

Disparities in Purchasing Power Parity: A Cross-Country Examination of Economic and Social Factors

Yichen Gan^{1*}, Xianting Cheng², Jiayue Feng³

¹*German Swiss International School, Hong Kong, China*

²*Olive Tree International Academy, Hangzhou, China*

³*Zhengzhou Foreign Language, Zhengzhou, China*

**Corresponding Author. Email: easongan0824@gmail.com*

Abstract. As a method to compare living standards and overall growth in different nations, purchasing power parity is a well-researched and established economic theory. The work examines the differences in countries' economic and statistical PPP correlations by focusing on three potential causes of purchasing power parity deviations: development status, trade openness, and population size. This work selects data over fifty years and over 120 countries to investigate how these factors influence PPP deviations using unsupervised machine learning and econometric techniques. In this work, regression analysis is applied to research the effects resulting from these factors. The result shows that developed countries have fewer PPP deviations compared with developing countries, greater trade openness is instrumental in reducing PPP deviations, and small populations tend to exacerbate these deviations. Our study offers insight into managing PPP in the long run and includes all types of countries.

Keywords: Purchasing Power Parity, linear regression, unsupervised machine learning, clustering, cross-country analysis.

1. Introduction

Purchasing Power Parity (PPP) is a cornerstone of economic theory because it reflects the proposition that, given no obstructions to trade and transportation, the prices of identical goods would be the same in all countries transacted in the same currency. This principle postulates that the forex rate linking two different currencies would be the same as the ratio of the different prices (in the country), thus reflecting the economy's purchasing power, which in turn is the basis for comparing productivity and living standards across nations. Firms may benefit from employing PPP in their analysis by pointing out the disparities in the exchange rates determined by the market as well as the ones needed by their prices. Further, PPP could make outflows more of an economic threat as it measures high deviations from PPP with the notion that something is wrong with a country's economy, like inflation or imbalanced export and import levels. Familiarity with elements of these issues is the foundation for disclosure of the intricate mechanisms underlining the global economy. This paper aims to categorize various countries according to the factors that affect the PPP, aiming to uncover relations and structures that impact its stability. Although the theoretical

version of the PPP reflects certain aspects of countries' economies, this paper focuses on extrapolating more subtle connections upon inspecting real-life data. We hope to find the best alignment pairs with the least forecast errors through our results. Eventually, the aim is to define policies and establish global understandings that would enhance economic growth, ensuring that the advantages of trade and economic integration are well-spread.

2. Literature review

Purchasing Power Parity (PPP) is a foundational concept introduced by Gustav Cassel in the early 20th century, suggesting that identical goods should have the same price when expressed in a common currency, eliminating arbitrage opportunities [1]. In international macroeconomics, the theory asserts that the exchange rates between two currencies should adjust to reflect changes in the price levels of the two countries. In close relation to the PPP, there is a vital effect called The Balassa-Samuelson effect [2], which posits that countries with higher productivity growth in the tradable goods sector will experience real exchange rate appreciation. This effect neglects the impact caused by other factors, including external shocks and policies. Afterward, Kenneth Rogoff coined the term "PPP puzzle" to describe the discrepancy between the theory and empirical findings [3]. Taking a long-term view, Yin-Wong Cheung and Kon S. Lai found a tendency for exchange rates to revert to parity in the long run, though PPP deviates in the short term by applying fractional cointegration techniques [4]. The outcome cannot be that pragmatic since the statistical means cannot be adapted sensibly when data varies rapidly. Engel found that non-tradables contribute to persistent deviations from PPP because their prices are determined by domestic supply and demand conditions rather than international arbitrage [5]. This model can be too simplified as it does not need to account for the reciprocity between non-tradable and tradable sectors. Another factor is researched by Frankel and Rose [6]. They found that floating exchange rate regimes tend to exhibit larger and more persistent deviations from PPP compared to fixed or pegged regimes. However, the model assumes that the exchange rate regime is the primary driver of PPP deviations, possibly neglecting other macroeconomic factors like inflation differentials or fiscal policies. Devereux and Engel concluded that such interventions could lead to persistent misalignments and deviations from PPP by reassessing how government policies and interventions impact PPP, which is complicated and difficult to predict [7]. What is more, Luca Antonio Ricci, Gian Maria Milesi-Ferretti, and Jaewoo Lee utilize a cross-country perspective to investigate the relationship between real exchange rate and productivity and trade. This model assumes that each economy reacts similarly to some economic variables, which is too limited and simplified. Michael D. Bordo, Lars Jonung, and Mark A. K. Taylor reevaluated the long-run behavior of real exchange rates illustrated by early mixed data perhaps resulting from short-period data limitations. From all the papers, factors that can impact PPP can be comprehended. However, only a few articles are investigating a substantial number of factors simultaneously. In this work, many countries are classified by some characteristics that can affect PPP, and this paper aims to identify the countries with the smallest PPP deviation, leading to fair trade all over the world.

3. Exploratory data analysis

3.1. Data description and extraction

Nominal GDP, real GDP, nominal net exports, and imports are extracted from the Penn World Table dataset [8], a comprehensive collection of economic indicators and data for various countries, which

provides crucial information for analyzing global economic performance and trends. Data from different sources are used to categorize countries based on economic metrics. Human Development Index was extracted from the 2023-2024 human development report [9]. Mortality Data and population data were extracted from the 2022 World Bank dataset [10].

3.2. Data preprocessing

In the data preprocessing phase, the initial idea was to remove rows with more than 20% missing values. The dataset had 647 missing values out of a total of 11,848 data points, resulting in a missing rate of approximately 5.46%. Additionally, a random forest algorithm was then applied to the cleaned dataset (with removed data). However, results indicate that the two datasets showed minimal variation from regression results. Given the minimal differences observed, the original dataset was retained for analysis (missing values are skipped in the calculation).

3.3. Initial analysis

To accurately analyze relatively large datasets, it is necessary to test the reliability of the data through a method that can be manually verified. In our case, dealing with economic data, a practical way to do this is by plotting data points on a pair of familiar axes. For instance, this paper plots the inflation rates against year for different countries. The countries are selected based on how well their historical economic trends and events are documented. With clearer evidence on the internet, it is more feasible to carry out cross-examinations, which can either consolidate the validity of the dataset or help identify any discrepancies.



