
Bayesian Modelling of Kenya Extreme Debt with Correction for Budgetary Leakage

Matabel Odin Odiaga^{*}, Samuel Musili Mwalili, Joseph Kyalo Mung'atu

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email address:

odinmatabel@yahoo.com (M. O. Odiaga), samuel.mwalili@gmail.com (S. M. Mwalili), j.mungatu@fsc.jkuat.ac.ke (J. K. Mung'atu)

^{*}Corresponding author

To cite this article:

Matabel Odin Odiaga, Samuel Musili Mwalili, Joseph Kyalo Mung'atu. Bayesian Modelling of Kenya Extreme Debt with Correction for Budgetary Leakage. *International Journal of Data Science and Analysis*. Vol. 4, No. 5, 2018, pp. 98-105. doi: 10.11648/j.ijdsa.20180405.14

Received: October 24, 2018; **Accepted:** November 8, 2018; **Published:** December 4, 2018

Abstract: Total public debt levels in Kenya are exponentially increasing due to rising budget deficit, poor public fund management as well as movement of various macro-economic indicators such as balance of payments, inflation, Gross Domestic Product, exchange rates, and grants leading to worries on whether or not the high debt levels would be sustainable in future. The major concern is that a huge portion of the country's revenue is committed to debt repayment and budgetary leakage strains the repayment efforts, thereby accelerating the country's debt unsustainability. This study sought to model extreme debt in Kenya with correction for budgetary leakage using a Bayesian approach to Extreme Value Theory (EVT) the main aim being to estimate the maximum debt tolerable for the country. A non-stationary Generalized Pareto Distribution (GPD) model is used for modeling the public debt extremes which depend on some covariates (macro-economic indicators) and Bayesian methods used to directly estimate the threshold and the GPD parameters. A major contribution of this study is the introduction of a compensator to allow for possible leakage due budgetary leakage through corruption, tax evasion, money laundering, and other forms of financial fraud, modelling it as a function of budget deficit. The established debt threshold is approximately KShs. 2 trillion which is the standard amount that should be borrowed, beyond which values are considered extremes. The results indicate that the movements in the macro-economic debt indicators significantly affect total public debt levels, and that budgetary leakage reduces Kenya's debt tolerance. The research concluded that the current debt level of around KShs. 5 trillion is still sustainable but high budgetary leakage may accelerate the country's long-run debt unsustainability. For further work, it is recommended to use a time-varying threshold to capture seasonality of the public debt series.

Keywords: Extreme Value Theory, Non-stationary GPD, Bayesian, Debt Indicators, Budget Deficit, Budgetary Leakage

1. Introduction

Rising debt levels in Kenya has been the focus of much attention and contention because of its serious implications on development and sustainability. The rate at which the country's income is growing is not matching the rate of spending and this translates to rising debt levels to cover for the deficit. Consequently, Ngunjiri noted that total public debt in Kenya as at January 2018 had for the first time surpassed the KShs. 4.5 trillion threshold set by the National Treasury [1]. This could mean that the country's debt is approaching unsustainable levels. The main problem is that the government is forced to find new ways of raising revenue such as through increasing taxes, but budgetary leakage through tax evasion, corruption, money laundering and other

commercialized criminal activities by public officials' strains government repayment efforts, thereby accelerating debt unsustainability for Kenya. Besides poor public fund management, the high debt levels can also be attributed to various movements in macro-economic indicators such as inflation, balance of payments, grants, Gross Domestic Product (GDP), and exchange rates just to name a few. Ngunjiri verifies that there have been numerous discussions on whether Kenya, as a nation is able to sustain the current levels of debt and as to whether the economy has the capacity to service the outstanding government debt [1]. Consequently, the study of debt sustainability and management has become relevant. Nandelenga studied the issue of debt in Kenya and identified a positive relationship between debt burden and defaults [2]. The study suggested

that high debt levels may threaten a country's ability to meet its debt obligations, and in turn, lead to unsustainability.

