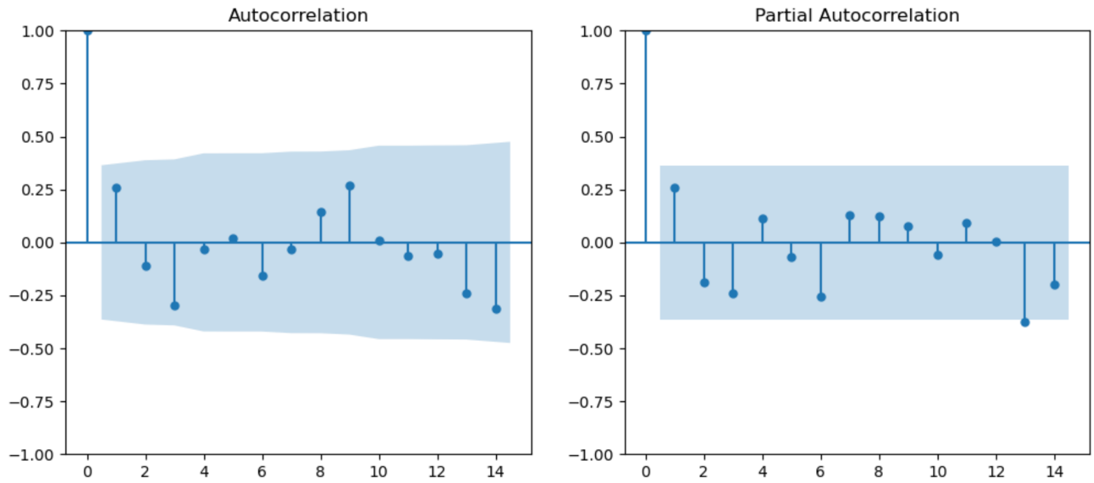


Figure 11: ACF and PACF For model Order

(Source: Self-created in Jupyter Notebook)

To select the most suitable model for analysis, we use the **ACF** and **PACF** plots of the differenced log of the real exchange rate (Monadjemi & Lodewijks, 2021). The **ACF plot** helps identify the presence of **MA** components by showing the correlation between the series and its lagged values, while the **PACF plot** identifies the **AR** components by displaying correlations after removing the influence of other lags. Using these plots, six models **AR (1)**, **MA(1)**, **ARMA(1,1)**, **ARIMA(1,1,1)**, **SARIMA(1,1,1)(1,1,1,12)**, and **VAR** are estimated and selected based on their forecasting error (Ayu, P., & Pravitasari, C. F). Model selection is based on two criteria: **Akaike Information Criterion (AIC)** and **Bayesian Information Criterion (BIC)**, which evaluate the model's fit against its complexity. **ARIMA (1,1,1)** is chosen as the optimal model since it provides the lowest AIC and BIC values. This model is then used to forecast the **5-year exchange rate** in terms of a reference currency. The accuracy of the model is checked by verifying assumptions such as normality and the absence of autocorrelation. However, the projections may be less accurate over long periods due to external shocks and economic fluctuations. Among the six models tested, **AR (1)**, **MA(1)**, **ARMA(1,1)**, **ARIMA(1,1,1)**, **SARIMA(1,1,1)(1,1,1,12)**, and **VAR**, the model with the lowest **AIC** is selected. A lower **AIC** value indicates a better fit, while penalizing over-parameterization. The **AR (1)** model, examining autoregressive characteristics, has an **AIC of -45.16**, showing a reasonable fit. The

