Determining Factors that Influence the Botswanian Pula and Their Respective Significance

PHILIP LINDGREN

RASMUS WICKLUND
Determining Factors that Influence the Botswanian Pula and Their Respective Significance

RASMUS WICKLUND
PHILIP LINDGREN
Determining Factors That Influence the Botswanan Pula and Their Respective Significance

Rasmus Wicklund       Philip Lindgren

May 23, 2018

Abstract

In this study, the Botswanan currency Pula is analysed and factors affecting the exchange rate between the Pula and the US Dollar are examined. Using a multiple linear regression analysis and the Cochrane-Orcutt transformation, an explanatory model is developed. Factors are chosen based on economic theory and are evaluated empirically through statistical methods. The final model shows that The Botswana Stock Index, The Real Interest Rate Spread Between Botswana and USA, Corruption Index, Business Confidence Index, Diamond Price Index and Unemployment Rate, are relevant and significant in determining changes in the exchange rate. Some factors were significant, but had to be removed from the model due to multicollinearity.
Sammanfattning

Acknowledgements

We would like to thank our supervisors at the Royal Institute of Technology (KTH), Prof. Anja Janssen at the Department of Mathematics and Prof. Hans Lööf at the Department of Industrial Economics and Management for the assistance before and during the study.
## Contents

1 Introduction ............................................. 6  
   1.1 Background ........................................... 6  
   1.2 Research Question .................................... 6  
   1.3 Goal and Purpose .................................... 7  
   1.4 Previous Research ................................... 7  
   1.5 Scope and Limitations ................................. 8  

2 Theory .................................................. 9  
   2.1 Economic and Financial Theory ......................... 9  
      2.1.1 Interest Rate ..................................... 9  
      2.1.2 Inflation Rate .................................... 9  
      2.1.3 Gross Domestic Product ........................... 10  
      2.1.4 Government Debt .................................. 11  
      2.1.5 Unemployment ..................................... 11  
      2.1.6 Current Account .................................. 11  
      2.1.7 Indices indicating Political and Socioeconomic Performance .... 12  
   2.2 Mathematical Theory ................................ 12  
      2.2.1 Time Series Analysis ............................... 12  
      2.2.2 Multiple Linear Regression ....................... 13  
      2.2.3 Underlying Assumptions of Multiple Linear Regression .......... 14  
      2.2.4 Ordinary Least Squares ............................ 14  
      2.2.5 Multicollinearity Measures ......................... 14  
      2.2.6 Regression Model Evaluation Methods ............... 16  
      2.2.7 Variable Selection Techniques ...................... 20  
      2.2.8 Measuring and Handling Autocorrelation .............. 20  
      2.2.9 Confidence Intervals of the Coefficients .......... 21  

3 Methodology ............................................. 22  
   3.1 Data Collection ....................................... 22  
      3.1.1 Response Variables ................................ 22  
      3.1.2 Regressor Variables ................................ 22  
   3.2 Data Preparation ..................................... 24  
      3.2.1 Scope of Data Selection ............................ 24  
      3.2.2 Data Preparation ................................... 24  
      3.2.3 Data packaging .................................... 25  
   3.3 Regression Model Analysis ............................ 25  
      3.3.1 Structure of Regression Model ...................... 25
3.3.2 Residual Analysis .................................................. 25
3.3.3 Transformation of the Variables ................................. 25
3.3.4 Multicollinearity Measures ....................................... 26
3.3.5 Outlier Diagnostics ................................................. 26
3.3.6 Variable Selection ................................................. 26
3.3.7 Model Comparison .................................................. 26

4 Results ........................................................................ 27
  4.1 Multicollinearity ....................................................... 27
     4.1.1 Initial Model ...................................................... 27
  4.2 Residual Analysis of Initial Model ............................... 27
  4.3 Transformation of the Variables ................................. 28
  4.4 Variable Selection .................................................... 28
  4.5 Analysis of Final Model ............................................ 29

5 Discussion .................................................................... 32
  5.1 Discussion of Residual Analysis .................................. 32
  5.2 Discussion of Multicollinearity ................................... 32
  5.3 Analysis of Final Model ............................................. 32
     5.3.1 Stock Index ........................................................ 32
     5.3.2 Botswana Real Interest Rate minus US Real Interest Rate 33
     5.3.3 Corruption Index ................................................. 33
     5.3.4 Business Confidence Index ..................................... 34
     5.3.5 Diamond Price Index ........................................... 34
     5.3.6 Unemployment Rate ............................................. 34
  5.4 Analysis of Excluded Variables ................................... 34
     5.4.1 Current Account .................................................. 34
     5.4.2 Ease of Doing Business Index ............................... 35
     5.4.3 Botswana Government Debt .................................. 35
     5.4.4 GDP per Capita .................................................. 35
     5.4.5 Botswana CPI/ US CPI ......................................... 35
  5.5 Conclusions ............................................................ 36
  5.6 Evaluation of Chosen Methods .................................... 37
  5.7 Further Studies ....................................................... 37

6 Appendix ..................................................................... 40
  6.1 Table 3 - Results from All Possible Regressors ............. 40
1 Introduction

1.1 Background

Since its independence, Botswana has exhibited extraordinary growth compared to the rest of Sub Saharan Africa. The growth can be attributed to several factors, such as a well-functioning and expanding export economy and a stable domestic environment for businesses, facilitated by prudent economic governance. Botswana’s economy is heavily dependant on the country’s mining output; 88% of all exports are diamonds. Pursuing continued international competitiveness while maintaining a stable business environment, has presented to be challenging for a small economy in a volatile region. As a consequence, Botswana has underwent several changes in its exchange rate regime. [2]

In 1976, Botswana introduced its own currency, Pula, part of Botswana’s strive of becoming a more independent and self-managed economy. Initially, the Pula was completely pegged to the US Dollar. This however, changed in 1980, following the breakdown of the Bretton Woods system, and the Pula became pegged to the South African rand and IMF’s currency basked SDR. The choice of the pegged currency’s counterparts was due to Botswana trade patterns; large exchange inflows to Botswana’s economy would cause it to appreciate, if pegged to a single currency such as the dollar. Additionally, an exchange rate regime based on free floating mechanisms was deemed unsustainable by a small and export driven economy, according to the World Bank. [2]

Botswana’s current exchange rate policy is based on a crawling peg to the South African Rand and the IMF’s currency basket SDR. With other words, Botswana’s currency is free to move within certain levels depending on the country’s economic targets. Currently, the currency is weighted with 55% towards the SDR and 45% towards the Rand. The smaller weight for the South African Rand is said to be due to its notable volatility, according to Bank of Botswana. [7]

1.2 Research Question

In this research study we aim to analyse the Botswanan currency Pula and determine which factors that affect it and their respective significance. This is done by regressing these factors against the currency and testing for adequacy of the model.
The research question that is attempted to be answered is: determining factors that influence the Botswanan Pula and their respective significance.