Another study by Kiptoo revealed that rising debt levels is a major problem for developing countries like Kenya [3]. The study identified production variables such as GDP, exports, proportion of domestic and external debt as contributing factors towards debt levels, and stressed on the need to examine the effect of these determinants on debt sustainability. Matiti identified factors such as exchange rates, budget deficit, inflation, and balance of payments as the key determinants of public debt and debt sustainability in Kenya [4]. Mathenge examine exports, imports and revenue as debt indicators that would influence a country's debt sustainability [5]. The findings from these studies underscore the need of modeling public debt in relation to the production variables or macro-economic factors. However, Kiptoo notes that most developing nations that rely on debt for economic growth face deep rooted problems in corruption that further threatens their sustainability [3]. Case in point, Ochieng' indicates that Kenya loses approximately one third of its budget to corruption and almost a similar amount is lost through tax evasion, money laundering, bribery and other forms of financial fraud [6]. For that reason, budgetary leakage as a result of financial fraud directly influences the total debt levels in Kenya, and may pose a threat to debt sustainability for the country.

Most of the studies have examined the nature of government borrowing and debt sustainability by considering various production variables. It is, however, prudent to investigate debt sustainability of a country by adjusting for budgetary leakage to determine the true picture of a country's debt tolerance. The rising debt levels in Kenya reflect the behavior of extreme events that depend on some covariates. This necessitated the choice of extreme value theory (EVT) for the study. The use of EVT is not novel. Mathenge estimated the public debt maxima for Kenya using the generalized extreme value (GEV) distribution [5]. The parameters of the GEV were estimated using the maximum likelihood method and debt maxima obtained from return levels. Another study by Smith applied extreme value theory in the estimation of maximum rainfall [7]. The study used Bayesian methods to estimate the parameters of the extreme value distribution and pointed the main advantage of Bayesian estimation of relaxing the regularity conditions provided for by the likelihood method (i.e. the shape parameter only exists for values > -0.5). De Paola, et al. and Cheng Linyin, et al. use non-stationary extreme value analysis to examine climate extremes depending on some covariates [8-9]. Jonathan, et al., Davison and Smith, and Tawn also apply EVT in a non-stationary setting to study extreme events that depend on time or some covariates [10-12].

In this paper, a Bayesian approach to EVT is used to estimate the maximum amount of debt tolerable for Kenya. Specifically, the non-stationary Generalized Pareto distribution (GPD) model which assumes that public debt extremes depend on some covariates is used. The non-stationary GPD model is applied to capture the effect of movements of various debt indicators on the public debt

maxima for Kenya. Bayesian estimation is used to relax the regularity conditions of the likelihood method and to improve precision of the estimates. Based on existing literature, the primary contribution of this study is the introduction of a compensator to allow for possible leakage due to budgetary leakage and modelling it as a function of one of the covariates. The focus is to model public debt extremes over an optimum threshold, and subsequently determine the return levels or the debt limits under two models: one adjusted and the other unadjusted for budgetary leakage.

2. Methodology

2.1. Extreme Value Theory

Extreme value theory offers a theoretical foundation of describing the stochastic behavior of extreme events. It is broadly categorized into the Block Maxima (BM) and Peaks over Threshold (POT) methods. The BM method involves dividing the observation period into non-overlapping periods of the same size and studying the maximum observation of each period. The POT approach involves studying observations that exceed a certain threshold.

Suppose Y_1, Y_2, \dots, Y_n is a sequence of independent random variables from a common distribution, and $M_n = \max[Y_1, Y_2, \dots, Y_n]$. Extremal Types Theorem specified in Fisher and Tippet postulates that for a sequence of constants $\{a_n\}$ and $\{b_n\} \in \mathbb{R}$, the limiting distribution of $(M_n - b_n)/a_n$ is non-degenerate and belong to one of the three standard extreme value distributions: Gumbel, Fretchet and Weibull, which can be combined to form the Generalized Extreme Value (GEV) family of distribution [13]. Refer to De Haan and Ferreira for a detailed description of these distributions [14]. The GEV is used as a natural model under the BM method but may result in wastage of public debt extremes data. Thus, the POT approach is considered where the public debt extremes beyond an optimum threshold are modeled using the Generalized Pareto Distribution given as:

$$G(z) = \begin{cases} 1 - \left[1 + \frac{\xi z}{\sigma}\right]^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{z}{\sigma}\right) & \text{if } \xi = 0 \end{cases} \quad (1)$$

Where $\sigma > 0$ is the scale parameter, ξ is the shape parameter and $z = Y - u$ are the excesses over a threshold u .