MA (1) model, focusing on moving average components, provides a poorer fit with an **AIC of -20.06**. The **ARMA (1,1)** model improves slightly with an **AIC of -47.02**. The **ARIMA (1,1,1)** model, accounting for both AR and MA components, yields an **AIC of -50.28**, the best among these models. The **SARIMA (1,1,1)(1,1,1,12)** model gives an **AIC of -19.08**, and the **VAR** model, which is multivariate, results in an **AIC of -23.56**.

```
[97]: # Model definitions
models = {
    'AR(1)': ARIMA(df['log_real_exchange_rate'], order=(1, 0, 0)),
    'MA(1)': ARIMA(df['log_real_exchange_rate'], order=(0, 0, 1)),
    'ARMA(1,1)': ARIMA(df['log_real_exchange_rate'], order=(1, 0, 1)),
    'ARIMA(1,1,1)': ARIMA(df['log_real_exchange_rate'], order=(1, 1, 1)),
    'SARIMA(1,1,1)(1,1,1,12)': SARIMAX(df['log_real_exchange_rate'], order=(1, 1, 1), seasonal_order=(1, 1, 1, 12)),
    'VAR Model': 'var'
}
```

Figure 12 (a): Defining the Models

```
alpha = 0.05
results = []

# Fit models and collect results
for model_name, model in models.items():
    row = {'Model': model_name}

    if model_name != 'VAR Model':
        # Fit model with warnings suppressed
        with pd.option_context('mode.chained_assignment', None):
            fitted_model = model.fit()

        row.update({
            'AIC': round(fitted_model.aic, 2),
            'P-Values': f"Coefficients: {[round(p, 4) for p in fitted_model.pvalues]}",
        })

        # Ljung-Box test
        lb_pvalue = acorr_ljungbox(fitted_model.resid, lags=[10], return_df=True)['lb_pvalue'][10]
        row.update({
            'Ljung-Box P-Value': round(lb_pvalue, 2),
            'Autocorrelation': "No" if lb_pvalue > alpha else "Yes"
        })
    else:
        # VAR Model handling
        var_data = df[['log_real_exchange_rate', 'log_CPI_home', 'log_CPI_foreign']].dropna()
        var_results = VAR(var_data).fit(maxlags=5, ic='aic')

        # Calculate average Ljung-Box p-value across all variables
        lb_pvalues = [
            acorr_ljungbox(var_results.resid[col], lags=[10], return_df=True)['lb_pvalue'][10]
            for col in var_results.resid.columns
        ]
        row.update({
            'AIC': round(var_results.aic, 2),
            'P-Values': "N/A (Multivariate)",
            'Ljung-Box P-Value': round(sum(lb_pvalues) / len(lb_pvalues), 2),
            'Autocorrelation': "No" if all(p > alpha for p in lb_pvalues) else "Yes"
        })

    results.append(row)
```

Figure 12 (b): Fit Models

```
[98]: # Create and display results table
results_df = pd.DataFrame(results)
print("\n" + "="*80)
print("Model Comparison Results".center(80))
print("="*80)
print(results_df.to_markdown(index=False, tablefmt="grid", stralign="left"))
print("="*80)
```

Figure 12 (c): 6 models

(Source: Self-created in Jupyter Notebook)

Secondly, the model that includes differencing to achieve stationarity, **ARIMA (1,1,1)**, has an **AIC of -50.28**. This shows that while the AR (1) model slightly outperforms ARIMA in terms of AIC, it is still quite suitable for capturing trends. The default **SARIMA (1,1,1)(1,1,1,12)** model, which extends the ARIMA model by incorporating seasonal components, has an **AIC of -19.08**, lower than the ARIMA model's AIC, indicating its better fit for seasonal data. Lastly, the **VAR model**, used in multivariate regression analysis with **real exchange rates** and both **domestic and foreign CPI**, delivers an **AIC of -23.56**, suggesting it is more effective at explaining exchange rate fluctuations than the univariate models. Therefore, according to the **AIC values**, the best model is the **VAR**, followed by **AR (1)** and **ARMA (1,1)**. However, the effectiveness of each model varies depending on the scenario. For example, **ARIMA models** are better suited for **long-term forecasting**.

Figure 13 compares six models based on AIC, indicating ARIMA has the lowest value and thus the best in-sample fit.