Figure 1. Scatterplot of UK inflation rate by year (1950–2019)

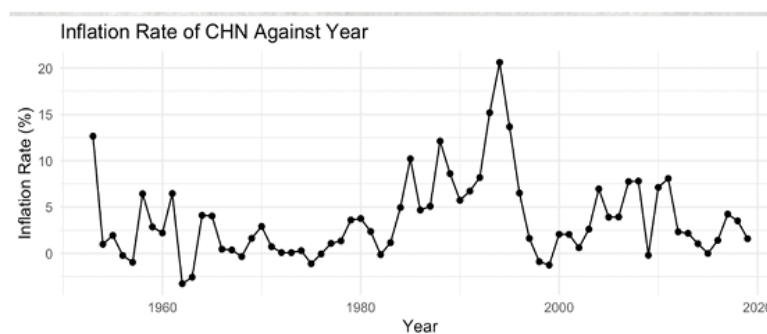


Figure 2. Scatterplot of China's inflation rate by year (1950–2019)

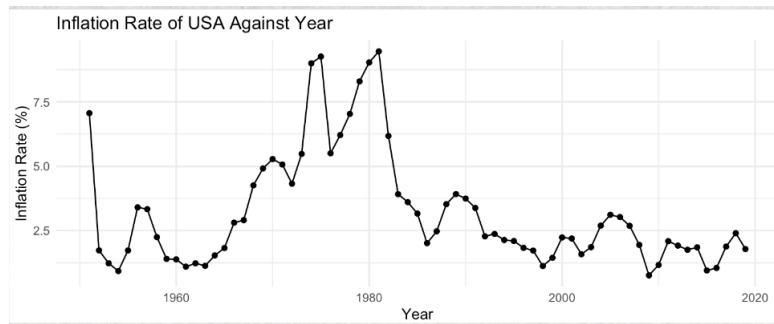


Figure 3. Scatterplot of United States inflation rate by year (1950-2019)



Figure 4. Scatterplot of Brazil's inflation rate by year (1950-2019)

As illustrated above, Figures 1, 2, and 3,4 show examples that helped support the reliability of the dataset. While figures 1 through 3 illustrate inflation rate patterns for more stable economies like the USA, GBR, and CHN (showcasing well-documented trends and responses to economic policies), figure 4 consolidates this by showing that the dataset accounts for major economic shocks and anomalies. All these figures visually display notable economic events, as well as important trends. For instance, figure 1 examines the UK inflation rate from 1950 to 2019. The early 1970s and early 1980s peaks most likely corresponded with periods of economic instability, including the oil crises, which resulted in recessions.

4. Method

After processing and conducting necessary validation checks on the dataset, this section explains how this paper explains economic factors' roles in influencing the PPP by examining statistical results. Firstly, this paper classifies countries using k-means clustering to understand what factors cause countries to be similar [11]. Then, the clusters of different countries will be analyzed before refining the classifications using targeted and more specific metrics. Finally, linear regression will be used to evaluate how well these categorized countries align with the PPP.

4.1. K-means clustering - country classification

Data points in a dataset often have underlying correlations that can be overlooked without the help of algorithms. When using the K-means clustering algorithm, the data points are used to calculate k-distinct partitions (non-overlapping subsets or clusters), where each row of data is assigned a partition based on their data values. Each cluster is represented by its centroid, the mean of all the points assigned to that cluster. The algorithm begins by initializing k cluster centroids, most of the

time chosen randomly. Each data point is assigned to the cluster whose centroid is closest to it, typically using the Euclidean distance formula. Formula (1) shows the Euclidean distance formula.

$$Distance(x, \mu_j) = \sqrt{\sum_{j=1}^d (x_j - \mu_{ij})^2} \quad (1)$$

Where x is a data point, μ_i is the centroid of the i -th cluster, and d is the number of dimensions. After all points are assigned to clusters, the centroids are recalculated as the mean of all points within each cluster. This process of assigning points and updating centroids repeats iteratively until the centroids stabilize, meaning that the assignments no longer change significantly or a predefined number of iterations is reached. The objective of k-means clustering is to minimize the total within-cluster variance quantified by the cost function. Formula (2) shows the cost function of the k-means clustering algorithm.

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_{ij}\|^2 \quad (2)$$

Here, C_i represents the set of points in cluster I , and μ_i is the centroid of cluster i . This cost function measures the compactness of the clusters, with lower values indicating more tightly packed clusters. The purpose of clustering countries is to provide insight into potential factors causing deviations from the PPP. After the countries are classified, analysis around their comparable statistical profiles can be conducted (and hopefully explain them with economic theories). For instance, if a cluster is small and deviates from the 'main' cluster with many countries, perhaps certain economic aspects they share are causing that (such as political and economic instabilities). Using k-means clustering as an initial process to filter and understand what countries have in common and what separates them, it is easier to develop better and more justified classification metrics.

Before carrying out the k-means clustering algorithm, the number of clusters (value of k) is a parameter that needs to be decided first, which involves a two-step process. Firstly, the Within-Cluster Sum of Squares (WSS) is first assessed [12]. By plotting WSS against different cluster counts, the "elbow" point is identified—adding more clusters results in only small decreases in WSS, indicating diminishing returns. This point suggests the ideal cluster number. Following this, additional statistical criteria, such as the silhouette score and Davies-Bouldin index, are employed to validate the cluster choice. These metrics assess clustering quality by measuring cluster cohesion (how closely related items are) and separation (how distinct each cluster is from others). Together, these methods ensure that the chosen clusters accurately represent the data structure, providing a solid basis for analyzing PPP deviations.

4.2. Categorical classification – country classification

As highlighted throughout this paper, grouping similar countries is essential for accurate analysis. The following section outlines the categorization criteria used to classify countries and their alignment with Purchasing Power Parity (PPP). The main categorization standards applied are as follows:

4.2.1. Development index

Countries are classified as 'Developed' or 'Developing' based on their Human Development Index (HDI). A country with an HDI greater than 0.8 is considered 'Developed.' HDI data was utilized in

three forms: the average value from 1990 to 2022, the average value with a small number of missing values, and the average value for 2019.

4.2.2. Trade openness

Based on trade metrics, countries are categorized into 'High Trade Openness' or 'Low Trade Openness'. The absolute sum of a country's real exports and imports is calculated, and then the Real GDP (RGDP) is divided by this sum to compute the trade openness index. Data from 1970 to 2019 was used for this calculation. Countries with a trade openness index greater than 0.7812432 are classified as 'High Trade Openness,' while those below this threshold are classified as 'Low Trade Openness.' The value 0.7812432 was chosen because it represents the median of the trade openness index, which effectively minimizes the impact of outliers.

4.2.3. Population

Countries are categorized as 'High Population' or 'Low Population' based on population data from 1960 to 2023. The average population per country was calculated and divided into 'large' and 'small' categories based on the median of these averages. Additionally, because the labor force is more closely related to economic factors than the total population, it was analyzed using the same median-based classification approach.

4.2.4. Health

Although health was considered a factor, it was excluded from the experimental analysis due to its limited impact on the results. Health was initially assessed by considering factors like wars and other variables. Since global conflicts have been minimal since 1960, the mortality rate per 1,000 people was used as a proxy for human health. The analysis considered countries' mortality rates with and without the influence of wars. Countries were then classified into 'High Mortality' or 'Low Mortality' based on the median of these rates (countries with rates equal to or greater than the median were classified as 'High Mortality'). With these three primary categorization standards—development index, trade openness, and population—linear regression analysis can now be employed to assess the significance and extent of these classifications.