According to De Haan and Ferreira, ξ is the most important parameter of the GPD as it determines the qualitative behavior of the tail of the distribution, such that if $\xi > 0$ the extremes have a Pareto distribution, when $\xi = 0$ an exponential distribution and when $\xi < 0$ a Pareto II type distribution [14]. This study focusses on public debt extremes which depend on some covariates or macro-economic indicators, hence the non-stationary GPD is applied. Covariate information can be added on either or both parameters but Northrop, et al. noted that observations beyond the threshold may not follow the GPD when covariate information are included in the shape parameter ξ [15]. Therefore, the covariates are modelled under the scale

parameter through the log-link function to ensure positivity.

$$\ln \sigma(x) = \phi_0 + \phi_1 X_1 + \dots + \phi_k X_k \quad (2)$$

The resulting non-stationary GPD model is given as

$$G(z) = \begin{cases} 1 - \left[1 + \frac{\xi z}{\sigma(x)}\right]^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{z}{\sigma(x)}\right) & \text{if } \xi = 0 \end{cases} \quad (3)$$

where $\sigma(x) = \exp(\phi_0 + \phi_1 X_1 + \dots + \phi_k X_k)$, X_1, \dots, X_k are the covariates, and $\phi_0, \phi_1, \dots, \phi_k$ are the coefficients of the covariates.

A compensator is introduced in the non-stationary GPD model to allow for possible budgetary leakage due to corruption, tax evasion, money laundering and other forms of financial fraud. According to Ochieng' financial fraud raises the country's annual budget deficit by inflating the total expenditures and deflating the total revenues [6]. The study therefore, modelled budgetary leakage as a continuous random process over an interval 0 to 0.5, such that budgetary leakage increases uniformly over the study period with 0 signifying minimum leakage and 0.5 representing maximum leakage. Budgetary leakage is defined by $\psi(\cdot) \sim \text{Uniform}(0, 0.5)$ and is a function of budget deficit in the non-stationary GPD model. The scale parameter in (3) is then defined as:

$$\sigma(x) = \exp(\phi_0 + \phi_1 X_1 + \dots + \psi(\phi_k X_k)) \quad (4)$$

where X_k is budget deficit. The underlying assumptions for the subsequent non-stationary GPD model are:

- (1) No multicollinearity among the covariates
- (2) The extreme observations $Z = \max[Y_1, Y_2, \dots, Y_n]$ are non-stationary to account for trends and shifts
- (3) The distribution of Y_1, Y_2, \dots, Y_n is non-normal, possibly light or heavy tailed.

$$\pi(\theta|y) = \left(\prod_{i=1}^n (1 - p_u)^{I(y_i < u)} \left(p_u \cdot \frac{1}{\sigma} \left(1 + \frac{\xi z}{\sigma}\right)_+^{-\left(1 + \frac{1}{\xi}\right)} \right)^{I(y_i > u)} \right) \times \left(\frac{1}{\xi u} a e^{-a(\xi+1)} \right) \times \left(\frac{p_u^{-0.5}(1-p_u)^{-0.5}}{\int_0^\infty p_u^{-0.5} e^{-p_u} dp_u \int_0^\infty p_u^{-0.5} e^{-p_u} dp_u} \right) \quad (8)$$

The uncertainty about parameter estimates is addressed by giving posterior weights to parameter values. The weights are defined by

$$\omega_i = \frac{\exp[\hat{T}(u_i)]}{\sum_{j=1}^k \exp[\hat{T}(u_j)]} \quad (9)$$

where $\hat{T}(u_j) = \sum_{r=1}^n \log \hat{f}(y_r | y_{(r)}, u)$, $y_{(r)}$ is the training data and y_r is the validation data. The training threshold with the highest threshold weight is chosen as the optimum threshold.

2.3. Non-Stationary GPD Parameter Estimation

Bayes theorem is used to obtain the posterior estimates of the parameters. As previously stated, covariate information is incorporated through the scale parameter, hence the parameters of interest in this case are the coefficients $\phi_0, \phi_1, \dots, \phi_k$ and the shape parameter ξ . To ensure posterior

2.2. Threshold Selection

The Bayesian leave-one-out cross-validation (LOOCV) as proposed by Northrop, et al., is used to select the best threshold based on the predictive ability at different extreme levels [15]. A threshold level is seen as a tuning parameter whose value is treated as known and fixed when the subsequent inferences are made. Under the Bayesian LOOCV method, the rate and number of threshold exceedances are considered, such that $u \sim \text{bin}(n, p_u)$ where p_u is $P(Y > u)$ and n is the number of exceedances leading to a Binomial Generalized Pareto (BGP) model with parameters, p_u, σ , and ξ .

