PREDICTION OF INFLATION RATES IN KENYA USING BINOMIAL LOGISTIC REGRESSION

BY

CARETAS AOKO ARIKO

A research project submitted to the Department of Applied Statistics and Actuarial Science in the School of Mathematics, Statistics and Actuarial Science in partial fulfillment of the requirements for the award of the degree of Masters of Science in Applied Statistics with IT of Maseno University.

SCHOOL OF MATHEMATICS,STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSIRTY

Declaration

This research project is my original work and has not been submitted for a degree or examination in this or any other university.

CARETAS AOKO ARIKO

MSC/MAT/00144/2014

SIGNATURE:....................................DATE:...............................

This report has been submitted for examination with my approval as a university supervisor.

DR.EDGAR OTUMBA

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSITY

SIGNATURE:......................................DATE:...............................

Acknowledgement

I am grateful and sincerely feel obligated towards all many people without whom this work would not have come this far. First of all, my thanks go to the Almighty God for the knowledge and good health while carrying out this study. I would like to express my most sincere gratitude to my supervisor Dr. Edgar Otumba for his guidance and encouragement to accomplish this project. I am extremely thankful and pay my gratitude to Killy Onyango, Toquin Zadrack, Cheryl Atieno, Vincent Lukio, Dr. Gabriel Dida and my parents who supported me as I pursued my studies. I would also like to thank my colleagues Erick Ng'eno, Catherine Gichuki and the entire KCB Kisumu West team for their moral support. Thanks are also extended to all the lecturers of the Department of Applied Statistics and Actuarial Science who prepared me during course work. Finally, I thank Cecelia Badia, George Ong'any, James Wekunda and all of my friends for their conscientious encouragement. To all those who made this work a success, omission does not mean lack of gratitude. Thank you very much and may the Lord bless you.

Dedication

To my parents Calvin Ariko Lukio and Consolata Akoth Ariko.

Abstract

Inflation is increasingly becoming an important parameter that determines financial performance of central banks worldwide. Although many empirical studies have used time series and vector auto regression models to analyze inflation, very few studies have relied on the predictive ability of the logistic model which has a more intuitive interpretation for key issues such as drivers of inflation. The primary concern for every country is growing expectations about inflation which has resulted to a specific need for developing probabilistic models for studying the interaction between inflation and the key drivers. In Kenya, a similar scenario is concern is experienced. The purpose of the study was to develop a predictive binomial logistic regression model of inflation rate with published data from the Central Bank of Kenya (CBK), to determine the associations between CBK rates and inflation using the logistic regression model and to analyze the accuracy level of the using the model on predicting likelihood of inflation exceeding the set target of 5%. For this purpose, various rates were used as independent variables and inflation (categorized as either "on target" or "above target") as the dependent variable with a binomial distribution. A binomial logistic regression was performed to ascertain the effects of these variables on the likelihood that inflation would exceed the set target. Results showed that increasing deposit, repo and reserve repo rates were associated with a reduction in likelihood of inflation exceeding the target, but increasing overdraft rates, interbank rates, USD exchange rates and interbank rates were associated with increase in the likelihood of inflation exceeding the target. The model was statistically significant at $\chi^2 = 32.482, p < .000$ and explained 68.0% (Nagelkerke R-Squared) of the variance in inflation by correctly classifying 95.0% of the cases.

TABLE OF CONTENTS

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Kenya is East Africa's economic hub with a gross domestic product (GDP) that is majorly generated by services such as tourism and travel. Agriculture is equally important as it contributes to approximately a third of the GDP every year. The GDP figures and growth rate are usually reported to investors after adjustment to inflation has been calculated. Increase in inflation means reduction in purchasing power of the money, reduction in consumption and therefore a decrease in GDP. Higher than expected inflation rate creates uncertainty and affect balance of payments which result in a higher cost of service and goods compared to a lower inflation rate. Low and stable inflation facilitates economic growth and helps uplift the living standards and economic welfare of all the people.

The mandate of calculating inflation numbers vests in the Kenya National Bureau of Statistics (KNBS). The KNBS revised the methodology of calculating inflation from the arithmetic to the geometric mean method in October 2009. Weights of the respective baskets were later revised in February 2010 to reflect changes in households' behavioral patterns stemming from changes in consumption patterns and introduction of new products. For instance, the major categories of goods and services that seem to drive inflation include the Food and non-alcoholic beverages (36.04%), Housing, Water, Electricity, Gas and Other fuels (18.30%) and Transport (8.66%). These three constitute a total weight of 63 percent leaving the other 9 categories to share the remaining 37 percent of the total CPI weight. Significant price changes in one or more of the individual commodities in these categories lead to changes in inflation rates depending on the direction and magnitude of the price changes.

Monetary policies consist of decisions and actions taken by the CBK to ensure that the money supply within the economy is consistent with growth and price objectives set by the Kenyan government. The main objective of monetary policy is to maintain price stability in the economy consequently maintaining a low and stable inflation. However, inflation can also be caused by the failure of aggregate supply to equal the increase in aggregate demand. As a result, inflation can be controlled by increasing the supplies of goods and services as well as reducing money incomes so as to control aggregate demand for goods and services since it threatens economic stability of any county, prompting reaction from policy maker,[1].