4.3. Linear regression

In simple terms, linear regression is a statistical method for modeling and explaining the relationship between a dependent variable and one or more independent variables. This technique helps identify relationships between variables by fitting a line that best represents the data points, allowing for predictions and insights into how one variable influences another [13]. Simple linear regression is best suited to analyze cross-sectional data (data with no trends or autocorrelation). Although this paper uses panel data, the variables created will account for time series trends (mentioned later in this section). This paper is interested in the causes of PPP deviations by examining similar countries (or grouped countries). Hence, linear regression can be an agreed method to examine the consistency of this theory on different subsets of countries. The theory must be converted and represented mathematically before the PPP can be analyzed systematically using linear regression. As the PPP suggests, in the long run, exchange rates will adjust so that the price of identical goods is the same across different countries when expressed in a common currency (this means that the exchange rate between two currencies should match the ratio of their price levels). Essentially, a basket of goods

should cost the same amount in both countries when converted into a single currency. This principle ensures that the relative price levels between two countries align with the exchange rate, as described by the following equation. Formula (3) shows that the exchange rate should equal relative prices.

$$E = \frac{P}{P^*} \quad (3)$$

E indicates the exchange rate of the domestic country against the US dollar, P indicates the price level in the domestic country, and P* indicates the price level of the USA. This equation can then be rearranged to give the 'absolute theory of PPP'. Formula (4) shows the transformed version, equating to 1.

$$\frac{E \times P^*}{P} = 1 \quad (4)$$

In this equation, it is expected to see that any changes in exchange rates or price levels will return to equilibrium by changes in another variable or a combination of 2. This paper will refer to the value $\frac{E \times P^*}{P}$ as the pppvalue for simplicity. Under complete satisfaction of the absolute PPP, by plotting the PPP value against year (ranging from 1950 to 2019, depending on available data collected for each country), it is expected to see a straight horizontal line at the value of 1 (y-axis). However, this was not the case. Observations made indicated that some countries exhibited ppp values that rose/dropped as time progressed, showing some trends, as illustrated below:

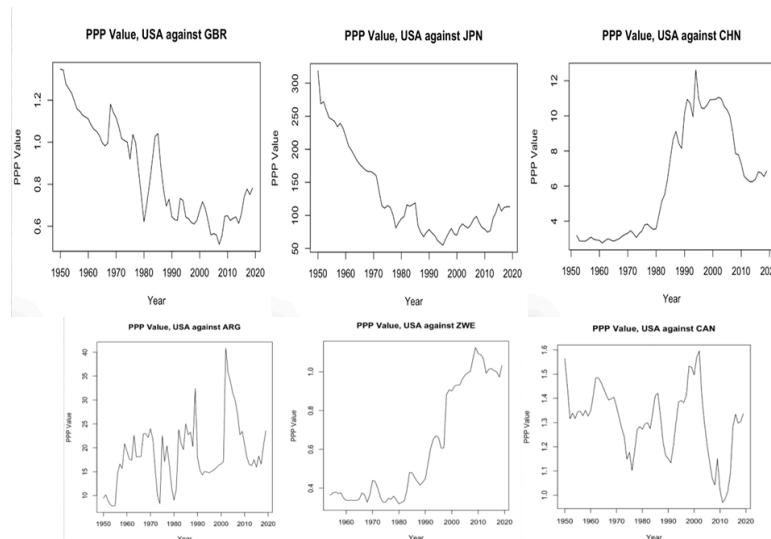


Figure 5. Scatterplot of pppvalue (against the USA) against year for the countries “GBR”, “JPN”, “CHN”, “ARG”, “ZWE” and “CAN”

Figure 5 illustrates that many countries exhibited curves with non-uniform fluctuations around arbitrary values (not 1). More importantly, there wasn't clear evidence to show the existence of this relationship. Therefore, it was necessary to alter the PPP there so it accounts for the unpredictable randomness.

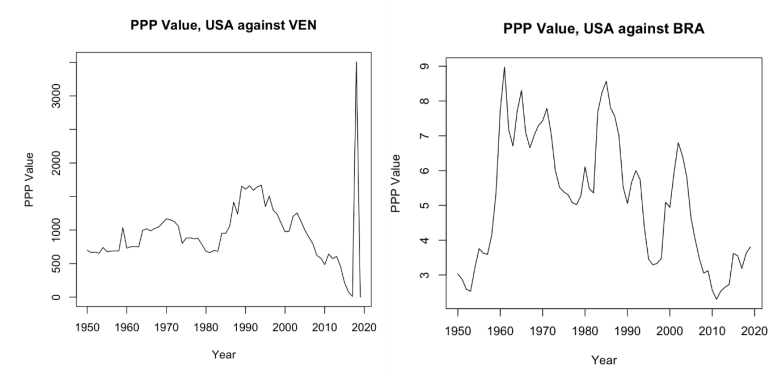


Figure 6. Scatterplot of pppvalue (against the USA) against year for the countries “VEN” and “BRA”

In Figure 6, it is worth mentioning that despite both Venezuela and Brazil experiencing Hyperinflation, Venezuela has seen more pronounced disruptions to Purchasing Power Parity (PPP), as evidenced by recent sharp peaks. Hyperinflation alone may not fully explain PPP deviations. Brazil's Hyperinflation, originating in the late 1990s, has likely been mitigated over time, resulting in less noticeable PPP deviations. In contrast, Venezuela's recent Hyperinflation, still being addressed, leads to more significant PPP deviations. Thus, while Hyperinflation is a relevant factor, its impact on PPP must be evaluated alongside other economic and policy responses for a comprehensive understanding. Therefore, to account for the random fluctuations, the formula is transformed to make the absolute PPP more flexible. For a domestic country, this paper aims to show that its pppvalues fluctuate around a constant k . Formula (5) shows the flexible version of the PPP.

$$\frac{E \times P^*}{P} = k \quad (5)$$

The natural logarithm is taken on both sides of the formula to achieve this. This transformation detrends the time-series data and facilitates linear regression by converting the data into percentage changes rather than raw trends [14]. Formula (6) shows the transformed version of formula(5) by taking the logarithm on both sides.

$$\log\left(\frac{E \times P^*}{P}\right) = \log(k)$$

$$\log(E) + \log(P^*) - \log(P) = \log(k) \quad (6)$$

Then, to eliminate the variable k (as the specificity of the value is not significant in further analysis), time is introduced to the equation. Knowing this is time series data (small explanation), two equations can be written with year t and $t-1$. Formula (7) shows the formula considering the data as a time-series data.

$$\log(E_t) + \log(P_t^*) - \log(P_t) = \log(k_t)$$

$$\log(E_{t-1}) + \log(P_{t-1}^*) - \log(P_{t-1}) = \log(k_{t-1}) \quad (7)$$

Then, because of the properties of logs and because k is assumed to be constant for our theory to hold, equation 2 is subtracted from equation 1, resulting in %change in exchange rate + %change in price levels in the USA - %change in price levels in domestic country = 0

Ultimately, the final equation is the main theory this paper will target to explain: %change in exchange rate = inflation difference (domestic country - USA). In summary, the adjusted (less absolute) PPP suggests that the percentage change in the exchange rate approximates the inflation differential between the domestic country and the USA, allowing for fluctuations around a country-specific constant K . This constant K accounts for variations due to unique economic conditions and institutional factors in each country. From the mathematical equation, three regression models are derived that will be experimented with in this paper.