Bayes theorem is used to estimate the parameters of the BGP and is defined by

$$\pi(\theta|y) = L(\theta; y, u) \times \pi(\theta) \quad (5)$$

Where $\theta = (p_u, \sigma, \xi)$, $\pi(\theta|y)$ and $\pi(\theta)$ are respectively, the posterior and prior densities of the BGP parameters, $L(\theta; y, u)$ is the likelihood function of BGP density and u is the training threshold.

The posterior inferences can then be used to show the differing parameter uncertainties across different thresholds. The generalized MDI prior for the GP parameters (σ, ξ) is considered and is defined by

$$\pi(\xi) \propto \frac{1}{\xi u} a e^{-a(\xi+1)} \quad (6)$$

For $a > 0$ and the Jeffrey's' prior, specifically the $\text{beta}(0.5, 0.5)$ for p_u which is a conjugate prior for the binomial distribution that yields a proper posterior density and given as

$$\pi(p_u) = \frac{p_u^{-0.5}(1-p_u)^{-0.5}}{\int_0^\infty p_u^{-0.5} e^{-p_u} dp_u \int_0^\infty p_u^{-0.5} e^{-p_u} dp_u} \quad (7)$$

The ensuing posterior density based on Bayes theorem is given by

propriety, the weakly informative independent normal priors with large variances are specified for ξ and $\phi_0, \phi_1, \dots, \phi_k$ such that;

The normal prior for ξ is defined by

$$p_n(\xi) \propto (2\pi\tau_0^2)^{-\frac{1}{2}} \exp - \left\{ \frac{1}{2} \frac{(\xi - \mu_0)^2}{\tau_0^2} \right\} \quad (10)$$

with mean μ_0 and variance τ_0^2 . The normal prior for $\phi_0, \phi_1, \dots, \phi_k$ is defined by

$$p_n(\phi_i) \propto (2\pi^* \tau_0^{*2})^{-\frac{1}{2}} \exp - \left\{ \frac{1}{2} \frac{(\phi_i - \mu_0^*)^2}{\tau_0^{*2}} \right\} \quad (11)$$

with mean μ_0^* and variance τ_0^{*2} . The posterior density based on Bayes theorem is given by

$$p(\xi, \phi|z) \propto L(\theta, z, X) p(\xi, \phi) \quad (12)$$

with the prior for ξ given in (10) and prior for ϕ_i s given in

(11). The likelihood function for the non-stationary GP model in (3) is given by

$$L(\Theta, z, X) = \left(\frac{1}{\sigma(x)}\right)^n \prod_{i=1}^n \left\{1 + \frac{\xi z}{\sigma(x)}\right\}^{-\frac{1}{\xi}-1} \quad (13)$$

$$p(\xi, \phi|z) \propto \left(\frac{1}{\sigma(x)}\right)^n \prod_{i=1}^n \left\{1 + \frac{\xi z}{\sigma(x)}\right\}^{-\frac{1}{\xi}-1} \times (2\pi\tau_0^2)^{\frac{1}{2}} \exp - \left\{\frac{1}{2} \frac{(\xi - \mu_0)^2}{\tau_0^2}\right\} \times (2\pi^* \tau_0^2)^{\frac{1}{2}} \exp - \left\{\frac{1}{2} \frac{(\phi_0 - \mu_0)^2}{\tau_0^2}\right\} \times (2\pi^* \tau_0^2)^{\frac{1}{2}} \exp - \left\{\frac{1}{2} \frac{(\phi_1 - \mu_0)^2}{\tau_0^2}\right\} \times \dots \times (2\pi^* \tau_0^2)^{\frac{1}{2}} \exp - \left\{\frac{1}{2} \frac{(\phi_k - \mu_0)^2}{\tau_0^2}\right\} \quad (14)$$

The posterior density is complex hence, computations are done using Markov Chain Monte Carlo methods via the Metropolis Hastings algorithm.

2.4. Model Validation

Quantitative methods including the Geweke and Heidelberger-Welch methods are used to determine whether the Markov chains converged. The Geweke diagnostic is used to check whether the first 10% and 50% of the simulations come from the same distribution. The test statistic is the Z-score. The Heidelberg and Welch diagnostic computes a test statistic that accepts or rejects the null hypothesis that the Markov chain is from a stationary distribution. A p-value > 0.05 for both methods means that Markov chains converged and come from a stationary distribution. The acceptance rate which represents the number of times a proposal value is accepted during the simulation process is also used to determine convergence diagnostics. An acceptance rate in the range 20-30% show that the Markov Chains converged.