1.3 Goal and Purpose

The goal of the study is to understand which factors affect the Botswanan Pula and to what extent. This will be tested for and formulated in a regression model, which is to be used for explaining the underlying dynamics and predicting future movements.

Having a general understanding of the underlying factors affecting the development of a country’s exchange rate and the cause-effect relationship, can be considered important for investing or initiating trade in that country, since stability and to a certain extent predictability, can be argued to be essential. The analysis would therefore be of value for any potential or current trade partners and investors in Botswana.

In addition, if Botswana were to decide to go back to a totally fixed currency, such an analysis would be useful in order to mitigate changes in the currency. The purpose of the project is to achieve a general understanding and transparency in the Botswanan Pula.

1.4 Previous Research

There have been a lot of research that investigates the relationship between external factors and exchange rates. Graciela Kaminsky, Saul Lizondo and Carmen M. Reinhart proposed an analysis of factors as indicators of currency crisis, which was published by IMF. In their research they found that international reserves, domestic credit, inflation, export growth and real GDP growth among others were good indicators [6]. Ravindran Ramasamy and Soroush Karimi Abar used a regression analysis in order to investigate the effects of macro-economic factors on the AUD/USD, Euro/USD and AUD/Euro currency pairs. They found that interest rate, inflation, balance of trade, corruption, GDP, Tax and Borrowing influenced the currency pairs [11]. However, previous mentioned researchers have looked at several currencies rather than a specific one, which to some extent generalises the effect of external factors on the currency. The underlying dynamic might be different for specific countries, especially for those in emerging regions. Therefore it might be more
relevant to look at previous research on currencies of countries in similar regions, such as other parts of Africa. Thaddeus and Nnneka tried to evaluate the effect of inflation and interest rates on the Nigerian currency and published their results in the Journal of Economics and Development Studies. As opposed to a regression analysis, Thaddeus and Nnneka employed a autoregressive distributed lag analysis. They found that the inflation had both short term and long term effects on the currency but that the interest rate were insignificant in the model [5].

1.5 Scope and Limitations

To limit the size of this study, some limitations have been posed on the scope. Included variables were: Botswana Stock Index, Real Interest Rate Spread, Current Account, Corruption Index, Business Confidence Index, Ease of Doing Business Index, EMEX Diamond Price Index, Botswana GDP per Capita, Botswana Unemployment Rate and Botswana CPI/US CPI. These were chosen based on economic theory and reasoning as well as mostly limited to factors affecting Botswana and not the US. Furthermore, the study is limited to regression analysis and does not evaluate other methods for analysing the effect of factors on a currency. Finally, the time span of which the data was collected from was limited to 2009-01-31 and 2016-01-31, mostly due to data availability.
2 Theory

2.1 Economic and Financial Theory

2.1.1 Interest Rate

A well documented and widely used theory to explain the dynamics of exchange rates, is the *Covered Interest Rate Parity*. The parity originates from international finance theory and is based on the principle of no arbitrage opportunities; it assumes that a holder of a currency can benefit from the risk free interest rate of the country in which the currency is being held. Thus, in the absence of no arbitrage opportunities, the development of the exchange rate between two currencies must follow the spread of the two interest rates. The relationship is illustrated in the following formula where $r_d$ and $r_f$ are the domestic and foreign interest rate compounded in $t$ respectively, $X_t$ is the exchange of domestic/foreign at time $t$.

\[
\frac{X_t}{X_0} = \frac{1 + r_d}{1 + r_f}
\]

This can also be supported by the market principle of supply and demand with the assumption that investors seeks to profit from a higher interest rate; an interest rate spread would therefore imply a higher demand for the currency with the higher interest rate and therefore appreciate in value. [8]

2.1.2 Inflation Rate

According to the theory of *Purchasing Power Parity (PPP)* the exchange rates of two countries should adjust so that the price levels are equal. PPP compares the price of a basket of goods and relative changes in prices. Thus, if the inflation rates differ between two countries, one would expect the currency of the country with the higher inflation rate to depreciate in value relative to the other currency. The most common statistic to measure inflation is the *Consumer Price Index (CPI)*. The inflation rate is the annual percent change in the index.
The relationship between changes in the exchange rate can be illustrated in the following formula where \( p_d \) and \( p_f \) is the domestic and foreign price index respectively and \( X \) is *domestic/foreign* exchange rate.

\[
\frac{dX(t)/dt}{X(t)} = \frac{dp_f(t)/dt}{p_f(t)} - \frac{dp_d(t)/dt}{p_d(t)}
\]

[13]

### 2.1.3 Gross Domestic Product

The *Gross Domestic Product (GDP)* is the total value of all final goods and services produced in a country in a specific time period, usually a year.

The GDP is the sum of consumer spending \((C)\), sales of investment goods and services \((I)\), government spending \((G)\), exports \((X)\) subtracted by the spending on imports \((IM)\). See formula below:

\[
GDP = C + I + G + X - IM
\]

The most common use of the GDP, is that it gives a measure of the size of an economy. This is very useful when comparing economic performance of two different years or two different nations.

One issue when comparing the GDP of two nations is that the more population a nation has, the more it will produce. Therefore it is reasonable to analyse GDP per capita, the GDP divided by the total population of the country. [9]

### The J-Curve

The theory of the J-curve poses an explanation to the relation between the value of the currency and trade balance. It suggests that a devaluation of the currency causes trade deficits in the short term but a trade surplus in the long term, thus shaping a J-shaped curve.[12]
2.1.4 Government Debt

Maintaining a healthy and high amount of government debt will allow a nation to increase government spending and reduce taxes and thus increase production. Although, if government debt gets relatively high investors might think that the nation will default on its debt and flee the country. A way to solve the issue of high debt is to print more money, which will cause inflation. Thus, a healthy level of debt will increase exports which in turn increases the value of the currency. If the debt gets too high investors will flee the currency because of the high risk of inflation, and thus the currency will depreciate in value. [9]

2.1.5 Unemployment

There exists a causal relation between the unemployment and inflation often measured by the Phillips Curve. It suggests that when unemployment is low the inflation rate should be high. This is because when unemployment is low, the people of a nation will have an increased purchasing power which leads to higher prices which leads to higher inflation. [9] Thus the correlation between the currency and unemployment should be positive. However, there are some argument to a reversed relationship; if unemployment rate is increasing, the economy might need stimulus in form of lower interest rates, which would cause the currency to depreciate [9].