The average inflation rate in Kenya was 4.69% in December 2018 as compared to 7.98% in December the previous year, recording a significant decrease. The average inflation over the past 5 years has been above the CBK's target of percent. It therefore follows that maintaining stable inflation in Kenya is still a challenge. These efforts have resulted in a large number of different models for inflation, which vary in terms of the variables assumed to drive inflation (including the Repo rates, reverse Repo rates, interbank rates, United States exchange rate or overdraft rates). In general, a variety of approaches have been used to model inflation, including single equation models, vector auto-regression (VAR) models and factor models. While there are advantages to systems approaches to modelling inflation, logistic regression models can also be very useful due to their simplicity. As an empirical estimation method, binomial logistical regression has been used in other areas of study such as deforestation analysis [2]; [3], agriculture [4]; [5], and urban growth modeling [6];[7].

The study explored the use of binomial logistic regression within the framework of the analysis of prediction accuracy of inflation rates as well as the contributions of varied CBK rates. Secondarily, the study established the importance of a target-based classification measure of inflation as a viable alternative to modeling inflation dynamics. As an empirical estimation method, binomial logistical regression has been used in several studies and proven to be very accurate.

1.2 Statement of the problem

Inflation modeling aims to understand its dynamic processes, and therefore interpretability of models used is becoming crucial. Statistical approaches such as us the use of VAR in time series models can readily identify the influence of independent variables like time and also provide a degree of confidence regarding their contribution. For the VAR models, the standard coefficients do not lend themselves easily into the interpretation of slopes in the classic regression equation. Other time series models such as the autoregressive integrated moving average and seasonal autoregressive integrated moving average have been constructed to demonstrate some form of a univariate stochastic behavior of inflation. However, inflation dynamics are not univariate in the sense that multiple variables contribute to the changes. Moreover, just because there is a time feature, it doesn't mean that inflation must only be a time series problem.

1.3 Objectives of the study

1.3.1 Main Objectrive

The purpose of the study was to develop a predictive logistic regression model of coordinated inflation rate with published data from the central bank that could be used by the public to project future scenarios.

1.3.2 Specific Objectives

i To determine the associations between CBK rates and inflation using the logistic regression model.

ii To examine the accuracy level of using CBK rates in predicting inflation.

1.4 Statical Hypothesis

H_0:There are no relationships between infation and the CBK rates.

H_a:There are relationships between inflation and the CBK rates

1.5 Research Questions

The following questions were addressed by the study:

1. How are CBK rates associated with the likelihood of inflation exceeding the set targets?

2. What are the efficiency levels of using these rates to predict inflation?

CHAPTER 2

LITTERATURE REVIEW

One category of inflation model is commonly referred to as the Phillips-curve model which has been a central model in macroeconomics. The original basis for this model is the framework developed by [8], which relied on the idea that low employment could be sustained by allowing high inflation to erode real wages and therefore boost labor demand.

Since the work of Milton Friedman, Edmund Phelps and others in 1968 presented a sharp critique of the Keynesian Philip curve. In particular, he criticized its treatment of expectations. He pointed out that if policy tried to keep output above its potential or equilibrium, then staff unions would get used to the high level of inflation and adjust their nominal wage demands upwards. This would result to higher inflation without the sustainable low employment. In 1970, empirical evidence seemed to subsequently back up Friedman's argument since there was a combination of high inflation and high unemployment that the Philips curve had seemed to rule out. Such models have been the pillar of inflation modeling in a number of central banks over the past three decades. This critique of the traditional Phillips curve, and the sense that it was due to inadequate modeling of expectations, was a major motivation for more research in 1970s led by Robert Lucas and Thomas Sargent. Besides rejecting the Phillips curve, these economists also questioned the whole basis for Keynesian economics, that is, the assumption that monetary policy could affect inflation. The study followed this latter approach by directly including monetary measures to control inflation expectations.

Another type of model that has been extensively used is the New-Keynesian Phillips curve (NKPC) which builds on the Phillips curve by deriving the model from microeconomic relationships between capacity utilization, costs, prices and nominal rigidities at the firm level. The resulting inflation equation then links the deviation of inflation from its expected level to either the output gap or real marginal cost.

The principal response of Keynesian economists to these theoretical critiques was to attempt to build models that incorporate rational expectations and that provide a microeconomic justification for monetary policy having at least short-run effects. The principal microeconomic rationale has been sticky prices. Without some type of price rigidity, it is difficult to rationalise the idea that there can be periods during which factors of production, such as labour, are under-utilized, with aggregate output being below its potential or equilibrium level. Once it is assumed that at least some prices are rigid, then not all markets are clearing instantaneously and aggregate output may sometimes be below what would be obtained when all prices move flexibly. Also, with sticky prices, an increase in the money supply can produce a short-run increase in real spending power and thus can boost real output.

The derivation of the closed economy version of the NKPC is now standard. In an open economy setting, the derivation is similar except that it is appropriate to view consumer price inflation as determined by a weighted average of domestic and imported inflation, with the latter equal to the change in real import prices and the former determined as per the standard closed economy NKPC, but with marginal cost approximated by both real unit labor costs and real import prices.