2.5. Return Level Estimation

Suppose p represents the probability that a return level Z_p is exceeded once every year, then the return period is defined by $1/p$. In this case, p depends on the rate of threshold exceedances, $\lambda_u = \frac{\text{Number of threshold exceedances}}{\text{total observations}}$. For the non-stationary case, the effective return levels which represent the return levels that would be estimated if the model were stationary at the parameter values for the different values of the covariates are obtained. The return levels are computed as

$$Z_p = \begin{cases} u + \frac{\hat{\sigma}(x)}{\xi} \left\{ \left(\frac{\hat{\lambda}_u}{p} \right) - 1 \right\} & \text{for } \xi \neq 0 \\ u + \hat{\sigma}(x) \log \left(\frac{\hat{\lambda}_u}{p} \right) & \text{for } \xi = 0 \end{cases} \quad (15)$$

Since $\hat{\sigma}(x) = \exp(\hat{\phi}_0 + \hat{\phi}_1 X_1 + \dots + \hat{\phi}_k X_k)$, the return level is also a function of the covariates. It follows that a positive trend in the scale parameter causes an increase in Z_p as p decreases or when the return period increases. Similarly,

where $\sigma(x) = \exp(\phi_0 + \phi_1 X_1 + \dots + \phi_k X_k)$ and $\Theta = (\phi_i, \xi)$ and $z = Y - u$, the excesses over threshold u . The subsequent posterior distribution is given by

a negative trend in the scale parameter causes a decrease in Z_p as p decreases. This means that the covariates in the scale parameter have an impact on the return levels.

3. Results and Discussion

3.1. Exploratory Data Analysis

The data used was monthly financial data from the period September 1999 to May 2018, which constituted total public debt, both domestic and foreign debt, expenditures, revenues, grants, Gross Domestic Product, inflation, KES/USD exchange rate, exports and imports for Kenya.

3.1.1. Properties of Public Debt Data

The properties of public debt data were explored and presented in Table 1. The kurtosis for domestic, foreign and total public debt are all less than 3 indicating that data are not normally distributed and possibly come from a heavy or light-tailed distribution as postulated in assumption 3 of the non-stationary model. At 5% significance level, the ADF test results indicate that the null hypothesis of non-stationarity fails to be rejected hence public debt data are non-stationary. This corresponds to assumption 2 of the model.

Table 1. Descriptive statistics of public debt data.

	Domestic Debt	Foreign Debt	Total Debt
Length of Data	223	223	223
Composition	46.70%	53.30%	100%
Min (KShs. 'Mn)	183417	311953	502253
Max (KShs. 'Mn)	2447619	2563074	4901288
Mean (KShs. 'Mn)	766266	788015	1553744
Kurtosis	0.16	1.48	0.79
Skewness	1.11	1.62	1.37
ADF Statistic	-2.55	-2.74	-3.41
ADF p-value	0.881	0.727	0.602

3.1.2. Normality Tests Using Histogram and QQ Plot

Figure 1 shows the density and QQ plots of public debt data. It is observed that in all the variables, the histograms/density tend to be skewed to the right hence, evidence of a heavy/light tailed distribution. From the normal QQ plots, most of the points do not coincide with the reference line. This gives further evidence that the data comes from a heavy/light tailed distribution.