2.1.6 Current Account

As stated earlier, a currency is affected by supply and demand, which according to theory implies that the exchange rate is set at the point of which demand and supply meets. To take this into account, one can analyse the current account, consisting of balance of payment on goods and services plus net factor income and net international services. In theory, the current account should be negatively correlated with the currency. For example, if imports were to increase of a country and further the current account, the supply of the related currency would also increase, causing it to depreciate. The relationship is usually described in the reversed order in common literature. [9]
2.1.7 Indices indicating Political and Socioeconomic Performance

There are many factors which are important to include when studying a developing country, such as Corruption Index, Business Confidence, Ease of Doing Business. These factors may be overlooked when analysing a developed country but are of the utmost importance to developing countries. According to [10] the consumer confidence index (CCI) has both a long-term and short-term relation to the exchange rate. The Corruption Index has been shown to have a positive correlation with GDP growth [14], as well as the Ease of Doing business index [1].

2.2 Mathematical Theory

2.2.1 Time Series Analysis

Time Series Model
A time series model for the observations \( x_t \) is a specification of the joint distribution of a sequence of random variables \( X_t \) of which \( x_t \) is a realisation.

Durbin-Watson Test
The Durbin-Watson test is based on the assumption that the errors of the model are a first order autoregressive process, i.e. \( \epsilon_t = \phi \epsilon_{t-1} + a_t \) where \( \epsilon_t \) is the error term, \( a_t \sim N(0, \sigma_a^2) \) and \( \phi \) is the lag one autocorrelation coefficient. The hypotheses of the test are

\[ H_0 : \phi = 0 \]
\[ H_1 : \phi > 0 \]

The Durbin-Watson test statistic is

\[ d = 2(1 - r_1) \]
\[ r_1 = \frac{\sum_{t=1}^{T-1} \epsilon_t \epsilon_{t+1}}{\sum_{t=1}^{T} \epsilon_t^2} \]

For uncorrelated errors the test statistic will be close to 2, since the one lag autocorrelation \( r_1 \) will be close to zero. Although, it is necessary to determine how far away from 2 the statistic must fall for us to conclude there is no autocorrelation. The following rules apply:

\[ d < d_L \] reject \( H_0 \)
\[ d > d_U \] do not reject \( H_0 \)
\[ d_L < d < d_U \] test inconclusive
The values on $d_U$ and $d_L$ depends on the number of predictors and the sample size[4].

**Autocorrelation** Autocovariance is defined as the covariance between two variables in the same times series. The *autocovariance function (AFC)* is defined as follows:

$$\gamma(h) = \text{Cov}(X_{t+h}, X_t)$$

where $h$ is the amount of lag, or the distance in time between the two variables. To calculate the autocovariance of observed data, the *sample autocovariance function* is used.

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n$$

The *sample autocorrelation function* is calculated as follows [3]:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

### 2.2.2 Multiple Linear Regression

Regression analysis is a statistical technique often used for investigating and modeling the relationship between variables. The goal is to find the best possible linear equation that describes the relationship between the response variable $y$ and the regressors $x_1, x_2, \ldots, x_k$[4]. The equation can be stated as follows:

$$y = X\beta + \epsilon$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

In the model we have $k$ dependent variables and $n$ observations. The error term $\epsilon$ is assumed to be independent and identically distributed variables (IID) with a $N(0, \sigma^2)$ distribution. $\beta$ is a vector of coefficients which is estimated using the least squares algorithm. [4]
2.2.3 Underlying Assumptions of Multiple Linear Regression

In order to perform linear regression some underlying assumptions must be met.

- The relationship between $y$, $X$ and $\epsilon$ is linear.
- There is no systematic error, $E(\epsilon) = 0$
- The error terms have constant variance, $Var(\epsilon_i) = \sigma^2$
- The error terms are normally distributed, $\epsilon_i \sim N(0, \sigma^2)$
- The errors are uncorrelated, $Cov(\epsilon_i, \epsilon_j) = 0, j \neq i$

If any of these assumptions is violated the regression model and results might be misleading. [4]

2.2.4 Ordinary Least Squares

The goal with the Ordinary Least Squares algorithm is to minimize the square of the error terms $\epsilon = y - X\beta$, i.e. minimizing $S(\beta) = (y - X\beta)^T(y - X\beta)$. The optimal $\beta$ is $\hat{\beta} = (X^TX)^{-1}X^Ty$.

$\hat{\beta}$ has the following properties:

- It is an unbiased estimator of $\beta$, $E[\hat{\beta}] = \beta$
- $Var(\hat{\beta}_j) = \sigma^2(X^TX)_{jj}$
- $Cov(\hat{\beta}_{ij}) = \sigma^2(X^TX)_{ij}, i \neq j$

[4]

2.2.5 Multicollinearity Measures

In most cases of multiple linear regression, the regressors are non completely orthogonal. Meaning, there exists some linear dependence among them. If the regressors are near perfectly linearly related, multicollinearity is said to exist. Common causes of multicollinearity can be the following:

- Collection of data
- Model constraints
Overdefined model

• Specification of model

Multicollinearity may have undesired effects on the model; among them are large variances and covariances of the regressor coefficients. It can be shown that the diagonal elements of the \( C = (X^T X)^{-1} \) matrix are \( C_{jj} = (1 - R_j^2)^{-1} \). This implies that high correlation and \( R_j^2 \) close to 1, causes large variances and covariances, since

\[
Var[\hat{\beta}] = \sigma (X^T X)^{-1},
\]

where \( R_j^2 \) is the \( R^2 \) of the model where the regressor \( j \) is used as dependent variable and the rest of the regressors is used as independent variables. In addition, multicollinearity can also cause coefficient estimates that are too large in absolute value.