Regression model 1:

$$(\text{change_xr})_t = \alpha + \beta(\text{inflation_diff})_t + \epsilon_t$$

The first regression model is formed by statistically interpreting the mathematical formula. This model can deal with time-series data because the trends and other time-series traits are accounted for. This model finds the correlation between the percentage change in exchange rates (change_xr) and the inflation differential between the domestic country and the USA (inflation_diff). In other words, this examines the more flexible version of the absolute PPP. As mentioned earlier, this regression model aims to determine if the intercept is close to zero and if the coefficient for (inflation_diff) approximates one. This regression model is used to verify the PPP for individual countries.

Regression model 2:

$$(\text{change_xr})_t = \alpha + \beta_1(\text{inflation_diff_domestic})_t + \beta_2(\text{inflation_diff_usa})_t + \epsilon_t$$

Similar to the first regression model, this model looks at time-sensitive data for individual countries. The variable (change_xr) represents the percentage change in exchange rates. Additionally, this model considers both the variables ($\text{inflation_diff_domestic}$) and ($\text{inflation_diff_usa}$). The inclusion allows us to separately account for the inflation differential specific to the USA (this is possible because of the logs taken in the mathematical derivation). This model is particularly useful for understanding how changes in exchange rates are influenced by the inflation dynamics in both countries rather than merely examining the difference in inflation rates alone. Although this model is better at accounting for factors that might influence the PPP in both countries, it introduces many difficulties in succinctly assessing the exact correlation to the PPP as a whole. Therefore, this regression model is not used in the following experiment.

Regression model 3:

$$(\text{mean_change_xr})_i = \alpha + \beta(\text{mean_inflation_diff}) + \epsilon_i$$

Finally, this third regression model looks at the mean of the previously mentioned variables. (mean_change_xr) Moreover, (mean_inflation_diff) represents the average values of exchange rate changes and inflation differentials over a specified period, respectively. Instead of looking at time, the subject of this model is the selected countries (as represented by i). This model evaluates how well the average percentage change in exchange rates correlates with the average inflation differential. This allows investigation into different periods, helping to understand the long-term relationship of the PPP.

To reiterate, the three regression models mentioned above are not typical predictive models. In this paper, the focus is on the coefficient of the regressor (mean_inflation_diff). As the y-value and the x-value are expected to have a linear relationship (in the mathematical formula derived above), the model is accurate when the coefficient is close to 1 and the y-intercept is close to 0.

As stated, this paper examines time-series data using linear regression models to explore the relationship between the percentage change in exchange rates and the inflation differential between a selected country and the USA. This approach serves as a proxy for evaluating the accuracy of the Purchasing Power Parity (PPP) theory. To ensure the validity of the regression results, it is essential to meet key assumptions underpinning Ordinary Least Squares (OLS) and the Gauss-Markov theorem. These include verifying linearity, ensuring residuals are normally distributed, confirming homoscedasticity (constant variance of residuals), and satisfying basic assumptions of independence and identically distributed (IID) errors. Addressing these assumptions is crucial for obtaining reliable and unbiased results and will be detailed in the subsequent diagnostics [15].

5. Experimental results and discussions

This section will explain the analysis behind improving the approach to assessing the causes of PPP deviations, using different regression models and analyzing different types of countries/categories.

5.1. Regression model analysis

The first regression model analyzes individual countries with the most available data over a specified period. This approach helps us understand deviations from PPP and evaluate the theory's accuracy from the perspective of one country. To ensure comprehensive results, countries are selected with varying traits, including continental differences, development disparities, and historical influences. From examining individual countries without transforming the data, the general trend shows weak correlations to the PPP. Firstly, extraordinarily high R^2 values and residual variances indicate the presence of outliers and high-influence points, likely caused by factors such as unstable economies and inept governmental policies, including Hyperinflation. These unusual observations are important because they can skew the R^2 value, creating a spurious correlation that appears much stronger than it is. Additionally, they may disrupt the normal distribution of residuals, which is a serious issue when using linear regression models. Finally, there are also indications of heteroscedasticity (detailed in section 5.3), which can skew results and make confidence intervals and hypothesis tests less reliable. Nevertheless, despite the influence of unusual observations on the model accuracy, they are handled cautiously in this paper and are not automatically removed in this analysis. Unusual observations (outliers or high leverage points) can sometimes provide valuable insights into economic conditions that might have been otherwise overlooked; this can potentially help devise better approaches to categorizing countries/explaining results. To summarize, the first regression model, which examines individual countries using annual data, is highly susceptible to

the distortion caused by outliers. It is necessary to analyze longer data periods, where temporary influences are mitigated as time progresses.

5.2. Unusual observations

Knowing that outliers and high leverage points can significantly distort regression model results (as they violate key OLS and Gauss-Markov assumptions), it is necessary to identify and address these observations (especially in countries with political or economic instability); this will prevent disruption and generation of spurious results that deviate away from the PPP [16]. To identify these unusual observations mathematically, Cook's distance was used, a method that measures a data point's influence on a linear regression model. It combines leverage (how far an independent variable deviates from its mean) and residuals (the difference between observed and predicted values) into what is known as Cook's distance (D_i) to identify observations that significantly impact the regression coefficients [17]. Cook's distance (D_i) is calculated as Formula (8) shows the formula for cook's distance.

$$D_i = \frac{\sum_{j=1}^n (\hat{y}_j - \hat{y}_{j(i)})^2}{p \bullet MSE} \quad (8)$$

Formula (8) shows the formula for the cook's distance. Where $\hat{y}_j$ is the predicted value, $\hat{y}_{j(i)}$ is the predicted value with the i - th observation removed, p is the number of parameters in the model, and MSE is the mean squared error. Typically, a Cook's distance greater than $4/n$ (where n is the number of observations) indicates a highly influential point. After modeling the dataset with the third regression model, This algorithm is applied to our paneled dataset (with all countries) over different periods (1970-2019, 1980-2019, 1990-2019). The results indicate that some countries, as indicated in Table 1, appear repetitively as outliers or high leverage points (individually or in subsets); this recurrence indicates that some economic/political challenges within these economies are persistently causing PPP deviations.