Though not the most realistic formulation of rigid or sticky prices, Calvo Pricing turns out to provide analytically convenient expressions, and has implications that are very similar to those of more formulations such that the form of price rigidity faced by a firm is as follows;

Each period, only a random fraction $(1-\theta)$ of firms is able to reset their price; all other firms keep their prices unchanged. When firms do get to reset their price, they must take into account that the price may be fixed for many periods. This is presumably done by choosing a log-price, z_t, that minimizes the "loss function"

$$L(z_t) = \sum_{k=0}^{\infty}(\theta, \beta)^k E_t(z_t - p^*_{t+k})^2 \tag{2.1}$$

where β is is between zero and one, and p^*_{t+k} is the log of the optimal price that the firm would set in period $t + k$ if there were no price rigidity.

In addition, [9] asserted that in empirical work, it is also relatively standard to include a backward looking inflation term in the NKPC to produce a better NKPC. The inclusion of both the level and change in real import prices through monitoring of changes in foreign currencies allows for a general specification, since imports or exports may fall in the major categories as weighted by the KNBS.

An alternative specification is to use a measure of the output gap as a proxy for real unit labor costs, given that these are proportional under certain conditions. Notably, it has been a point of considerable concern as to whether including the output gap defined as the deviation of actual gross domestic product from its probable level as the driving variable for inflation provides a better representation than other proxies for real marginal cost. This could reflect the possibility that either the output gap could be too imperfectly measured to be a useful proxy for real unit labor costs, or that the conditions under which they are proportional to may not hold.

Milton Friedman in the late 1960s challenged the Keynesian theory by arguing that government intervention would destabilize the economy. Monetarists strongly challenge the Keynesian view that government spending stimulates national output. Monetarists assume a crowding out effect of government spending on private investment, especially if the latter is deficit financed. The monetarist school of thought believes that the major cause of inflation is excess money supply. Monetarists advocate for the use of fixed money growth rate rules in order to ensure monetary stability in the economy which formed the foundation of our model.

Many studies on inflation in African countries researched on the quantity theory and the supply of money. Interest rates, Reserve ratios, Foreign exchange rates, and other fiscal and monetary measures are sometimes added to the models to account for exchange regimes and imported inflation through internationally traded goods. An example is [10] in Guinea, who investigated the short-term and long run relationship between inflation and money supply concluding that there was long-term positive relationship between money supply and inflation.

In Nigeria, [11], analyzed the main sources of fluctuations in inflation in Nigeria using the framework of error correction mechanism. They found that the lagged consumer price indices, expected inflation, petroleum prices and real exchange rate significantly propagated the dynamics of inflation in Nigeria. They also found that the level of output was insignificant in the parsimonious error correction model. However, the coefficient of the lagged value of money supply was found to be negative.

Sudan, [12], studied inflation dynamics caused by money supply using various models; the single equation model, the structural vector-auto regression model and a vector error correction model to check the robustness of the estimated parameters in the models since there was limited data coverage and possible structural breaks.

Angola, [13], used a Vector Error Correction (VEC) model, to analyze the factors that affect the inflationary process in Angola. The two concluded that the inflation path was largely affected by exchange rate movements. Results also showed that excess money supply led to increase in demand.

Ethiopia, [13] used monthly data from 2000-2009 to model inflation in Ethiopia by including error correction mechanisms for food and non-food prices. They separately specified long-run relationships for the monetary, domestic food, and external food and non-food sectors but ignored long-run effects of energy prices. They found that the external sector largely determines inflation in the long run. Specifically, domestic food prices adjust to changes in world food prices, measured in local currency and non-food prices adjust to changes in world producer prices. Domestic food supply shocks also have a strong effect on inflation but it is a short-run effect. The evolution of money supply does not affect food prices directly, though money supply growth significantly affects non-food price inflation in the short run. Hence, in the long run, money supply seems to be adjusting passively to demand.

Predictably, recent studies based on the quantity theory usually found that money supply certainly drives inflation. Another study by [15] also found that domestic food price increases were caused by aggregate demand in Ethiopia, except in 2007-2008 when cereal production was very low. Aggregate demand in turn, was driven by excess money supply. However, this was capped by strict monetary policies in the period of 2008 to 2009. [16] , supported the role of money supply by using estimating error correction models concluding that money supply was by far the most important cause of inflation, accounting for over 50% of the variation in the long run.

In Kenya, empirical studies such as; [17] used inflation as the independent variable of interest rate, using a first-order autoregressive model to simulate inflation. [17] linked the realization of inflation with other variables using a cascade approach. Wilkie's original model (1986) included dividends, dividend yields, and interest rates. [18] updated his earlier work by expanding on the structural form of the processes used to represent key variables in his "stochastic investment model." The research included several appendices that fully developed the time series tools used throughout the presentation including co-integration, simultaneity, vector auto-regression (VAR), autoregressive conditional hetero-scedasticity (ARCH), and forecasting. He also estimated parameters for each equation of the model by looking at data from 1923-1994 and performed tests on competing models for fit. As in the 1986 model, Wilkie's updated model simulated inflation as an autoregressive process which drives all of the other economic variables including dividend yields, long-term interest rates, short-term interest rates, real estate returns, wages, and foreign exchange.[19] studied the dynamics of inflation during 1974 - 1996; a period characterized by external shocks and internal disequilibria and found that inflation in was influenced by changes in maize-grain prices indicating a non-negligible role for agricultural supply constraints in the inflation process. They also found that the exchange rate, foreign prices and terms of trade determined inflation in the long run.[20] studied monetary policy reaction function . The study established that the central bank of Kenya had been targeting money supply, when making its monetary policy decisions. The results indicated that Central Bank of Kenya has been successful in controlling inflation, at least for the greater period in the sample. The study established that the CBK performed well in the implicit objective of short run interest rate management.[21] focused on the exchange rate and oil prices using the generalized Phillips curve discussed. The results showed that both variables drive inflation in the short run, but that the exchange rate is by far the most important variable. The study by [16] reported that monetary expansion is a key driver of inflation in and accounted for about 30 percent of the variation in the long run. In fact, the USD exchange rate seemed to explain a large part of the variation according to its coefficient, but no further details were provided.[22] estimated a VAR model that included GDP, money supply, fiscal expenditure and exchange and interest rates using innovation accounting and concluded that money supply was the main driver of prices of commodities. In addition, [23] reported results from a monetary model with more predictor variables. The parameters were calibrated which allowed for a more complex model specification.