There are several ways of detecting multicollinearity, two of those are:

**Variance Inflation Factor**

As stated previously, \( C_{jj} \) can be expressed as \((1 - R_j^2)^{-1}\) which is large if linear dependence exists. One way to investigate if multicollinearity is present, is therefore simply to examine the diagonal elements of \( C \); the diagonal elements are called *variance inflation factors* (VIF). VIFs exceeding 10, indicates multicollinearity. The VIF formula is summarised below:

\[
VIF_j = C_{jj} = \frac{1}{1 - R_j^2}
\]

**Condition Number**

The eigenvalues of \( X^T X \) can be used as a measurement for multicollinearity. Small eigenvalues indicate that linear dependence exists between the columns of \( X \). A method of quantifying that phenomenon is to examine the *condition number*, defined below where \( \lambda \) is the eigenvalues of \( X^T X \). A condition number higher than 1000 indicates severe multicollinearity; between 100 and 1000 implies some multicollinearity; if values are lower than 100, there is generally no problem with multicollinearity.

\[
\kappa = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}
\]
2.2.6 Regression Model Evaluation Methods

In order to analyse the obtained model three simple statistic measures can be used:

**Regression Sum of Squares**
The regression sum of squares equals the explained variability in the model.

\[
SS_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2
\]

where \(\hat{y}_i\) is the i:th estimate of the dependant variable and \(\bar{y}\) is the mean of all observations.

**Residual Sum of Squares**
The residual sum of squares is a common statistic for measuring the total difference of the model and the observed values.

\[
SS_{Res} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

**Total Sum of Squares**
The total sum of squares measures the volatility in the observed values.

\[
SS_T = SS_R + SS_{Res} = \sum_{i=1}^{n} (y_i - \bar{y})^2
\]

[4]

**Coefficient of Determination**
The coefficient of determination, commonly called \(R^2\) is a measure of how good the fitted model represents the observed values. It is calculated as follows:

\[
R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{Res}}{SS_T}
\]

If the \(R^2\) is close to 1 it suggests that the model represents the data well and if it is closer to 0 the model failed to represent the data[4].

**Adjusted \(R^2\)**
In many cases the adjusted \(R^2\) statistic is preferred to the Coefficient of Determination, because it punishes the addition of new regressors. Adjusted \(R^2\), commonly
denoted $R^2_{Adj}$ only increases when the mean squared error decreases which is not always the case when adding new variables to the model. It is calculated as follows:

$$R^2_{Adj} = 1 - \frac{SS_{Res}/(n-p)}{SS_T/(n-1)}$$

[4]

PRESS Residuals

The PRESS residual is calculated by deleting the $i$th observation and fitting the model on the remaining $n-1$ data points and calculate the difference between the model fitted on all $n$ observations.

$$e_{(i)} = \hat{y}_i - \hat{y}(i) = \frac{e_i}{1 - h_{ii}}$$

where $h_{ii} = x_i^T (X^T X)^{-1} x_i$.

If there is a large difference between $e_i$ and $e_{(i)}$, observation $i$ is considered a point which fits the data well but will predict on new data poorly[4].

R-student

To diagnose whether a point is an influential point, the externally studentized residual, or the R-student is a recommended statistic. It is calculated as follows:

$$t_i = \frac{e_i}{\sqrt{S^2_{(i)} (1 - h_{ii})}}, \quad i = 1, 2, ..., n$$

$$S_{(i)} = \frac{(n-p) MS_{Res} - e^2_i / (1 - h_{ii})}{n-p-1}$$

$S_{(i)}$ is an estimate of $\sigma^2$ with the $i$:th observation removed.

PRESS Statistic

The PRESS statistic is the sum of squared PRESS residuals, described above. It is used as a measure of how well the model with perform in predicting new observations. The formula is stated below:

$$PRESS = \sum_{i=1}^{n} \left( \frac{e_i}{1 - h_{ii}} \right)^2$$
A small value of the PRESS statistic is desired[4].

**Akaike Information Criterion**
The Akaike Information Criterion (AIC) is a penalized log-likelihood measure.

\[ AIC = -2 \ln(L) + 2p \]

*L* is the likelihood function for the model and *p* is the number of parameters in the model. Thus it measures the log-likelihood of a model which is penalized by the amount of parameters used. The *AIC* of a ordinary least squares model is calculated as follows:

\[ AIC = n \ln\left(\frac{SS_{Res}}{n}\right) + 2p \]

When choosing the amount of parameters to use one usually compares the *AIC* of the different models constructed and chooses the model with the lowest AIC[4].

**Bayesian Information Criterion**
There exists a few Bayesian extensions of the *AIC*. The one used in this report is called the Schwartz criterion. The formula is:

\[ BIC = -2 \ln(L) + p \ln(n) \]

The main difference between the *BIC* and the *AIC* is that the *BIC* penalizes adding more parameters as the sample size *n* increases. Similar to the *AIC*, the *BIC* for ordinary least squares models is calculated as:

\[ BIC = n \ln\left(\frac{SS_{Res}}{n}\right) + p \ln(n) \]

[4]

**Mallow’s *C*<sub>p</sub> Statistic** Mallow’s *C*<sub>p</sub> statistic is used to assess the fit of a regression model. It takes into account both the variance and the bias of the model and is calculated as follows:

\[ C_p = \frac{SS_{Res}(p)}{\hat{\sigma}^2} - n + 2p \]

As it punishes the number of predictors in the model it helps to detect overfitting. A small value of *C*<sub>p</sub> is desired[4].
t-test
The t-test is commonly used to test the significance of a coefficient $\beta_j$ corresponding to a specific predictor $x_j$ in a regression model. The hypotheses of the test are

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

If the null hypothesis is rejected we can with a predetermined confidence level deem the predictor $x_j$ significant and include it in the model.

The t-test statistic for predictor $x_j$ is defined as

$$t_j = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 (X^T X)^{-1}}}$$

where $\hat{\sigma}^2$ is set equal to $MS_{Res}$.

We reject the null hypothesis if

$$|t_j| > t_{\alpha/2,n-k-1}$$

where $\alpha$ is the chosen significance level of the test[4].