Table 1. Analysis summary of unusual observations

Country	Key Events/Characteristics	Impact on PPP
Angola (AGO)	Oil Price Crash (2014-2016), Civil War (1975-2002)	Significant economic instability, leading to deviations from PPP
Bolivia (BOL)	Political Crisis (2019), Hyperinflation (1980s)	Disrupted economic equilibrium, affecting PPP alignment
Brazil (BRA)	Hyperinflation (1990), Economic Recession (2014-2016), Political Scandals (2014-present)	Fluctuating economic policies and instability, causing PPP deviations
DR Congo (COD)	Second Congo War (1998-2003), Hyperinflation (1990s)	Conflict and economic mismanagement leading to deviations from PPP
Nicaragua (NIC)	Political Crisis (2018), Sandinista Revolution (1979)	Political upheaval contributing to deviations from PPP
Peru (PER)	Economic Reforms (1990s), Internal Conflict (1980s-1990s), Commodity Price Shocks (2000s)	Economic Fluctuations Affecting PPP Adherence

These countries are removed from the dataset due to their anomalous economies and historical context. While examining these countries, it is obvious that the theory of PPP is limited in extreme conditions. Hence, in the holistic approach of this paper, it is necessary to prevent these countries

from disrupting the analysis. Additionally, it is worth mentioning that although Brazil was identified as an outlier (using Cook's distance algorithm), Brazil still strongly correlated with the PPP when analyzed individually with annual data. This implies that PPP holds validity and lack thereof, depending on whether countries are assessed individually or in aggregate. In other words, some countries may act as 'relative' outliers influenced by specific events. Therefore, it is necessary to analyze the collective trends (paneled countries) and individual discrepancies in the analysis of PPP.

5.3. Regression diagnostics

After excluding these unusual observations, improvements surrounding Gauss Markov's OLS assumptions should be expected. Therefore, an extensive analysis of the residuals was conducted. In summary, a Q-Q plot was employed to check the normality of the residuals, and a plot of residuals against fitted values was used to assess the linearity of the regression model. By verifying that these assumptions are met, the regression model results will be more reliable and accurate. As mentioned in section 4.3, the Gauss-Markov theorem assumes that residuals from the regression model are normally distributed, exhibit constant variance, and have a linear relationship with the fitted values (otherwise, a non-linear relationship in the model is suggested, i.e., quadratic or cubic); any violations of these assumptions may lead to unreliable or biased regression results [18]. As illustrated in Figure 7, the Q-Q plot on the left (before removing unusual observations) is rather flat and close to paralleling the x-axis. This suggests that the residuals are not normally distributed. On the other hand, the plot on the right (after removing unusual observations) showed a higher resemblance to a slanted 45-degree line, indicating the dataset's univariate normality.

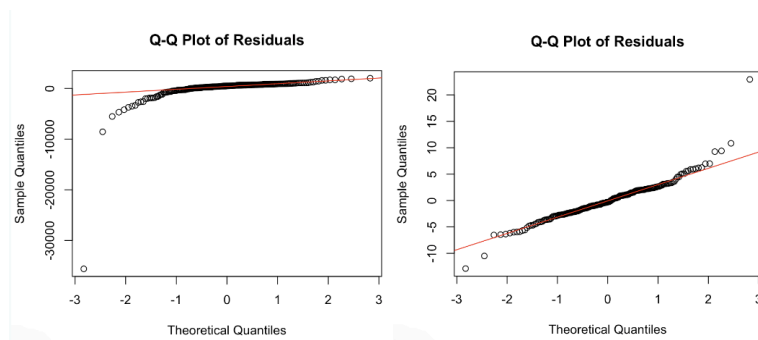


Figure 7. Q-Q plot of residuals before (left) and after (right) removing unusual observations

Additionally, Figure 8 shows the differences in the plot of residuals against fitted values before and after removing unusual observations. From the plot on the right of the figure, it is visible that there are no obvious patterns that suggest non-linearity, as they are randomly scattered around the x-axis (as opposed to the plot on the left, which shows a rather vertical pattern between the datapoints, suggesting not only non-linearity but also heteroscedasticity).

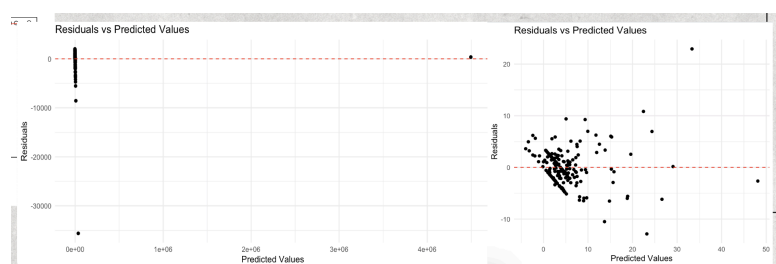


Figure 8. Scatterplot of residuals vs predicted values before (left) and after (right) removing unusual observations

Finally, heteroscedasticity is one of the other problems noticed when regression models were fitted with individual countries initially (e.g., Brazil, Japan). The Breusch-Pagan test was used to determine whether the variance of the residuals was constant [19]. As the results indicated, the BP value was 47.278, suggesting an extremely small p-value of 6.158×10^{-12} (way below the significance level of 0.05), which suggests that the null hypothesis that homoscedasticity is present should be rejected. To solve this problem, weighted least squares (WLS) regression was applied (make sure predictors have minimum variance and are unbiased) [20]. In short, WLS regression assigns different weights to data points based on their variance, minimizing the sum of weighted squared residuals. After applying this algorithm, the BP value decreased to 0.1722, meaning that the p-value is not 0.6782, which is higher than the significance value (0.05). This means that the null hypothesis is accepted that homoscedasticity is present, or heteroscedasticity is no longer an issue.

5.4. Time as an influence

After removing unusual observations and making sure assumptions are satisfied, experiments can be conducted based on analyzing outstanding countries' economies. As mentioned earlier, historical context has a great impact on the establishment of the PPP. More importantly, it is essential to determine the time being examined (the interval over which the mean is calculated for the third regression model, i.e., 1980-2019) to ensure reliability and consistency in the following categorical analysis (section 5.6). In this paper, the optimal time frame is the one with minimal economic disturbances (major events such as the 1973 and 1979 oil shocks, the 1971 collapse of the Bretton Woods system, the 2008 Financial Crisis, China's rise in the early 2000s, the 2010 European debt crisis, and the 2014 oil price crash are considered outliers that may skew PPP analyses). This is done using the third regression model across various periods.

Table 2. Regression analysis results across various time periods

Time Period	Intercept	Coefficient (mean_inflation_diff)	R ²
1970-2019	-3076.937(*)	186.082(***)	0.719
1980-2019	4.266	1.062(***)	0.751
1990-2019	4.744	0.990(***)	0.601
2000-2019	2.127	1.051(***)	0.281
2010-2019	-741.6681	259.6480	0.9999
2015-2019	2.729 (***)	1.285 (***)	0.745
2000-2010	291.690	1.390	0.000

Then, by analyzing the statistical significance of the results shown in Table 2, we conclude that although the PPP is expected to hold stronger in longer periods of time, it depends on the overall stability of economies. From 1970 to 2019, the regression results show a high coefficient for mean inflation differences (186.082) and a significant intercept (-3076.937), with an R^2 of 0.719. This suggests that the economic shocks and policy changes in the 1970s heavily influenced the relationship between inflation differences and PPP, explaining 71.9% of the variance. Narrowing the focus to 1980-2019, the results indicate a negligible intercept (4.266) and a significant coefficient (1.062), with an improved R^2 of 0.751, suggesting a stable and strong relationship between inflation differences and PPP during this period. This model explains 75.1% of the variance, highlighting that the selected time frame effectively captures the PPP dynamics. From looking at this, we can infer that the period 1970-1980 is likely the cause of PPP deviations.