According to previous preliminary findings, imported food price shocks and poor harvests explained some of the inflation dynamics, both in 2008 and 2011. However, [19] studied the dynamics of inflation in Kenya during 1974 – 1996; a period characterized by external factors and internal disequilibria and found that inflation was influenced by changes in maize prices indicating a non-negligible role for agricultural supply constraints in the inflation process. Besides, they also found that the USD exchange rate, import and export prices and terms of trade determined inflation in the long run. Largely, the findings were conflicting. For instance, studies by [24] and [25] who concluded that money supply was the major driver of inflation contradicts the study by [26] that found that exchange rate movements and changes in oil prices were the most important factors determining inflation while the contribution from monetary variables was quite small. Nonetheless, majority concluded that loose monetary policy drove inflation trend.

Related reviews, monetary measures include credit control, demonetization and issuance of new currency while fiscal measures include reduction in unnecessary expenditure, increase in savings, surplus budget and public debts.

One of the important monetary measures is monetary policy. The central banks adopt a number of methods to control the quantity and quality of credit. For instance, they could raise the commercial bank rates, sell securities in the open market, raises the reserve ratio, and adopt a number of selective credit control measures, such as raising margin requirements through Repos and Reverse Repos and regulating consumer credit. However, monetary policy can only be helpful in controlling inflation due to demand-pull factors. In this study, focus was selective money supply control measures and stabilization of Kenyan shillings against the U.S dollar.

Another monetary measure is to demonetize currency of higher denominations. Such measures are usually adopted when there is abundance of black money in the country. In extreme cases, the government issues a new currency in place of the old currency. In June 2019, the CBK demonetized old Ksh.1000 notes for new notes by setting an expiry deadline of 30^{th} September 2019. The value of commercial bank deposit rate is also fixed accordingly. Such a measure is usually adopted when there is an excessive issue of notes and there is hyperinflation within a country. It is believed to be very effective measure but inequitable for it hurts the small depositors the most.

The government may also reduce unnecessary expenditure on non-development activities as a fiscal measure in order to curb inflation. This also puts a check on private expenditure which is dependent upon government demand for goods and services. Though this measure is always welcome, it becomes difficult to distinguish between essential and non-essential expenditure. Therefore, this measure should be supplemented by taxation. Another measure that can be used is to increase savings on the part of the people. This tends to reduce disposable income with the people, and hence personal consumption expenditure. However, due to the rising cost of living, the people may not be in a position to save much voluntarily. The government therefore advocates for compulsory savings or provident fund cum-pension schemes where the saver gets his money back after some years through compulsory deductions submitted to National Social Security Fund (NSSF) among other pension schemes. All such measures may increase savings and are likely to be effective in controlling inflation. Adoption of surplus budget is also important. For this purpose, the government gives up deficit financing and instead has surplus budgets. It would mean collecting more in revenues and spending less. At the same time, the government would borrow more to reduce money supply with the public.

Fiscal and monetary measures alone cannot help in controlling inflation. Though not within the scope of this study, supply-push inflation can be controlled by non-monetary and non-fiscal measures which aim at increasing aggregate supply and reducing aggregate demand directly. These include;

(a) Increase Production

The following are measures that could be adopted to increase production:

- One of the foremost measures to control inflation is to increase the production of essential consumer goods like food, clothing, fuel, sugar, housing and transportation.
- If there is need, raw materials for the above named products could be imported on preferential basis to increase the production of essential commodities,
- Efforts should also be made to increase productivity. For this purpose, industrial peace should be maintained through agreements with trade unions, binding them not to resort to strikes for some time,
- The policy of rationalization of industries should be adopted as a long-term measure.
- All possible help in the form of latest technology, raw materials, financial help and subsidies should be provided to different consumer goods sectors to increase production.

(b) Rational Wage Policy

Another important measure is to adopt a rational salary and income policy. The government can implement this by freezing wages, incomes, profits, dividends, and bonus even though such a drastic measure is effective for short term solutions. A longer term solution would be to link increase in salaries to increase in productivity. The policy will control wages and at the same time increase productivity, and hence raise production of goods and services.

(c) Price Control

Price control and rationing is another measure of direct control to check inflation. Price control means fixing an upper limit for the prices of essential consumer goods. They are the maximum prices fixed by law and anybody charging more than these prices is punished by law.

(d) Rationing

Rationing helps in distributing consumption of scarce goods so as to make them available to a large number of consumers. It is applied to essential consumer goods such as cereals, milk and sugar with the aim of stabilizing the prices of necessaries and assures distributive justice. However, rationing is very inconvenient for consumers as it leads to artificial shortages, corruption and black marketing.