F-test
In order to test whether there is any linear relationship between the response variables $y$ and any of the predictors $x_1, x_2, ..., x_k$ one can construct a test for significance of regression with the hypotheses

$$H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$$

$$H_1 : \beta_j \neq 0 \text{ for any } j$$

An appropriate test statistic is the F statistic defined below

$$F_0 = \frac{SS_R/k}{SS_{Res}/(n-k-1)} = \frac{MS_R}{MS_{Res}}$$

$F_0$ follows a $F_{k,n-k-1}$ distribution. [4]

We chose to reject the null hypothesis $H_0$ if $F_0 > F_{\alpha,k,n-k-1}$ and conclude that there is at least one predictor with significance[4].

19
2.2.7 Variable Selection Techniques

All Possible Regressors
This method consists of trying all possible combinations of regressor variables and evaluate the different models according to some statistic, such as the BIC or $R^2_{adj}$. It starts with only the intercept, $\beta_0$, and successively adds the other regressors until every combination has been tried[4].

One has to consider the amount of calculations needed for this method. If K is the number of regressors used in the model this method fits and evaluates $2^K$ models which might not be feasible if the number of regressors is high[4].

2.2.8 Measuring and Handling Autocorrelation

As mentioned previously, one of the underlying assumptions for multiple linear regression is independent errors. This is usually not the case when regressors and response variables are time series since the error terms tend to show autocorrelated structures. This will have several effects on the OLS regression model:

- The OLS coefficients are not minimum-variance estimates
- If there is positive autocorrelation between the error terms, the residual mean square may underestimate the actual error variance. As a consequence, confidence and predictor intervals may be smaller that they actually are

In order to measure the amount of autocorrelation in the errors the Durbin-Watson statistic is often used.

If the errors follow a autoregressive process (AR(p)-process) the Cochrane-Orcutt method can be used to handle the autocorrelation.

The Cochrane-Orcutt method transforms the data in the following way (simple regression case):

$$y'_t = y_t - \phi y_{t-1}$$

$$= \beta_0 + \beta_1 x_t + \epsilon_t - \phi(\beta_0 + \beta_1 x_{t-1} + \epsilon_{t-1})$$

$$= \beta'_0 + \beta'_1 x'_t + \epsilon'_t$$
where $\beta'_0 = \beta_0(1 - \phi)$, $x'_t = x_t - \phi x_{t-1}$ and $\epsilon'_t = \phi \epsilon_{t-1} + a_t, a_t \sim N(0, \sigma^2_a)$. With the following regression, $\phi$ can be estimated:

$$\hat{\phi} = \frac{\sum_{t=2}^{T} \epsilon_t \epsilon_{t-1}}{\sum_{t=1}^{T} \epsilon_t^2}$$

Now $\phi$ is estimated by $\hat{\phi}$ and inserted into the equations above. Then OLS is applied to the transformed data which will result in different residuals on which the Durbin-Watson test is applied to see if there still is any autocorrelation in the model. If there is, the same procedure is applied again. [4]

### 2.2.9 Confidence Intervals of the Coefficients

When the assumptions of multiple regression are met the vector of estimates, $\hat{\beta}$, follows a normal distribution with mean $\beta$ and covariance matrix $\sigma^2(X'X)^{-1}$. Thus each coefficient $\hat{\beta}_j$ follows a normal distribution with mean $\beta_j$ and variance $\sigma^2(X'X)^{-1}_{jj} = \sigma^2 C_{jj}$. Then,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 C_{jj}}} \sim t_{n-p}$$

Thus the confidence interval at significance level $\alpha$ for the j:th coefficient can be calculated as:

$$\hat{\beta}_j - t_{\alpha/2,n-p}\sqrt{\sigma^2 C_{jj}} \leq \beta_j \leq \hat{\beta}_j + t_{\alpha/2,n-p}\sqrt{\sigma^2 C_{jj}}$$

[4]
3 Methodology

3.1 Data Collection

3.1.1 Response Variables

The response variable is the BWP/USD exchange rate daily closing price. It was fetched from the Reuter’s Eikon terminal located in the KTH library.

3.1.2 Regressor Variables

The regressor variables consist of different factors that, according to economic theory and general consensus, could affect the exchange rate.

- Botswana Stock Index
  This data is comprised by 3611 observations of the daily closing prices of the Botswanan stock index.

- US 10 Year Government Bond
  This data is comprised by 3462 observations of the daily closing yield to maturity of the US 10 year government bond. This was used as a proxy for the US risk-free nominal interest rate.

- Botswana 10 Year Government Bond
  It consists of 88 observations of the monthly closing yield to maturity of the Botswanan 10 year government bond. This was used as a proxy for the Botswanan risk-free nominal interest rate.

- US Consumer Price Index
  It consists of 241 monthly observations of the US Consumer Price Index. This was used for the calculation of the US inflation rate.

- Botswana Consumer Price Index
  It consists of 266 monthly observations of the Botswanan Consumer Price Index. This was used for calculation of the Botswanan inflation rate.

- Botswana’s Population
  It consists of 59 yearly observations of the population of Botswana.
• Botswana GDP  
  It consists of 43 yearly observations of the Botswanan GDP.

• Botswana Current Account  
  It consists of 55 quarterly observations of the Botswanan Current Account.

• Botswana Unemployment  
  It consists of 27 yearly observations of the Botswanan unemployment rate.

• Botswana Business Confidence Index  
  It consists of 14 yearly observations of the Botswanan Business Confidence Index.

• Botswana Ease of Doing Business Index  
  It consists of 14 yearly observations of the Botswanan Ease of Doing Business.

• Botswana Corruption Index  
  It consists of 14 yearly observations of the Botswanan Corruption Index, where 0 indicates strong corruption and 100 very little corruption.

• Diamond Price Index  
  It consists of 2593 daily observations of the EMEX Diamond Price index.