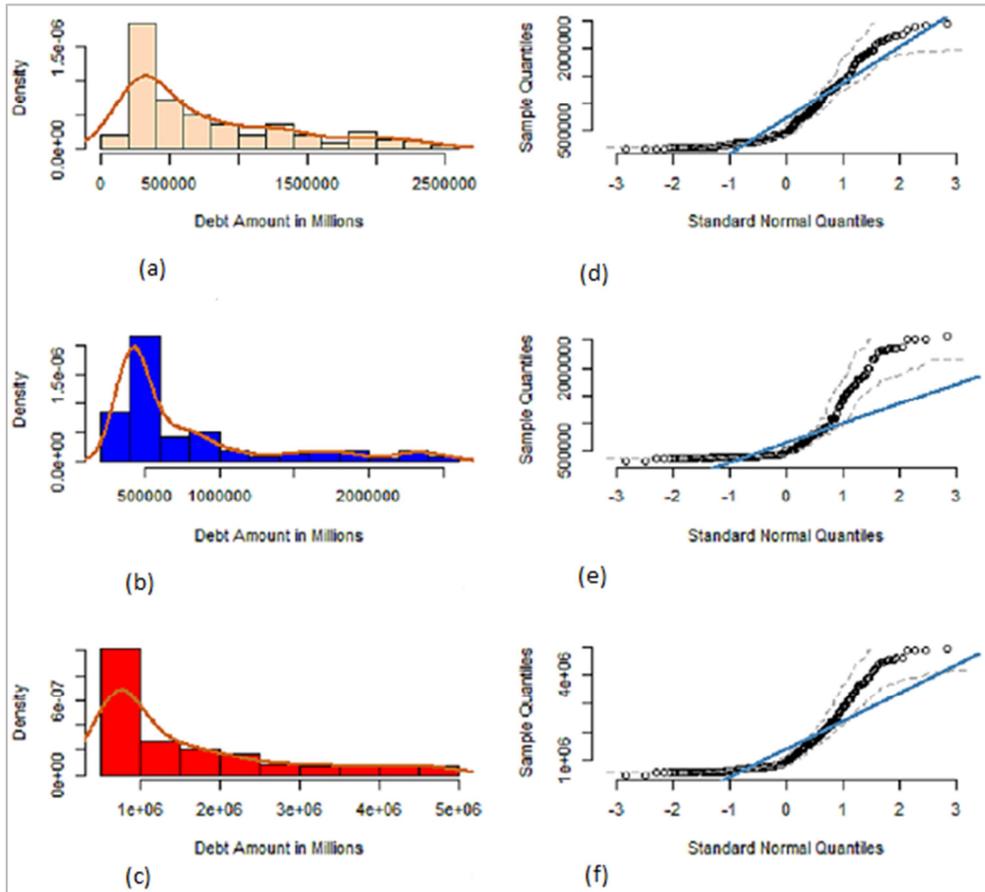


Figure 1. Density/histograms (a), (b), (c) and QQ plots (d), (e), (f) for domestic, foreign and total public debt respectively.

3.1.3. Trends in Kenya Public Debt Data

Time series plots of public debt data are shown in Figure 2. There is an increasing trend in total debt since 2000 to 2018 this confirms the non-stationary assumption of public debt series.

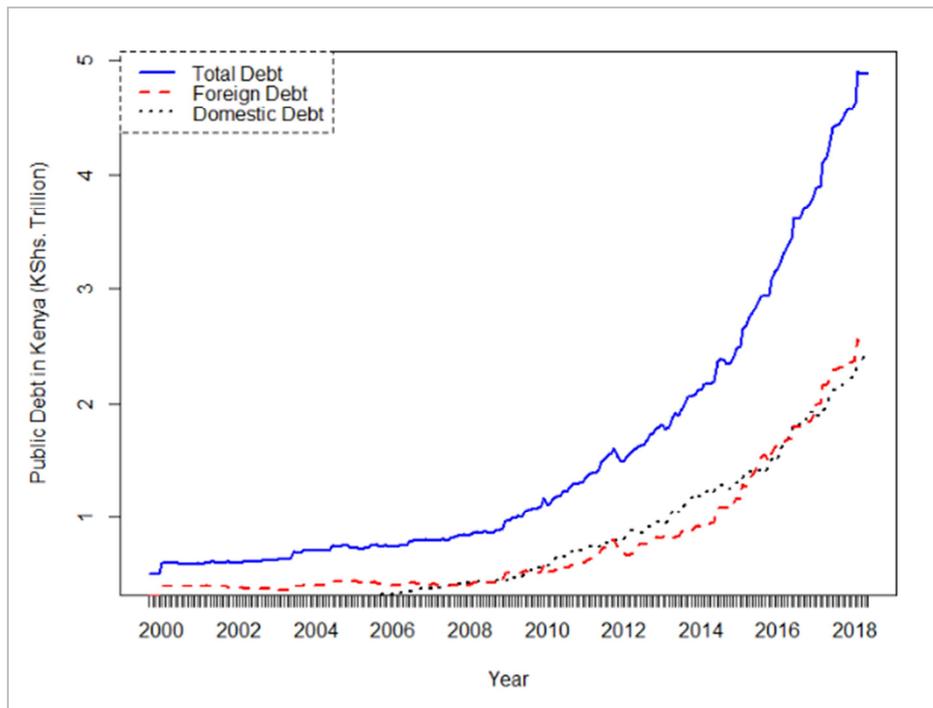


Figure 2. Time series plots for Kenya public debt.