High inflation is detrimental to an economy since it is associated with the following effects:

i. Devalues real incomes of fixed income earners such as ”inua jamii” beneficiaries in Kenya and pensioners.

ii. Results in demand for salary increase and frequent strikes for increased salary by unionists to protect their members against the rising cost of living. A common example is frequent strike organized by Kenya National Union of Teachers (KNUT) and the Medical Practitioners.

iii. Lowers the overall purchasing power of the currency in terms of units of goods and services that a unit of that currency can purchase.

iv. Discourages local and foreign investors

v. Reduces the value of savings.

vi. Discourages lending as lenders of money make losses in periods of high inflation. Thus financial institutions are hesitant to lend during periods of high inflation resulting in lower production and declining economic growth.

Despite having many researchers compare models and use almost similar models (mostly time series models) in forecasting, the use of binary regression in forecasting in the country has been low. On the contrary, Logistic regression model seems to be more reliable in forecasting financial scenarios and business downfalls.

It is therefore safe to say that the effect of inflation very much depends upon the ability of people to predict it and there being mechanisms available for them to mitigate it. Inflation actually levies taxes on those who failed to anticipate it or who were in no position to protect them against it and redistributes it to those who were smart enough or lucky enough to anticipate it and take appropriate action. There is no obvious correlation between those who gain or lose from inflation.
"The denial to economics of the dramatic and direct evidence of the 'crucial' experiment does hinder the adequate testing of hypotheses; but this is much less significant than the difficulty it places in the way of achieving a reasonably prompt and wide consensus on the conclusions justified by the available evidence. It renders the weeding-out of unsuccessful hypotheses slow and difficult. They are seldom downed for good and are always cropping up again.[27]

This study will add to knowledge base, provide a lasting solution to policy makers and evaluate the effects of selected monetary measures on inflation in Kenya.

CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

3.1 Data collection

Secondary data was obtained from published data by CBK spanning from September 2013 to August 2018 (60 months). This was because some CBK rates were either too high or missing in the later years and therefore not consistent with current data.

3.2 The Binomial Regression Model

Since the probability of an event must lie between 0 and 1, it is impractical to model probabilities with linear regression techniques, because the linear regression model allows the dependent variable to take values greater than 1 or less than 0. The logistic regression model is a type of generalized linear model that extends the linear regression model by linking the range of real numbers to the 0-1 range.

Binomial Logistic model is used in dichotomous dependent variables. In this study we were interested in predicting the likelihood of inflation exceeding the set target. In the study, we used categories 0 and 1 where 0 means inflation falls within the target and 1 for inflation exceeding the target (5 percent).

Considering the independent variables deposit rates (DEP), overdraft rate (OD), Repo (REP), Reverse repo (RREP), interbank rates (IR) and USD rates (USD) and dependent variable, inflation (IF),We defined Yi to be equal to 1 if inflation was above target the target of 5 percent and 0 if inflation was equal to or below 5 percent.

We viewed y_i as a realization of a random variable Y_i that can take the values zero and one with probabilities $1 - \pi_i$ and π_i respectively.
The distribution was given by $Pr\{Y_i = y_i\} = \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$,for $Y_i = 0; 1$

Since we had 60 independent observations $y_1, ..., y_{60}$, and the i^{th} observation can be treated as a realization of a random variable Y_i. Y_i (Inflation) has a binomial distribution

$$Y_i \sim \beta(n_i, \pi) \tag{3.1}$$

with binomial denominator n_i and probability π and individual data $n_i = 1$ for all i. This defined the stochastic structure of the model.

The probability distribution function is given by

$$Pr\{Y_i = y_i\} = \binom{n_i}{y_i} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i}$$

The expected value and variance of inflation (Y_i) is given by

$$E(Y_i) = \mu_i = \pi_i, \text{ and}$$

$$var(Y_i) = \sigma_i^2 = \pi_i(1 - \pi_i) \tag{3.2}$$

The next step in defining a model for our data concerned the systematic structure. The probabilities of π_i depended on a vector of observed covariates x_i. We therefore assumed that π_i was a linear function of the covariates giving

$$\pi_i = x_i'\beta \tag{3.3}$$

where β was a vector of regression coefficients. The model (3.3) is called the linear probability model. This model was estimated from individual data using ordinary least squares (OLS).

One problem with this model was that the probability π_i on the left-hand-side has to be between zero and one, but the linear predictor $x_i'\beta$ on the right-hand-side can take any real value, so there was no guarantee that the predicted values wouldl be in the correct range unless complex restrictions were imposed on the coefficients. The solution to this was to transform the probability to remove the range restrictions, and model the transformation as a linear function of the covariates.

First, we moved from the probability π_i to the odds.

$$odds_i = \frac{\pi_i}{1 - \pi_i} \tag{3.4}$$

that is defined as the ratio of the probability to its complement, or the ratio of favorable to unfavorable response. If the probability of an event is a half, the odds are one-to-one or even.

Secondly, we took logarithms, calculating the logit / log-odds.

$$\eta_i = logit(\pi_i) = log\frac{\pi_i}{1-\pi_i} \tag{3.5}$$

The logit transformation is one-to-one. The inverse transformation is called the anti-logit, and allowed us to go back from logits to probabilities.