Following the 1980-2019 period, the PPP becomes gradually weaker. Until the 2010-2019 period, displaying an extremely high coefficient (259.648) and R^2 (0.9999), implying an almost perfect fit; however, the negative intercept (-741.6681) and high values indicate potential overfitting or the influence of specific events within this short timeframe. Additionally, more recent periods (2015-2019) also exhibited a strong and significant relationship with a high R^2 of 0.745, indicating that the recent years still capture the PPP dynamics well. However, by looking at the intervals before (2000-2010), it argues that perhaps this strong correlation is spurious and that it is not a result of economic stability but rather other omitted factors. To summarize, although the recent data from 2015-2019 suggests a strong short-term relationship, the near-perfect fit in the period of 2010-2019 (and 2000-2010) raises concerns about overfitting or the influence of specific recent events. Thus, the period 1980-2019 is selected, providing a robust foundation for understanding the long-term relationship of the PPP through our selected variables.

5.5. Cluster analysis

As explained in detail in section 4.1, this paper first focuses on deciding the number of clusters for our dataset.

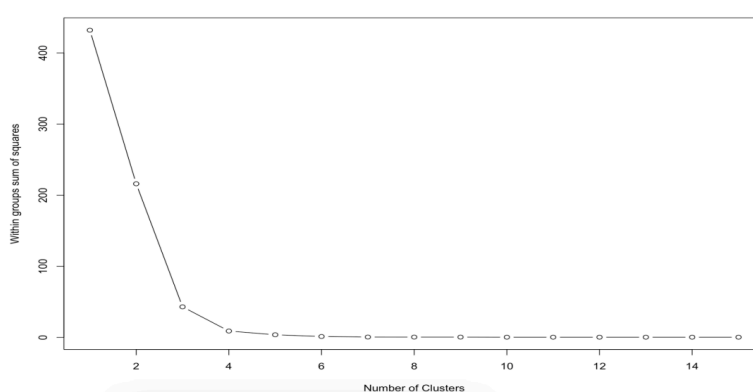


Figure 9. Plot of the comparison between the sum of squares within the group and the number of extracted clusters

Using the elbow method, a plot of the Within-Cluster Sum of Squares (WSS) against the number of clusters was generated to determine the optimal cluster count. As shown in Figure 9, the WSS decreases rapidly from one cluster to three clusters, indicating a significant improvement in cluster cohesion. Beyond three clusters, the reduction in WSS becomes marginal, suggesting that additional

clusters contribute only incrementally to the overall clustering quality. This sharp drop, followed by a gradual decline, supports the recommendation to select three clusters as the optimal solution (see Figure 10). This choice balances the trade-off between the number of clusters and the explanatory power of the clustering solution.

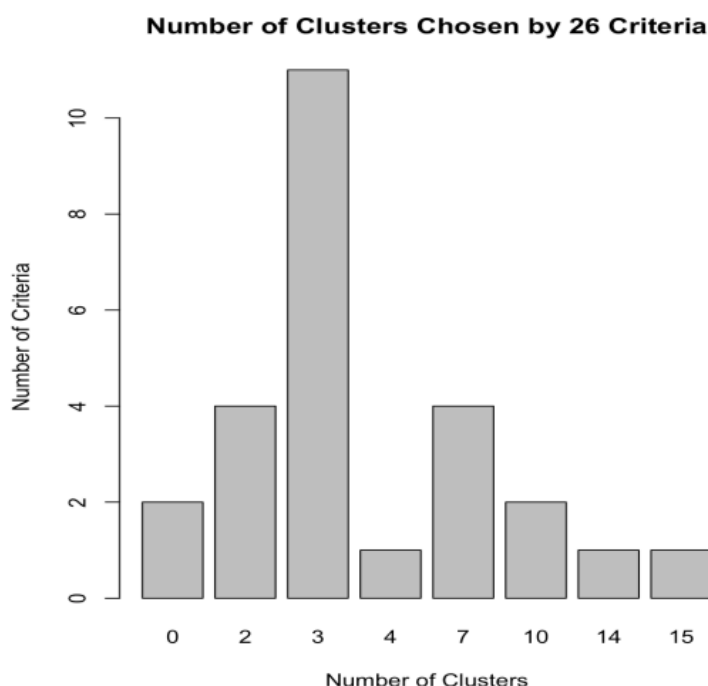


Figure 10. The number of clusters recommended using the 26 indicators in the NbClust package

Countries are segmented based on their economic data using the K-means clustering algorithm (using 3 clusters). The clustering results predominantly revealed a single large cluster encompassing most countries, with only a few countries classified into smaller, distinct clusters. The countries in these smaller clusters were thoroughly investigated to understand the unique economic factors contributing to PPP deviations. The following analysis explains how specific variables or conditions that deviate from the general pattern are decided and evaluated.

Cluster 1 includes most countries (206 countries, excluding the removed unusual observations). This indicates that there is something 'similar' statistically about their economies. In other words, their economic conditions should be aligned. However, not all countries are consistent with the PPP (necessitating further classification); otherwise, the PPP should be evident from the start when analyzing the paneled data.

Cluster 2 consists of Vietnam, Indonesia, Iran, South Korea, and Colombia. The K-means clustering algorithm suggests that these countries are statistically inconsistent with the countries in cluster 1 (perhaps due to their unique economic conditions). To summarize the potential causes of these deviations, countries are examined individually. Firstly, Vietnam's rapid economic growth and heavy reliance on manufacturing and exports causes gravitation away from a stable economy. Similarly, Iran and Columbia face economic challenges due to international sanctions and a heavy reliance on oil exports (Iran) and commodities such as oil and coffee (Columbia). This implies that the balance and stability of exports and imports (within the countries' financial capabilities; otherwise, they could be in debt) are essential to satisfy the PPP. Secondly, Indonesia's significant regional disparities and dependence on natural resources lead to variability. This may suggest that

the distribution of wealth may be a factor that causes PPP deviations, and as the population increases, this is likely bettered (as well as health and the reduction of wars).

Cluster 3 includes only Venezuela. This cluster is significant as it only includes one country. It not only suggests that this country is statistically anomalous but also implies that its economic history is vastly important (to help understand the economic causes of PPP deviation). Venezuela has been experiencing Hyperinflation (2010s-present) and Currency Controls (2003-present). Venezuela has yet to have time to stabilize or recover. This consistent instability led to extreme economic policies and PPP deviations. Additionally, it might be interesting to look at why Cook's distance (section 5.2) did not classify Venezuela as an unusual country. This might be because the recent economic turmoil has been evened out (as the mean was taken for a given period) during the calculation process. This suggests that perhaps Venezuela would show a stronger correlation with the PPP if data before the 2000s was examined).