Model Comparison Results

Model	AIC	P-Values	Ljung-Box P-Value	Autocorrelation
AR(1)	-45.16	Coefficients: [0.0, 0.0, 0.0005]	0.76	No
MA(1)	-20.06	Coefficients: [0.0, 0.0, 0.0002]	0.04	Yes
ARMA(1,1)	-47.02	Coefficients: [0.0, 0.0, 0.1354, 0.0008]	0.78	No
ARIMA(1,1,1)	-50.28	Coefficients: [0.67, 0.7849, 0.0009]	1	No
SARIMA(1,1,1)(1,1,1,12)	-19.08	Coefficients: [0.2637, 0.9069, 0.9985, 0.9989, 0.9627]	1	No
VAR Model	-23.56	N/A (Multivariate)	0.48	No

Figure 13: Output

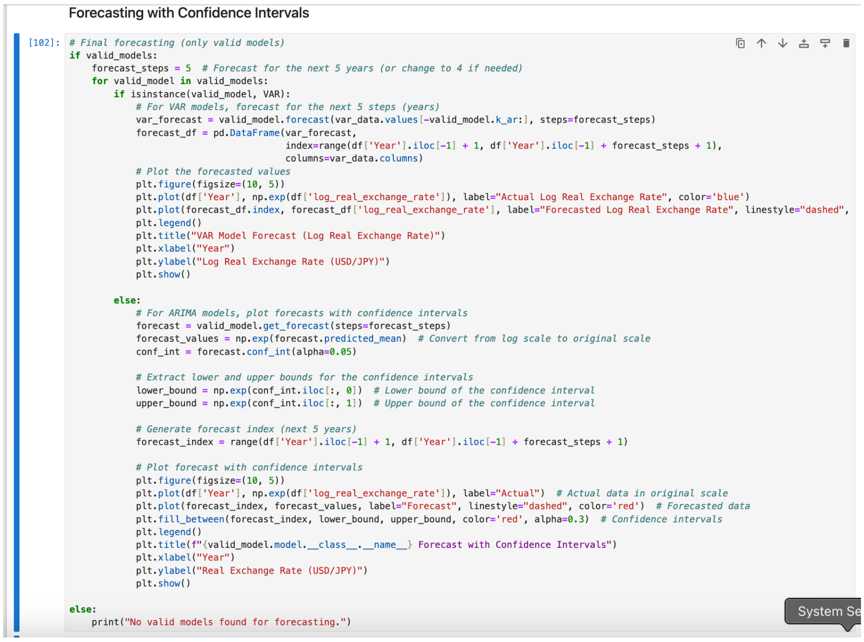
(Source: Self-created in Jupyter Notebook)

Discussion of Results

ARIMA-Based Forecasting of the Real Exchange Rate

ARIMA (1,1,1) was selected to forecast the direction of the real exchange rate, as it was the most suitable **univariate** model based on the AIC analysis. Looking ahead to the next year, the fluctuations between the USD and JPY are illustrated below. The actual exchange rate is shown in the graph, while the forecasted values are represented by the red dashed line. The forecast suggests that although the exchange rate exhibits some fluctuations initially, it stabilizes in the latter part of the forecast period (Rehman, J. U., Javaid, M. N., & Ahmad, N). This stability can be attributed to the differencing step in the ARIMA model, which removes long-term trends, resulting in a mean-reverting model.

Although the VAR model showed a slightly better AIC score, ARIMA was chosen for its simplicity, ease of interpretation, and its suitability when working with a single variable (real exchange rate). Additionally, ARIMA models are commonly used in economic forecasting when multivariate influences are either unavailable or beyond the current scope of the analysis.