Solving for π_i in the equation,

$$\pi_i = logit^{-1}(\eta_i) = \frac{e^{\eta_i}}{1+e^{\eta_i}} \tag{3.6}$$

We assumed further that the logit of the underlying probability π_i was a linear function of the predictors.The logit of the underlying probability π_i is a linear function of the predictors.

$$logit(\pi_i) = x_i\acute{}\beta, \tag{3.7}$$

where X_i is a vector of covariates and β is a vector of regression coefficients. This defines the systematic structure of the model.

Exponentiating equation 3.7 we find that the odds for the i^{th} unit are given by

$$\frac{\pi_i}{1-\pi_i} = exp\{x_i\acute{}\beta\} \tag{3.8}$$

The model defined in Equations 3.7 and 3.8 is a generalized linear model with binomial response and link logit. Note, incidentally, that it is more natural to consider the distribution of the response Y_i than the distribution of the implied error term $Y_i - \mu_i$

3.3 Hypothesis Testing

The Wald test was used to test the hypothesis

$$H_0 : \beta_j = 0$$

And the significance of the coefficients of the predictor varriables by calculating the ratio of the estimate to its standard error

$$z = \frac{\hat{\beta}_j}{\sqrt{v\hat{a}r(\hat{\beta}_j)}}$$

The Omnibus Tests of Model Coefficients was also used to test the null hypothesis that there were no relationships between inflation and the selected CBK rates.The Pearson's chi-squared goodness of fit statistic, for the model was obtained by;

$$\chi_P^2 = \sum_i \frac{n_i(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i(n_i - \hat{\mu}_i)}$$

Such that each term in the sum is the squared difference between observed and fitted values divided by the variance of y_i,

3.4 Model Estimation

In the linear regression model, the coefficient of determination, R-Squared, summarizes the proportion of variance in the dependent variable associated with the predictor (independent) variables, with larger R-Squared values indicating that more of the variation is explained by the model, to a maximum of 1. For regression models with a categorical dependent variable, it is not possible to compute a single R-Squared statistic that has all of the characteristics of R-Squared in the linear regression model, so these approximations are computed instead.

Estimating the parameters for logistic models involved numerically approximating the solutions of nonlinear equations. The basis of the likelihood function for the 60 independent binomial observations was obtained by getting the product of densities given by the probability distribution function of inflation. The log likelihood was then obtained by taking logs of both sides.

$$logL(\beta) = \sum\{y_i log(\pi_i) + (n_i - y_i)log(1-\pi)\}$$

These were summarized using;

- **Cox and Snell's R-Squared** which is based on the log likelihood for the model compared to the log likelihood for a baseline model. However, with categorical outcomes, it has a theoretical maximum value of less than 1, even for a perfect model.

- **Nagelkerke's R-Squared:** This is an adjusted version of the Cox & Snell R-square that adjusts the scale of the statistic to cover the full range from 0 to 1.
- **McFadden's R-Squared:** This is also another version, based on the log-likelihood kernels for the intercept-only model and the full estimated model.

The classification table was also used to show the practical results of using the logistic regression model.

- Cells on the diagonal were results for correct predictions.
- Cells off the diagonal were results for incorrect predictions

3.5 Diagnotic Checks

The model having been identified and the parameters estimated, diagnostic checks were then applied to the fitted model. To determine whether the logistic regression model reasonably approximated the behavior of the data used, Hosmer-Lemeshow goodness of fit statistic reports were evaluated. This statistic is the most reliable test of model fit for binary logistic regression, because it aggregates the observations into groups of "similar" cases. Diagnostic residual plots such as the change in deviance versus predicted probabilities and Cook's distances versus predicted probabilities were used.

The measure of discrepancy between observed and fitted values is the deviance statistic, which was given by

$$D = 2\sum\{y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) + (n_i - y_i)\log\left(\frac{n_i - y_i}{n_i - \hat{\mu}_i}\right)\}$$

Where y_i is the observed and μ_i is the fitted value for the i^{th} observation. This statistic is twice a sum of (observed times log of observed over expected), where the sum is over both within target and in excess of the target. In a perfect fit, the ratio observed over expected is one and its logarithm is zero, causing the deviance to be equal to zero.

CHAPTER 4

RESULTS AND DISCUSSIONS

In building this model, six variables were selected based on their correlation with the dependent variable. The dependent and independent variables were classified as Table 4.1 and 4.2 respectively: Table 4.3 shows the statistical regression model for prediction, while Table 4.9 shows the logistic regression coefficients, Wald Test and Odds ratio for each of the predictors. Employing a 0.05 criterion of statistical significance repo, overdraft, USD, and deposit rates had significant partial effects. Inverting the odds ratio for each variable revealed that there was higher likelihood of inflation falling within the target of 5 percent for one point increase in deposit rate, repo and reverse repo rates. Conversely, one point increase in the USD exchange rate, overdraft and inter-bank rates increased, the likelihood inflation would exceed the target.

Reverse repo rate is the rate at which the central bank of a country borrows money from commercial banks within the country. It is a monetary policy instrument which is used to control the money supply by the CBK. An increase in reverse repo rate means that commercial banks will get more incentives to park their funds with the CBK, thereby decreasing the supply of money in the market .Repo rate is also used to control inflation. In the event of inflation, increasing repo rate acts as a disincentive for banks to borrow from the central bank. This reduces the money supply in the economy and thus helps in arresting inflation.