To conclude, this section provides insight into the major economic reasons that might cause PPP deviations; however, it is not specific enough to categorize countries. Significant PPP deviations in smaller clusters suggest that outlier countries like Vietnam, Iran, Indonesia, and Venezuela have unique economic factors—such as unbalanced exports, regional disparities, and reliance on volatile or specialist exports—contributing to these deviations. Therefore, to refine our categorization of the countries, the focus is shifted towards further distinguishing within this broad cluster, partitioning this large cluster into more specific subgroups based on observed economic factors (such as differentiating between developed and developing economies).

5.6. Categorical analysis

To better understand the economic factors leading to deviations from the PPP, countries are analyzed by categorizing them into specific groups. Using the metrics from section 4.2 and the rationale for variable selection (obtained by analyzing countries in clusters 2 and 3), this section presents a statistical and economic analysis. Table 3 summarizes the regression results for each classification:

Table 3. Statistical results of the influence of each research variable on PPP deviations

Categories	Influence of Research Variables on PPP					
	Development Index		Trade openness		Population	
	Developed	Developing	High trade openness	Low trade openness	Large population	Small population
Statistical measurements						
Intercept	-1.174** *	-4277.400	-33.601	132.684	-1.751	-3.688
Coefficient	1.138***	189.600** *	3.783***	5.478	1.365***	2.190***
Inflation_diff						
R^2	0.961	0.729	0.507	0.010	0.940	0.317
Adjusted R^2	0.960	0.726	0.502	-0.002	0.939	0.310

Note: These results are generated using the third regression model over the time period 1980-2019. Countries with high influential points are removed collectively to ensure data coherency. Additionally, the coefficient of inflation_diff measures the slope of the

regressor. The stars next to each data point indicate how significant the value is according to the standardized t-test.

As mentioned earlier, when assessing how well the countries fit, the model depends on the intercept and the coefficient of the regressor. The intercept and coefficient of the regressor explain the relationship between each variable X and Y , such as whether it is positive or negative and whether the variables are significant or not. In other words, the closer the coefficient is to 1, and the closer the intercept is to 0, the better fit is explained. The following analysis will strictly follow this premise. Additionally, to evaluate the fit of each country's data to the model, we use the adjusted R-square, which indicates the effectiveness of the fit. Adjusted R-square was focused rather than R-square because the adjusted R-square formula accounts for overfitting, offering a more accurate measure of model performance.

5.6.1. Development index

As the plot on the left of Figure 11 suggests, the results indicate that developed countries exhibit a stronger fit to the regression model, with an adjusted R^2 of 0.960 and a regressor (inflation_diff) coefficient of 1.138. This suggests that the pppvalue fluctuates around a constant k , for all countries in the category. In contrast, the plot on the right of Figure 11 suggests that the fitted regression model justifies for less data points, as developing countries show a much higher coefficient of 189 and a lower adjusted R^2 of 0.726, indicating greater variability and less predictive accuracy.

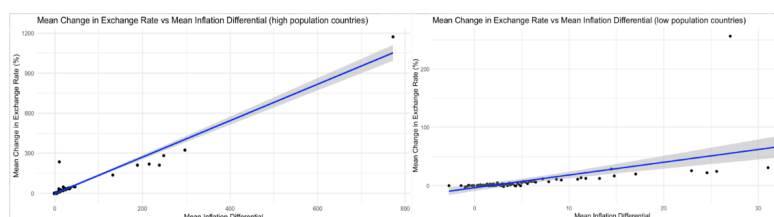


Figure 11. Scatterplot of countries categorized by development (x and y variables from the third regression model), developed countries on the left and developing countries on the right

These results can be explained as developed countries typically function within a more stable economic environment. This results in higher production efficiency, optimizing resource utilization, and sustained economic growth. Additionally, these countries experience lower and more stable inflation rates, which help maintain the purchasing power of currency and reduce economic volatility. Greater price stability, with less fluctuation in prices, thereby enhancing market predictability and consumer confidence [21].

5.6.2. Trade openness

The classification of countries by 'high' and 'low' trade openness revealed minimal differences. The points are rather randomly scattered, showing little or no correlation with the fitted model. There are many highly influential points that influence the slope of the line, causing deviations from our hypothesis. Although 'high trade openness' countries had coefficients and intercept closer to expected values, the model still showed poor fit, indicated by a low adjusted R^2 . Additionally, the insignificance of the coefficient for low trade openness countries suggests that trade openness alone might not be a strong predictor of PPP deviations, as it fails to account for other influential factors that may better explain the variation.

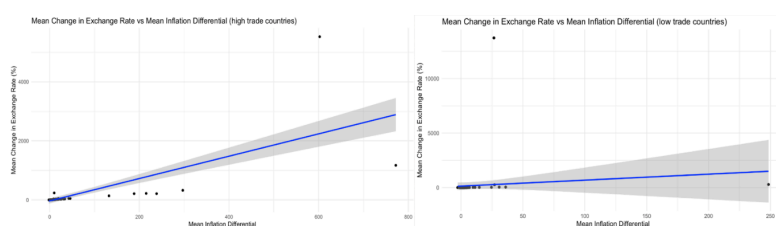


Figure 12. Scatterplot of countries categorized by trade openness (x and y variables from the third regression model), high trade openness countries on the left, and low trade openness countries on the right

In simple terms, the expansion of free trade facilitates the flow of goods and services across borders. This results in a reduction in commodity price differentials through the balancing of supply and demand across nations, which contributes to price convergence on a global scale. In an enlarged market scope, it provides producers with broader sales channels and increasing overall economic efficiency.

5.6.3. Populations

Classifying countries by population size revealed more significant differences. 'High-population' countries showed less PPP deviation, reflected by a stronger model fit and higher adjusted R^2 , suggesting more consistent and predictable economic factors, and vice versa. Similar to the development index classification, both coefficients are significant, indicating that population size, like development status, plays a crucial role in explaining PPP deviations. Larger and more developed economies tend to exhibit more stable and predictable inflation dynamics. Notably, both regressions show fewer outliers, reflecting lower variability and greater consistency with this metric (see Figure 13).

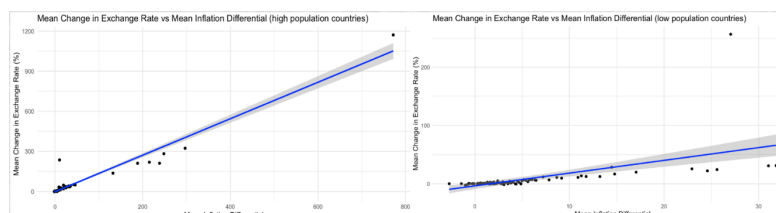


Figure 13. Scatterplot of countries categorized by population (x and y variables from the third regression model), high-population countries on the left and low-population countries on the right

Economists argue that nations with larger populations create extensive internal markets, providing a robust foundation for both consumption and production. As a result, prices within these markets tend to be more stable due to diversified demand and the benefits of economies of scale. Also, the larger market capacity offers enterprises continuous growth opportunities, which contributes to long-term economic stability.