The regressors became as follows:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>Botswana Stock Index ($BSI$)</td>
</tr>
<tr>
<td>$x_2$</td>
<td>Real Interest Rate Spread, Botswana minus USA ($RIRS$)</td>
</tr>
<tr>
<td>$x_3$</td>
<td>Current Account ($CA$)</td>
</tr>
<tr>
<td>$x_4$</td>
<td>Corruption Index ($CI$)</td>
</tr>
<tr>
<td>$x_5$</td>
<td>Business Confidence Index ($BCI$)</td>
</tr>
<tr>
<td>$x_6$</td>
<td>Ease of Doing Business Index ($EDBI$)</td>
</tr>
<tr>
<td>$x_7$</td>
<td>EMEX Diamond Price Index ($EDPI$)</td>
</tr>
<tr>
<td>$x_8$</td>
<td>Botswana Government Debt ($BGD$)</td>
</tr>
<tr>
<td>$x_9$</td>
<td>Botswana GDP per Capita ($BGDPC$)</td>
</tr>
<tr>
<td>$x_{10}$</td>
<td>Botswana Unemployment Rate ($BUR$)</td>
</tr>
<tr>
<td>$x_{11}$</td>
<td>Botswana CPI/US CPI ($RCPI$)</td>
</tr>
</tbody>
</table>
3.2 Data Preparation

3.2.1 Scope of Data Selection

Due to certain limitations of the information collection infrastructure in Botswana, obtaining the correct data proved to be challenging. Only variables with complete data from trusted sources were considered.

3.2.2 Data Preparation

Macroeconomic data is published at different time intervals, i.e. the exchange rate is published daily and the unemployment rate is published yearly.

The assumption that the average estimation of each variable should be a linear interpolation between each published value was made. Because of this, the data with longer time intervals between publications were interpolated.

The data that were published monthly were interpolated as well, as the dates of publications differed.

To calculate the inflation rates of the specific countries, monthly percentage changes was computed of the respective Consumer Price Index. This was used together with the nominal interest rates in order to calculate the real interest rate, according to the formula below.

$$r_{real} = \frac{1 + r_{nominal}}{1 + r_i}$$

The difference of the real interest rates of each respective country is later used as a predictor in the models.

The absolute purchasing power parity (APPP) of a currency pair is calculated as in the formula below:

$$APPP = \frac{CPI_{Botswana}}{CPI_{US}}$$

The APPP is used as proxy for the purchasing power parity predictor variable in the models.

As mentioned in section 2.1.3 a good measure of the productivity of a nation is
the GDP per capita. This is calculated by dividing the GDP and the total population.

3.2.3 Data packaging

When doing regression on a time series one often wants to reduce the amount of autocorrelation of the error terms. Thus, the data was split into monthly time intervals.

3.3 Regression Model Analysis

Using the data presented in section 3.1 and prepared in section 3.2 together with the programming language R and the IDE RStudio, the regression model was constructed and analysed.

3.3.1 Structure of Regression Model

The initial model can be written as:

\[ y = \beta_0 + \beta_1 x_1 + \ldots + \beta_{12} x_{12} + \epsilon \]

The predictors \( x_1, x_2, \ldots, x_{12} \) are presented in section 4.1.2.

3.3.2 Residual Analysis

To determine if autocorrelation is present, the Durbin-Watson test was used. Furthermore, the R-student was calculated and the R-student residuals were plotted against the fitted values \( \hat{y} \). A QQ-plot was constructed as well. These two plots were inspected for any violations of the normality assumptions of the regression model, as well as homoscedacity.

3.3.3 Transformation of the Variables

In order to handle the autocorrelation of the error terms, the Cochrane-Orcutt method was used to transform the data and construct a new model.
3.3.4 Multicollinearity Measures

To measure the amount of multicollinearity of each model the Variance Inflation Factor (VIF) of each variable was calculated, as well as the condition number of each model.

3.3.5 Outlier Diagnostics

To check whether there is outliers the Studentized Residuals, Hat-values, Covratio and Cook’s distance was calculated and inspected. No points with severe values were encountered thus none were removed.

3.3.6 Variable Selection

All Possible Regressors algorithm was considered to choose suitable models. The best model for each of the test statistics BIC, $R^2_{adj}$ and $C_p$ was chosen, resulting in three different models. A fourth model was fitted with regressors the Botswanan unemployment rate and the diamond price index, called the Diamond Model.

The Durbin-Watson test was applied to each of the models in order to make sure that autocorrelation of the residuals, was not present.

The condition number was calculated for each of the models, in order to check for severe problems with multicollinearity.

3.3.7 Model Comparison

To choose between the three constructed models using the All Possible Regressions algorithm, a 4:th statistic, the PRESS statistic, was applied. The model with the lowest PRESS statistic was chosen as the final model. Later, the AIC statistic is used to compare the final model and the Diamond Model.
4 Results

4.1 Multicollinearity

4.1.1 Initial Model

The Variance Inflation Factors of the initial model are presented below. The predictors Government Debt \((x_8)\), Botswana CPI/ US CPI \((x_{11})\) and the Ease of Doing Business \((x_6)\) index were removed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1)</td>
<td>37.1256</td>
</tr>
<tr>
<td>(x_2)</td>
<td>18.3886</td>
</tr>
<tr>
<td>(x_3)</td>
<td>5.5656</td>
</tr>
<tr>
<td>(x_4)</td>
<td>15.8794</td>
</tr>
<tr>
<td>(x_5)</td>
<td>3.4715</td>
</tr>
<tr>
<td>(x_6)</td>
<td>20.5560</td>
</tr>
<tr>
<td>(x_7)</td>
<td>5.3770</td>
</tr>
<tr>
<td>(x_8)</td>
<td>398.2403</td>
</tr>
<tr>
<td>(x_9)</td>
<td>107.4753</td>
</tr>
<tr>
<td>(x_{10})</td>
<td>35.4710</td>
</tr>
<tr>
<td>(x_{11})</td>
<td>344.1692</td>
</tr>
</tbody>
</table>

The condition number of the initial model is just above 500000 before removing the variables stated above. The condition number was reduced to about 91 after removing the variables.