High inflation means that local goods increase in price quicker than imported goods. Therefore local goods become less competitive. Demand for exports will fall, and therefore there will be less demand for the foreign currency. Also, local consumers will find it more attractive to buy imports. Therefore they will supply more KES to be able to buy imports. This increase in the supply of KES decreases the value of KES and vice versa

Table 4.1: Dependent Variable Encoding

Original Value	**Internal Value**
Inflation on target of 5%	0
Inflation exceeds target of 5%	1

Table 4.2: Dependent Variables

Variable Name	**Description**	**Classes**
DEP	Deposit rates	Continuous
OD	Overdraft rates	Continuous
USD	USD exchange rate	Continuous
REP	Repo rate	Continuous
RREP	Reverse Repo rate	Continuous
IRP	Weighted interbank rates	Continuous

The general logistic regression equation for predicting odds

Table 4.3: Logistic regression equation for predicting inflation

$\log(\sqcap/(1-\sqcap))$
where
$\sqcap$ = -4.158 -4.189*DEP+0.372*OD+0.319*USD-0.239*REP-0.283*RREP+0.632*IR

From the 6 variables, the following estimates and their p-values are displayed below

Table 4.4: Significance of the variables in the logistic regression

		B	**S.E**	**Wald**	**df**	**Sig**	**Exp(B)**
Independent Variables	DEP	-4.189	1.537	7.426	1	0.036	0.015
	OD	0.372	0.349	1.135	1	0.028	1.451
	USD	0.319	0.226	1.993	1	0.0015	1.375
	REP	-0.239	0.203	1.388	1	0.039	0.788
	RREP	-0.283	0.38	0.553	1	0.057	0.754
	IR	0.632	0.404	2.444	1	0.078	1.882
	Constant	-4.158	22.053	0.036	1	0.058	0.016

From the intercept only model, the $ln(odds) = 1.494$. Exponentiation of both sides of the expression gives the predicted odds $[Exp(B)] = 4.455$ which was the predicted odds of inflation going above the target.

Table 4.5: The intercept only model

		B	**S.E**	**Wald**	**df**	**Sig**	**Exp(B)**
Step 0	Constant	1.494	.334	20.049	1	.000	4.455

The Omnibus Tests of Model Coefficients was used to test the null hypothesis that adding the CBK rates to the model had not significantly increased our ability to predict inflation. The Chi-Square of 32.482 on 1 df, was significant at .000. This is a. We therefore rejected the null hypothesis and accepted the alternative that these variables surely contributed to the model fit.

Table 4.6: Omnibus Tests of Model Coefficients

		Chi square	df	Sig
Step 1	Step	32.482	6	.000
	Block	32.482	6	.000
	Model	32.482	6	.000

In summary, the results showed that the -2 Log Likelihood statistics was 24.687. This statistic measured how poorly the model predicts the decisions the smaller the statistic the better the model. Although SPSS does not give us this statistic for the model that had only the intercept, the calculated value was 57.169. This confirmed that adding the 6 independent variables reduced the -2 Log Likelihood statistic by $57.169 - 24.687^a = 32.482$, the Chi-square statistic from the Omnibus Test.

Table 4.7: Model Summary

Step	**-2 Log Likelihood**	**Cox & Snell R-Square**	**Nagelkerke R Square**
1	24.687^a	.418	.680

Staying true to the objective of the study, the analysis of the accuracy of the model is also given below. Assuming that the two class labels of inflation $i \in 0, 1$, and p_i,
The classification process worked as follows:

I. A classification rule first assigns a score $s(x)$ to all applications based on a transformation of the probabilities.

   i. $f_i(s)$ represents the probability density function (PDF)

   ii. $F_i(s)$ represents the cumulative distribution function of the scores, where $F_i(s) = \sum_{i=0}^{1} P_i = 1$

   iii. $0 \leq 0.5 \leq 1$ represents classification threshold

II. A decision was made based on t •

i. If $s > 0.5$ the observation was classified as belonging to class 1

ii. If $s \leq 0.5$ the observation was classified as belonging to class 0

The introduction of the predictor variables to the model reflected an overall success rate of 95%. The results from inflation classification table showed that for each case, the predicted response was correct if that case's model-predicted probability was greater than the cutoff value of 0.5.Whereas a total of 57 out of 60 cases were correctly predicted, only 3 out of 60 cases were incorrect predictions.

The analyses of the observed and predicted inflation categories showed that 81.8% of inflation on target and 98% of inflation rates above target were correctly predicted.

Classification Table[a]

	Observed		Predicted		
			InfationCat		Percentage Correct
			On Target	Exceeds target	
Step 1	InfationCat	On Target	9	2	81.8
		Exceeds target	1	48	98.0
	Overall Percentage				

a.The cut value is .500

Test of Model Fit

Goodness-of-fit statistics helps to determine whether the model adequately describes the data. The non-significant Chi- square statistic in Table 4.8 indicated that the data fitted the model well.The Hosmer-Lemeshow tests the null hypothesis that predictions made by the model fit perfectly with observed group memberships.

Table 4.8: Hosmer and Lemeshow Test

Step	**Chi-square**	**df**	**Sig**
1	45.387	8	.60

CHAPTER 5

CONCLUSION

The explosion of computing power in modeling gives statistical analysts tremendous tools for more refined prediction ability unlike traditionally when deterministic calculations that were long and cumbersome were used. Modern approaches to statistical modeling begin by specifying the underlying economic and financial environments based on sophisticated mathematical equations, and then incorporate product-specific features which are commonly related to other conditions. This approach yields a much richer understanding of the associations between factors and covariates.