6. Discussions

6.1. Main discussions (advantages and disadvantages)

This study examined the influence of temporal factors on the correlation between Purchasing Power Parity (PPP) and inflation disparities, analyzing data over various time intervals. It strategically

focused on economically stable systems while excluding countries with atypical economic conditions to ascertain the veracity of PPP over extended durations. Findings indicate that PPP is more likely to be valid over longer periods, particularly when the analysis is confined to countries with similar economic profiles, where a higher correlation with PPP was observed during the same time frames. This underscores the critical importance of selecting appropriate temporal scopes and homogeneous country groups to reduce the influence of outliers and economic volatility, thus enhancing the accuracy of insights into the long-term dynamics between inflation and PPP. Furthermore, the study highlighted the necessity of considering a spectrum of factors, rather than isolating individual variables, to understand the mechanisms underpinning PPP.

6.2. Limitations & future works

Unpredicted economic shocks deal with impacts that cannot be explained through a linear analysis, especially when time and severity are unknown factors. As exemplified by Venezuela, the aftermath of periods of economic instability likely results in PPP deviations. However, it is statistically flawed and can be improved. To address these impacts, the concept of a lagged regression model (introduce lags using the Akaike Information Criterion (AIC)) can be incorporated into future research [22]. In essence, it accounts for the delayed economic effects and calculates the statistical results based on this. Additionally, our model encountered challenges in integrating variables and assessing them collectively, which, to a certain extent, limits our ability to explore the complex interactions among multiple factors. Therefore, to improve the somewhat limited research, it is recommended that future research employ alternative analytical methods such as panel data or time series analysis. Moreover, integrating previously omitted variables could also lead to a more thorough analysis. For example, using dummy variables to differentiate between groups of countries within a single analytical model can help enhance the understanding and identification of group-specific effects and interactions, which could, in turn, inform more precise conclusions [23].

7. Conclusion

Through the outcomes obtained from statistical methods, this work proves that a specific arrangement of factors can contribute to PPP deviations. First, whether the country is developed is crucial to PPP value because developed countries operate in a more stable economic environment, indicating more stable prices, higher productivity efficiency, and lower and more stable inflation rates [24]. Secondly, countries with higher trade openness can reduce PPP deviations since free trade increases the flow of goods and services between countries, reducing commodity price differentials by balancing supply and demand across borders. This leads to more uniform prices globally. Larger populations create larger internal markets, providing a broader base for consumption and production. Prices within these markets are more stable due to diversified demand and economies of scale. Based on these conclusions, this paper aims to propose some policies that can be implemented to restore PPP deviations. For developing countries, governments should implement policies that stimulate economic growth, such as investing in education, research and development, infrastructure, increasing productivity, and reducing PPP deviations. Some trade-restriction countries need to encourage trade liberalization, such as reducing tariffs and eliminating trade barriers, which can be instrumental in entering the global market. Governments should enhance population health. Governments should foster regional economic collaboration and maintain monetary and fiscal policies to control inflation and currency fluctuations.

References

- [1] Cassel, G. (1918). Abnormal Deviations in International Exchanges. <https://doi.org/10.2307/2223329>.
- [2] Samuelson, P.A. (1964). Theoretical Notes on Trade Problems. <https://doi.org/10.2307/1928178>.
- [3] Rogoff, K. (1996). The Purchasing Power Parity Puzzle. <https://scholar.harvard.edu/rogooff/publications/purchasing-power-parity-puzzle>.
- [4] Cheung, Y. - W. and Lai, K.S. (1993). A Fractional Cointegration Analysis of Purchasing Power Parity. <https://doi.org/10.2307/1391310>.
- [5] Engel, C. (1993). Real exchange rates and relative prices: An empirical investigation. <https://ideas.repec.org/a/eee/moneco/v32y1993i1p35-50.html>.
- [6] Frankel, J.A. and Rose, A.K. (1996). A panel project on purchasing power parity: Mean reversion within and between countries. <https://ideas.repec.org/a/eee/inecon/v40y1996i1-2p209-224.html>.
- [7] Devereux, M.B. and Engel, C. (2003). Monetary Policy in the Open Economy Revisited: Price Setting and Exchange-Rate Flexibility. <https://www.jstor.org/stable/3648623>.
- [8] University of Groningen (2017). The Database | Penn World Table | Productivity | University of Groningen.
- [9] Nations, U. (2024). Human Development Report 2023-24. <https://hdr.undp.org/content/human-development-report-2023-24>.
- [10] World Bank (2022). World Development Report 2022: FINANCE for an Equitable Recovery.
- [11] Banoula, M. (2023). K-means Clustering Algorithm: Applications, Types, and Demos <https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm#:~:text=K%2DMeans%20clustering%20is%20an>.
- [12] Kabacoff, R. (2015). R in action : data analysis and graphics with R. Shelter Island, Manning.
- [13] Bertsimas, D. (2017). 2 Linear Regression | The Analytics Edge | Sloan School of Management. <https://ocw.mit.edu/courses/15-071-the-analytics-edge-spring-2017/pages/linear-regression/>.
- [14] Andy (2019). Logarithmic Transformation in Linear Regression Models: Why & When.
- [15] Weisberg, S. (2014). Applied linear regression. Chichester: Wiley-Blackwell.
- [16] Lecture 20: Outliers and Influential Points. (n.d.). Available at: <https://www.stat.cmu.edu/~larry/=stat401/lecture-20.pdf>.
- [17] mathworks (n.d.). Cook's Distance - MATLAB & Simulink. <https://www.mathworks.com/help/stats/cooks-distance.html>.
- [18] Greene, W.H. (2003). Econometric Analysis. Pearson Education, India.
- [19] Wikipedia Contributors (2020). Breusch-Pagan test. https://en.wikipedia.org/wiki/Breusch%E2%80%93Pagan_test.
- [20] PennState: Statistics Online Courses. (n.d.). 13.1 - Weighted Least Squares | STAT 501.
- [21] Atkeson, A. and Burstein, A. (2007). Pricing-to-Market in a Ricardian Model of International Trade. <https://doi.org/10.1257/aer.97.2.362>.
- [22] mathworks (n.d.). Time Series Regression VIII: Lagged Variables and Estimator Bias - MATLAB & Simulink Example.
- [23] Wikipedia Contributors (2019). Dummy variable (statistics). [https://en.wikipedia.org/wiki/Dummy_variable_\(statistics\)](https://en.wikipedia.org/wiki/Dummy_variable_(statistics)).
- [24] Atkeson, A., & Burstein, A. (2007). Pricing-to-Market in a Ricardian Model of International Trade. <https://doi.org/10.1257/aer.97.2.362>