4.2 Residual Analysis of Initial Model

The Durbin-Watson test statistic was 1.05 and the autocorrelation function showcased strong autocorrelation between the error terms. In order to solve for this, the Cochrane-Orcutt method was used. The plot of the autocorrelation function is shown below:
4.3 Transformation of the Variables

The Durban-Watson test statistic after the transformation is 1.9220. The new autocorrelation function is shown below:

4.4 Variable Selection

The resulting models, $y_{BIC}$, $y_{R^{2}_{adj}}$ and $y_{C_p}$ of the All Regressors Algorithm with the best $R^{2}_{adj}$, $C_p$ and BIC values are presented in table 3 in the appendix. All three models show a significance of regression at 99 % with $R^{2}_{adj}$ over 0.9.
The fourth model, $y_{Dia}$, had a $R^2_{adj}$ of 0.1307 and a significance of regression at 99%. The PRESS statistic of the models are shown in the table below:

<table>
<thead>
<tr>
<th></th>
<th>$y_{BIC}$</th>
<th>$y_{R^2_{adj}}$</th>
<th>$y_{Cp}$</th>
<th>$y_{Dia}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.6670</td>
<td>1.6786</td>
<td>1.6670</td>
<td>2.2483</td>
</tr>
</tbody>
</table>

The AIC values of $y_{Cp}$ and $y_{Dia}$ are shown below:

<table>
<thead>
<tr>
<th></th>
<th>$y_{Cp}$</th>
<th>$y_{Dia}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-92.8760</td>
<td>-65.9756</td>
</tr>
</tbody>
</table>

Thus our final model is $y_{Cp}$ and the corresponding formula is shown below:

$$y = \beta_0 + \beta_1BSI + \beta_2RIRS + \beta_4CI + \beta_5BCI + \beta_7EDPI + \beta_{10}BUR$$

### 4.5 Analysis of Final Model

The correlation between each of the variables in the final model:

The plot of the residuals vs fitted values, the autocorrelation function and the QQ-normal plots are shown below:
The significance of each coefficient is presented in the table below:
<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Standard Error</th>
<th>Significance level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0.1108</td>
<td>0.0163</td>
<td>99.9%</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.4983</td>
<td>0.0610</td>
<td>99.99%</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.2041</td>
<td>0.0376</td>
<td>99%</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.3185</td>
<td>0.0376</td>
<td>99.9%</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-0.0905</td>
<td>0.0345</td>
<td>95%</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>0.1775</td>
<td>0.0340</td>
<td>99.9%</td>
</tr>
<tr>
<td>$\beta_{10}$</td>
<td>0.1863</td>
<td>0.0350</td>
<td>99.9%</td>
</tr>
</tbody>
</table>

The confidence intervals at the 5% significance-level for each of the coefficients are shown below:

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0.0784</td>
<td>0.1432</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.6199</td>
<td>-0.3767</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0603</td>
<td>0.3478</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.3935</td>
<td>-0.2436</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-0.1593</td>
<td>-0.0218</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>0.1098</td>
<td>0.2452</td>
</tr>
<tr>
<td>$\beta_{10}$</td>
<td>0.1166</td>
<td>0.2559</td>
</tr>
</tbody>
</table>

The Durban-Watson statistic of the final model is 1.9007 and the $R^2_{\text{adj}}$ is 0.903. A plot of the final fit is presented below to the left and the prediction with prediction intervals to the right. The blue line is the exchange rate, the red is the estimated exchange rate and the green ones represent the prediction interval.
5 Discussion

5.1 Discussion of Residual Analysis

In section 4.2 the Durbin-Watson test statistic of the initial model was 1.0515. As this is lower than \( d_L \) we reject the null hypothesis, which indicates autocorrelation between the error terms. This violates the assumption of multiple linear regression that the error terms should be uncorrelated and must therefore be handled. The method chosen was to transform the data with the Cochrane-Orcutt method. The final model, which is fitted on the transformed data has a Durbin-Watson test statistic of 1.9872 which is higher than \( d_U \). Thus we no longer have a problem with autocorrelation of the residuals.

The plot of the residuals vs fitted values of the final model in section 4.5 indicates that the model is relatively homoscedastic since the residuals are randomly spread around zero. The QQ-plot suggest that the residuals are normally distributed.

Thus none of the assumptions of multiple linear regression is violated and we can make inferences such as confidence intervals and various test.

5.2 Discussion of Multicollinearity

The condition number of our final model is 56.2876, which indicates small multicollinearity in the final model. Thus there is risk that the estimated coefficients of the regression model have large variances and covariances. By inspection of the correlation plot in section 4.5 some of the variables have strong correlation. This is expected when handling macroeconomic factors, as they usually do affect each other in several different ways.

5.3 Analysis of Final Model

5.3.1 Stock Index

The Botswana Stock Index had a coefficient of -0.4983 in the model with a significance exceeding 99%. Both a positive and a negative correlation between the currency and the stock market can be expected, depending on the dynamics of the
country’s economy; increasing stock prices indicates a healthy and stable economy, which would give rise to stronger confidence in that country; on the contrary, increasing stock prices might be partly due to decreasing interest rates, which would cause investors to move funds. In this case, the latter relationship could be argued for. This can also be supported by the fact that the stock index has a negative correlation with interest rate; a decreasing interest spread would according to theory, dilute the currency. However, more analysis would be required to establish that relations with certainty.

5.3.2 Botswana Real Interest Rate minus US Real Interest Rate

The Botswana Real Interest Rate minus the US Real Interest Rate had a positive correlation with a coefficient of 0.2041. As described in the previous subsection as well as in the macro economic theory section, a positive correlation between the interest spread and the currency is expected. A increase in Botswana’s relative interest rate would increase capital flows to the economy and the currency would therefore appreciate.

5.3.3 Corruption Index

The Corruption Index had a coefficient of -0.3185 with a statistical significance above 99%. This is somewhat inconsistent with theory. Since a high corruption index indicates very little corruption, one would assume that increases in the index would imply stronger confidence in the country’s economy and therefore cause the currency to appreciate. An explanation could be due to the relationship with interest rates supported by their negative correlation; lower interest rates had a positive correlation with the stock index, which itself indicates the overall performance of the economy. Rising stock prices could perhaps cause corruption to decrease and therefore the index to gain. With other words, lower interest rates causes a healthier economy in short term, which causes the index to increase.
5.3.4 Business Confidence Index

The Business Confidence Index had a coefficient of -0.0905. As with the stock index, both a positive and negative correlation may exist; increasing confidence may imply that corporations will hire and invest more, which would increase inflation and therefore decrease the currency; on the contrary, an increase in the index may indicate stronger confidence in the country and a rise in its currency. By solely examining the negative coefficient, one might argue that the first case is true. However, the Business Confidence Index showcased a strong negative correlation with the relative CPI, which would imply some other dynamic that needs explanation. One hypothesis, could be that decreasing confidence would require economic stimulus in the form of lower interest rates, which would cause the currency to depreciate. This can be supported by a positive correlation of 0.65 between the confidence index and the interest spread.