The logistic regression model and its underlying mathematical structure that was presented in the study surely provide an integrated framework for sampling from a wide range of future options. The model produced predicted coefficient values for various CBK rates and their associations to inflation in Kenya. This model can later be incorporated into a variety of applications. It is hoped that its application will form a basis to facilitate its use in recent advances in research, economic and statistical modeling.

REFERENCES

[1] F. Bashir, S. Nawaz, K. Yasin, U. Khursheed, J. Khan and M. J. Qureshi, "Determinants of inflation in Pakistan: An econometric analysis using Johansen co-integration approach," Australian Journal of Business and Management Research (AJBMR), p. 71, 2011.

[2] J. Geoghegan, S. C. Villar, P. Klepeis, P. M. Mendoza, Y. Ogneva-Himmelberger, R. R. Chowdhury, B. Turner Ii and C. Vance, "Modeling tropical deforestation in the southern Yucatan peninsular region: comparing survey and satellite data," Elsevier, pp. 25-46, 2001.

[3] L. C. Schneider and R. G. Pontius Jr, "Modeling land-use change in the Ipswich watershed, Massachusetts, USA," Elsevier, pp. 83-94, 2001.

[4] S. Serneels and E. F. Lambin, "Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model," *Elsevier*, pp. 65-81, 2001.

[5] S. J. Walsh, T. W. Crawford, W. F. Welsh and K. A. Crews-Meyer, "A multiscale analysis of LULC and NDVI variation in Nang Rong district, northeast Thailand," Elsevier, pp. 47-64, 2001.

[6] J. Landis and M. Zhang, The second generation of the California urban futures model. Part 1: *Model logic and theory, UK: London, England: SAGE Publications* , 1998.

[7] F. Wu and A. G.-O. Yeh, Changing spatial distribution and determinants of land development in Chinese cities in the transition from a centrally planned economy to a socialist market economy: a case study of Guangzhou, UK: London, England: *Sage Publications* , 1997.

[8] A. W. Philips, "The relation between unemployment and the rate of change of money wage rates in the United Kingdom 1861–1957," pp. 283-299, 1958.

[9] J. Fuhrer and G. Moore, "Inflation persistence," MIT Press, 1995.

[10] R. Blavy, "Inflation and Monetary Pass-Through in Guinea," IMF Working Paper WP/04/223, 2004.

[11] O. E. Olubusoye and R. Oyaromade, "Modelling the inflation process in Nigeria," *African Economic Research Consortium (AERC)*, 20008.

[12] K. Moriyama, "Investigating inflation dynamics in Sudan," *International Monetary Fund*, 2006.

[13] N. Klein and A. Kyei, "Understanding Inflation Inertia in Angola," I*MF Working Paper* WP/09/98. , 2009.

[14] D. J. Durevall, A. Loening and Y. A. Birro, "Inflation Dynamics and Food Prices in Ethiopia," *Scandinavian Working Papers in Economics (S-WoPEc)*, No 478, 2010.

[15] S. Rashid and N. Minot, "Are staple food markets in Africa efficient? Spatial price analyses and beyond," 2010.

[16] AfDB, "Inflation Dynamics in Selected East African Countries: Ethiopia, Kenya, Tanzania and Uganda," *AfDB Brief*. African Development Bank, Tunisia, 2011.

[17] A. D. Wilkie, "A stochastic investment model for actuarial use," Transactions of the Faculty of Actuaries, 39: 341-403., 1985.

[18] D. Wilkie, "More on a Stochastic Model for Actuarial Use," British Actuarial Journal, pp. 777-964, 1995.

[19] D. Durevall and N. Ndung'u, "A Dynamic Model of Inflation of Kenya, 1974-96," Journal of African Economies 10(1): 92-125, 2001.

[20] H. Rotich, M. Kathanje and I. Maana, "A monetary policy reaction function for Kenya," 13^{th} *African Econometrics Society Conference, Pretoria*, 2008.

[21] M. Kiptui, "Oil Price Pass-Through into Inflation in Kenya," Kenya School of Monetary Studies Research Centre, Nairobi., 2009.

[22] R. N. Misati, E. Nyamongo, L. Njoroge and S. Kaminchia, ”Feasibility of inflation targeting in an emerging market: Evidence from Kenya,” Journal of Financial Economic Policy, 4(2): 3 âĂŞ 19., 2012.

[23] IMF, ”Kenya: Third Review Under the Three-Year Arrangement Under the Extended Credit Facility and Request for Modification of Performance Criteria–Staff Report; Press Release,” Country Report No. 12/97, Washington DC: International Monetary Fund, 2012b.

[24] A. Aminu and I. A. Raifu, ”Firm Financial Status and Investment Behaviour: Evidence from Manufacturing Firms in Nigeria,” Sciendo, pp. 73-92, 2019.

[25] R. Barnichon and S. Peiris, ”Sources of Inflation in Sub-Saharan Africa,” *Journal of African Economies*, pp. 729-746, 2008.

[26] T. Ryan and W. Milne, ”Analysing inflation in developing countries: an econometric study with application to Kenya,” Taylor & Francis, 1998.

[27] M. Friedman and M. FRIEDMAN, ”Essays in positive economics,” *University of Chicago Press*.