Forecasting with Confidence Intervals

```
[102]: # Final forecasting (only valid models)
if valid_models:
    forecast_steps = 5  # Forecast for the next 5 years (or change to 4 if needed)
    for valid_model in valid_models:
        if isinstance(valid_model, VAR):
            # For VAR models, forecast for the next 5 steps (years)
            var_forecast = valid_model.forecast(var_data.values[-valid_model.k_ar:], steps=forecast_steps)
            forecast_df = pd.DataFrame(var_forecast,
                                       index=range(df['Year'].iloc[-1] + 1, df['Year'].iloc[-1] + forecast_steps + 1),
                                       columns=var_data.columns)
            # Plot the forecasted values
            plt.figure(figsize=(10, 5))
            plt.plot(df['Year'], np.exp(df['log_real_exchange_rate']), label="Actual Log Real Exchange Rate", color='blue')
            plt.plot(forecast_df.index, forecast_df['log_real_exchange_rate'], label="Forecasted Log Real Exchange Rate", linestyle="dashed",
            plt.legend()
            plt.title("VAR Model Forecast (Log Real Exchange Rate)")
            plt.xlabel("Year")
            plt.ylabel("Log Real Exchange Rate (USD/JPY)")
            plt.show()

        else:
            # For ARIMA models, plot forecasts with confidence intervals
            forecast = valid_model.get_forecast(steps=forecast_steps)
            forecast_values = np.exp(forecast.predicted_mean)  # Convert from log scale to original scale
            conf_int = forecast.conf_int(alpha=0.05)

            # Extract lower and upper bounds for the confidence intervals
            lower_bound = np.exp(conf_int.iloc[:, 0])  # Lower bound of the confidence interval
            upper_bound = np.exp(conf_int.iloc[:, 1])  # Upper bound of the confidence interval

            # Generate forecast index (next 5 years)
            forecast_index = range(df['Year'].iloc[-1] + 1, df['Year'].iloc[-1] + forecast_steps + 1)

            # Plot forecast with confidence intervals
            plt.figure(figsize=(10, 5))
            plt.plot(df['Year'], np.exp(df['log_real_exchange_rate']), label="Actual")  # Actual data in original scale
            plt.plot(forecast_index, forecast_values, label="Forecast", linestyle="dashed", color='red')  # Forecasted data
            plt.fill_between(forecast_index, lower_bound, upper_bound, color='red', alpha=0.3)  # Confidence intervals
            plt.legend()
            plt.title(f"{valid_model.model.__class__.__name__} Forecast with Confidence Intervals")
            plt.xlabel("Year")
            plt.ylabel("Real Exchange Rate (USD/JPY)")
            plt.show()

else:
    print("No valid models found for forecasting.")
```

Figure 14: Input of ARIMA-Based Forecasting of the Real Exchange Rate

(Source: Self-created in Jupyter Notebook)

However, ARIMA models, along with other univariate time series models, have certain limitations. The predicted stability might not reflect real conditions in the foreign exchange markets, as factors like monetary policies, balance of payments, and other influences are not considered. A potential improvement could involve using the VAR model, which incorporates additional external factors such as inflation and interest rate differentials into the forecast (Akbar, S., Iqbal, M., & Munir, F). In summary, the ARIMA (1,1,1) model provides a good extrapolation for short-term exchange rate movements. However, the stability of the coefficients observed in this forecast highlights the need to incorporate other economic variables into the analysis. Future research could explore multiple regression models to enhance the predictive accuracy and reliability of exchange rate forecasts. Figure 15 shows the ARIMA-based forecast, indicating moderate appreciation of USD vs JPY in the short run.