5.3.5 Diamond Price Index

As stated in the background, Botswana’s economy is heavily dependant on the diamond output; 88% of exports are attributed to diamonds. Therefore, a positive correlation could be expected, since increasing diamonds prices within certain limits, could cause greater inflows to the Botswanan currency. The Diamond Price Index had a coefficient of 0.1775 and a significance of 99.9% in the regression model.

5.3.6 Unemployment Rate

The unemployment rate had a coefficient of 0.1863. This is in line with the theory stated in section 2.1.5 and thus validates that if the unemployment rate is low, inflation will increase, which will decrease the value of the currency.

5.4 Analysis of Excluded Variables

5.4.1 Current Account

The Current Account was removed due to decrease of fit. In theory, the Current Account should have a significant effect on the currency; as described in the macro
economic theory section, Current Account effectively measures the demand and supply of a currency. A potential explanation to the decrease of fit, can be that the reporting of Current Account might lag the actual effect.

5.4.2 Ease of Doing Business Index

The Ease of Doing Business Index was removed due to a high VIF. The index is according to theory supposed to be important when analysing currencies, since it can indicate confidence in the economy. The information is likely already provided by other regressors such as the Corruption Index and Business Confidence Index, which causes the multicollinearity.

5.4.3 Botswana Government Debt

The government debt was removed due to multicollinearity. If a government takes on too much debt, confidence in the economy might decrease causing the currency to depreciate. Information about this effect however, might already be provided by other regressors such as the Confidence Index and to some extent the interest rate and the stock index.

5.4.4 GDP per Capita

The GDP per capita had a positive coefficient when it was included. This can be explained by the phenomena called the J-curve, described in the economic theory section. The theory suggests that a currency devaluation causes trade deficits in the short term, but in the long run results in a trade surplus. Thus, it has a short term positive correlation. There are many other ways in which the GDP per capita affects the exchange, but that is not in the scope of this analysis.

5.4.5 Botswana CPI/ US CPI

When the relative CPI of Botswana and the USD was included, it had a strong negative correlation with the exchange rate. As described in the economic theory section, that is well argued for; if domestic inflation increases relatively to foreign inflation, the purchasing power decreases of that currency and such, the value.
Additionally, the relative CPI showcased a high VIF, which was the reason why it was removed from the final model. Many of the macro-economic factors are known to affect each other and there is rarely isolated cause-effect relationships. From a mathematical standpoint, this would mean that high coefficient values also implies multicollinearity, since variables affecting the currency are likely to also affect the relative CPI. All this taken into account, high VIF together with high coefficient values, is expected.

5.5 Conclusions

In this study, the Botswanan Pula was analysed and factors determining the exchange rate between the Pula and the US Dollar are investigated using a multiple linear regression model. Due to the time series characteristic of the data, the Cochrane-Orcutt method was used, which solved the problem of autocorrelation. Potential factors are included according to macro economic theory and reasoning. The empirical research in this study, found that several factors did affect the currency: the Botswana Stock Index, the Real Interest Rate Spread between Botswana and the US, the Corruption Index, Business Confidence Index, Diamond Price Index and Unemployment Rate. Factors removed, either due to multicollinearity or decreasing of model fit, are: Current Account, Ease of Doing Business Index, Botswana Government Debt, GDP per Capita and Botswana CPI/US CPI.

Factors included in the model and their respective coefficient are by and large consistent with economic theory and reasoning; the Real Interest Rate Spread showcased a positive correlation with the currency, which is a fundamental relationship according to economic theory; the Diamond Price Index was expected to have positive correlation since Botswana’s economy is heavily dependant on diamonds; Unemployment Rate had a positive coefficient, mainly consistent with the relationship with inflation. The resulting coefficients of some variables are to some extent inconsistent with theory; the Stock Index could be expected to have positive correlation but showcased negative, which could be explained by the relationship between interest rates and stock prices; the Corruption Index had a negative coefficient that could perhaps be explained by interest rates effect on the economy and therefore the corruption; Business Confidence Index had a negative coefficient contradicting economic reasoning that would argue for that stronger confidence in the economy would increase exchange flows and therefore the demand of the currency.
5.6 Evaluation of Chosen Methods

The Cochrane-Orcutt method makes it possible to fulfill the underlying assumptions of linear regression when doing regression on a time series as it handles the issue of autocorrelation in the error terms. It does however, depend on the fact that the correlation structure of the error terms is an $AR(p)$-process, which might not be fulfilled. Other methods, such as fitting the parameters of a $ARMA(p,q)$-process using maximum likelihood could be used instead of the Cochrane-Orcutt [8].

Using linear interpolation of the variables within the time intervals which are too broad may be incorrect. For example, the GDP of a country does not follow a linear curve. Fitting a more advanced mathematical models to these variables could yield results closer to the real values. However, this was not within the scope of this study.

As the PRESS-statistic does not punish the amount of regressors used in a model, some other statistic, such as the AIC, might be more suitable for choosing a final model [4].

5.7 Further Studies

As described in the section above, the Cochrane-Orcutt method is not suitable for all problems of this sort. Using other mathematical methods such as Prais Winsten regression, which doesn’t omit the first observation as Cochrane-Orcutt does, or time series models is recommended.

The choice of macroeconomic factors in this analysis originates from general macroeconomic theory, which might not be suitable when analysing a fast growing country in a volatile region. Therefore, we suggest a more thorough analysis of which factors affect the exchange rate of a country such as Botswana.
References


6 Appendix

6.1 Table 3 - Results from All Possible Regressors

<table>
<thead>
<tr>
<th>$y_{BIC}$</th>
<th>$y_{R^2_{adj}}$</th>
<th>$y_{C_p}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$x_1$</td>
<td>$x_1$</td>
</tr>
<tr>
<td>$x_7$</td>
<td>$x_4$</td>
<td>$x_5$</td>
</tr>
<tr>
<td>$x_{10}$</td>
<td>$x_5$</td>
<td>$x_9$</td>
</tr>
<tr>
<td>$x_{11}$</td>
<td>$x_9$</td>
<td>$x_7$</td>
</tr>
<tr>
<td></td>
<td>$x_7$</td>
<td>$x_{10}$</td>
</tr>
<tr>
<td></td>
<td>$x_{10}$</td>
<td>$x_{11}$</td>
</tr>
<tr>
<td></td>
<td>$x_{11}$</td>
<td></td>
</tr>
</tbody>
</table>