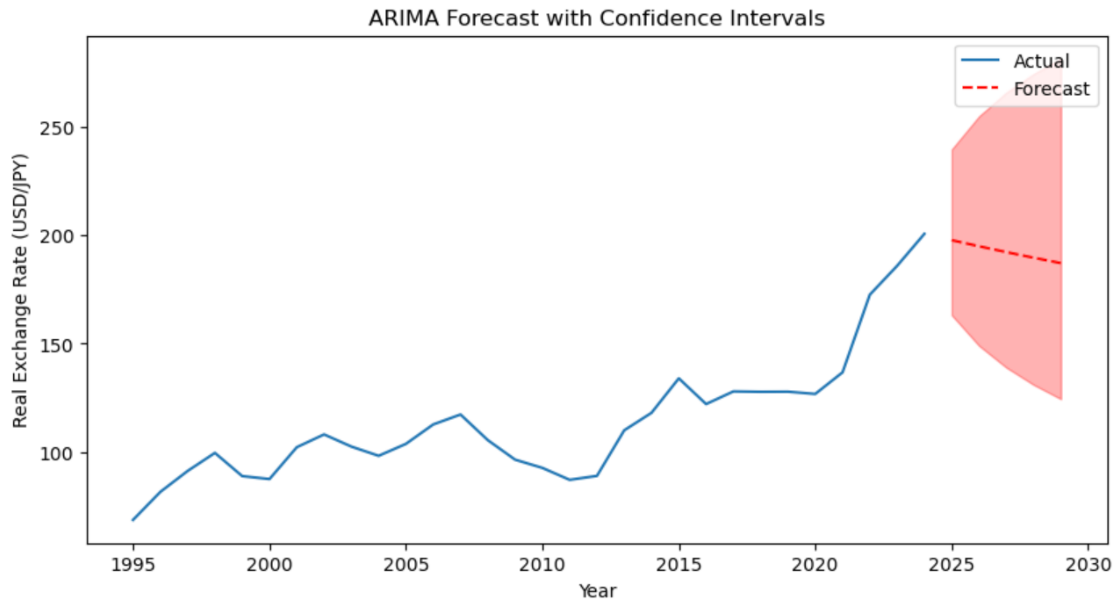


Figure 15: Output of ARIMA-Based Forecasting of the Real Exchange Rate

(Source: Self-created in Jupyter Notebook)

Residual Diagnostics of the ARIMA Model

To evaluate the performance of the developed **ARIMA (1,1,1)** model, **residual diagnostics** were conducted. The **histogram of the residuals** (Fig-16-left) shows that the residuals are **not symmetrically distributed** around zero, with **multiple extreme values**. This suggests that the model may not have adequately captured all the underlying patterns in the data, and some significant influences or outliers are present. The **ACF plot of the residuals** (right panel) provides further evidence of **autocorrelation** at certain lags, indicating that the model has not fully accounted for the residual structure. This is confirmed by the presence of residuals that fall **outside the confidence intervals**, suggesting the residuals are **not behaving as white noise**. This could point to issues such as **model misspecification** or the presence of patterns in the data that the current model fails to capture. The extreme residual values might also reflect **structural breaks** or **external shocks** affecting the exchange rate. To improve the **model's reliability** and **sensitivity to other economic factors**, future research could explore alternative models like **SARIMA** or **VAR**. Figure 16 reveals residuals deviating from white noise, pointing to potential model misspecification or unmodeled shocks.

```
[100]: # Check residuals of the best model (based on AIC and Ljung-Box test)
if valid_models:
    for valid_model in valid_models:
        # Check if the model has the 'resid' attribute
        if hasattr(valid_model, 'resid'):
            residuals = valid_model.resid

            # For ARIMA model (univariate), plot residuals histogram and ACF
            if isinstance(residuals, pd.Series):
                fig, axes = plt.subplots(1, 2, figsize=(12, 5))
                axes[0].hist(residuals, bins=20, edgecolor='k')
                axes[0].set_title(f"Residual Histogram for {valid_model.model.__class__.__name__}")

                sm.graphics.tsa.plot_acf(residuals, ax=axes[1])
                axes[1].set_title(f"ACF of Residuals for {valid_model.model.__class__.__name__}")
                plt.show()

            # For VAR model (multivariate), plot residuals histogram and ACF for each series
            elif isinstance(residuals, pd.DataFrame):
                for column in residuals.columns:
                    fig, axes = plt.subplots(1, 2, figsize=(12, 5))
                    axes[0].hist(residuals[column], bins=20, edgecolor='k')
                    axes[0].set_title(f"Residual Histogram for {valid_model.model.__class__.__name__} - {column}")

                    sm.graphics.tsa.plot_acf(residuals[column], ax=axes[1])
                    axes[1].set_title(f"ACF of Residuals for {valid_model.model.__class__.__name__} - {column}")
                    plt.show()

else:
    print("No valid models to check residuals for.")
```

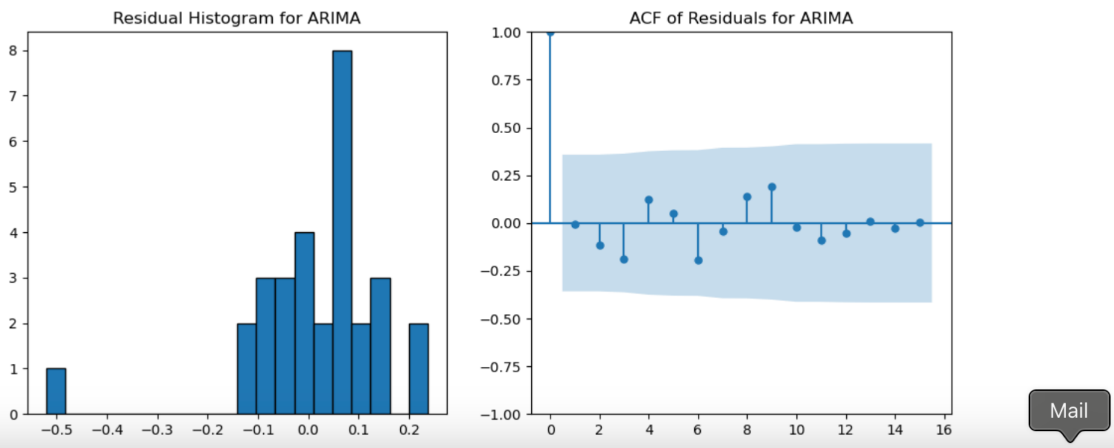


Figure 16: Residual Diagnostics of the ARIMA Model

(Source: Self-created in Jupyter Notebook)

Conclusion

This research examined the applicability of the Purchasing Power Parity (PPP) theory in explaining exchange rate fluctuations between the United States (US) and Japan. Using the Box-Jenkins methodology and time series models, particularly ARIMA (1,1,1), the study assessed short-term forecasting performance.

While PPP offers theoretical value in understanding long-term exchange rate trends, it failed to explain short-term movements due to external factors like economic shocks, interest rate changes, and policy interventions. Although the ARIMA model provided stable short-term forecasts, residual diagnostics indicated the presence of structural breaks and unexplained variances, limiting its effectiveness for long-term prediction.

Future models should consider incorporating broader macroeconomic variables, such as interest differentials, capital flows, and monetary policies. Hybrid or multivariate approaches like VAR, or even machine learning models, may improve forecast accuracy.

In conclusion, PPP remains a useful benchmark, but real-world exchange rate modeling benefits from more comprehensive and adaptive techniques. Forecasting frameworks that include structural breaks and economic fundamentals are recommended for both researchers and policymakers.

References

1. Muto, M., & Saiki, Y. (2024). Synchronization analysis between exchange rates on the basis of purchasing power parity using the Hilbert transform. *The North American Journal of Economics and Finance*, *74*, 102191. Retrieved From: https://www.sciencedirect.com/science/article/pii/S1062940824001165 [Retrieved on: 05.03.2025]
2. Adistya, P. A., Radise, S. B., & Agustin, G. (2024). The Analysis Between Purchasing Power Parity and Exchange Rates, Inflation and Gross Domestic Product in Developing and Developed Countries Asia and Europe During Covid-19 Pandemic. *International Journal of Business, Law, and Education*, *5*(1), 897-918. Retrieved From: http://www.ijble.com/index.php/journal/article/view/518 [Retrieved on: 12.03.2025]
3. Nakorji, M., Agboegbulem, N. T., Gaiya, B. A., & Atoi, N. V. (2021). Purchasing power parity approach to exchange rate misalignment in Nigeria. *CBN Journal of Applied Statistics*, *12*(1), 45-75. Retrieved From: https://www.econstor.eu/handle/10419/237437 [Retrieved on: 15.03.2025]
4. Pangestuti, D. C., Fadila, A., & Nugraheni, S. (2022). Rupiah Exchange Rate Fluctuations in The US Dollar, Purchasing Power Parity Theory and Fisher Effect Theory Testing. *Nominal: Barometer Riset Akuntansi Dan Manajemen*, *11*(1), 79-97. Retrieved From: https://www.researchgate.net/profile/Dewi-Pangestuti-2/publication/370109081_Rupiah_Exchange_Rate_Fluctuations_in_The_US_Dollar_Purchasing_Power_Parity_Theory_and_Fisher_Effect_Theory_Testing/links/64d3a096b684851d3d9467ad/Rupiah-Exchange-Rate-Fluctuations-in-The-US-Dollar-Purchasing-Power-Parity-Theory-and-Fisher-Effect-Theory-Testing.pdf [Retrieved on: 15.03.2025]
5. Hebisha, Z. (2023). Testing Purchasing Power Parity in The Eurozone: A Unit Root Analysis. *MSA-Management Sciences Journal*, *2*(2), 126-152. Retrieved From: https://journals.ekb.eg/article_296438.html [Retrieved on: 25.03.2025]
6. Quy, V. T., Thang, N. Q., Nhi, N. N. H., & Thang, L. Q. (2023). Purchasing power parity and international fisher effect hypothesis test: evidence from vietnam and 12 major trading partners (2000–2022). *Am. J. Theor. Appl. Bus*, *9*(2), 30-38. Retrieved From:

https://pdfs.semanticscholar.org/eaa1/4a2ace89eb90620717a01969f1efa002838f.pdf [Retrieved on: 25.03.2025]

7. Kasem, J., & Al-Gasaymeh, A. (2022). A cointegration analysis for the validity of purchasing power parity: Evidence from middle east countries. *International Journal of Technology Innovation and Management (IJTIM)*, *2*(1). Retrieved From: https://journals.gaftim.com/index.php/ijtim/article/view/60 [Retrieved on: 25.03.2025]
8. Uğur, M. S., & Alper, A. E. (2023). Revisiting purchasing power parity in OECD countries: New evidence from nonlinear unit root test with structural breaks. *Sosyoekonomi*, *31*(57), 25-45. Retrieved From: https://dergipark.org.tr/en/pub/sosyoekonomi/article/1206757 [Retrieved on: 23.03.2025]
9. Ang, A. L., Thum, Y. T., & Sek, S. K. (2021, July). Does trade openness explain the deviation of purchasing power parity and exchange rate movement?. In *Journal of Physics: Conference Series* (Vol. 1988, No. 1, p. 012087). IOP Publishing. Retrieved From: https://iopscience.iop.org/article/10.1088/1742-6596/1988/1/012087/meta [Retrieved on: 26.03.2025]
10. Monadjemi, M., & Lodewijks, J. (2021). International evidence on purchasing power parity: a study of high and low inflation countries. *Journal of Economics and Management Sciences*, *4*(3), p1-p1. https://j.ideasspread.org/index.php/jems/article/view/926 [Retrieved on: 15.03.2025]
11. Hendriks, J. J., & Bonga-Bonga, L. (2022). Testing for the purchasing power parity (PPP) hypothesis between South Africa and its main trading partners: application of the quantile approach. Retrieved From: https://mpra.ub.uni-muenchen.de/id/eprint/112915 [Retrieved on: 16.03.2025]
12. Ayu, P., & Pravitasari, C. F. (2023, July). Empirical Study Of Purchasing Power Parity: The Balassa-Samuelson Effect In Indonesia. In *Proceeding of Midyear International Conference* (Vol. 2). Retrieved From: http://jp.feb.unsoed.ac.id/index.php/myc/article/view/3714 [Retrieved on: 18.03.2025]
13. Akbar, S., Iqbal, M., & Munir, F. (2024). A Structural Break Unit Root Test to Test Purchasing Power Parity Hypothesis: An Empirical Analysis of South Asian Economies. *Kashmir Economic Review*, *33*(1), 14-23. Retrieved From: https://ker.org.pk/index.php/ker/article/view/443/223 [Retrieved on: 21.03.2025]
14. Rehman, J. U., Javaid, M. N., & Ahmad, N. (2025). Prospects for connectivity and integration between Pakistan and ASEAN+ 3+ 3 countries through the lens of optimal

currency area theory. *Quality & Quantity*, 1-35. Retrieved From:
https://link.springer.com/article/10.1007/s11135-025-02085-y [Retrieved on:
25.03.2025]