3.1.4. Trends in Selected Debt Indicators

Computation of the Variance Inflation Factors (VIF) detected multicollinearity between the predictor variables which led to the elimination of variables such as revenues, expenditures, exports, imports and exchange rates. A $VIF > 5$ showed high multicollinearity. The remaining

covariates required to build a parsimonious model included balance of payments, inflation, grants, Gross Domestic Product and budget deficit. Time series plots of the selected debt indicators in Figure 3 show presence of trend and fluctuations over time implying that data are not stationary.

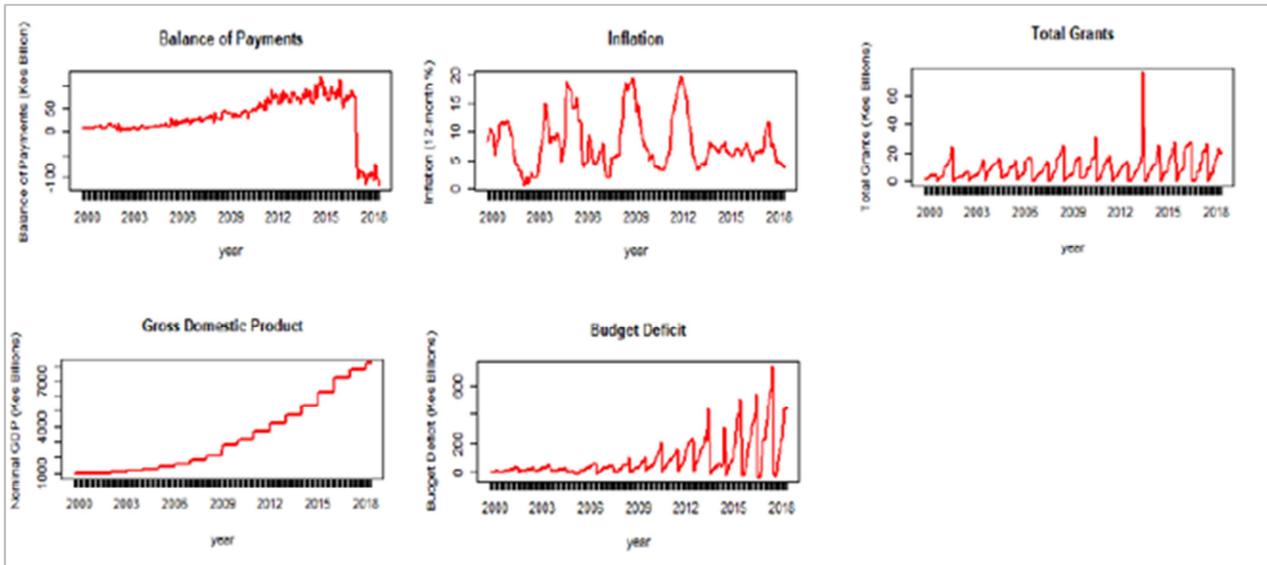


Figure 3. Time series plots of selected debt indicators.

3.2. Threshold Estimation

Computations are done by random sampling from the posterior distribution in (8) using R packages *threshr* and *revdbayes* published by Northrop, et al. [15]. 191 training thresholds set between 0% and the 95% sample quantiles were used. The Bayesian estimates of the parameters were obtained from samples of 10,000 iterations. The optimum threshold was determined at the 77% sample quantile where the training threshold had the largest threshold weight, and this corresponded to KShs. 1.8948 trillion. This is the standard or optimal amount to be borrowed by Kenya, beyond which the excess is modeled as public debt extremes.

3.3. Non-stationary GPD Parameter Estimation

Parameter estimates computed from the posterior density in (14) were obtained after 10,000 iterations based on different values of the proposal distributions. Poor starting values for the parameters were selected in the simulation process to minimize bias and the proposal values for standard deviation were continuously changed to allow for convergence of the Markov chains. The posterior means of the estimates are computed after a burn in of 500 iterations. Table 2 shows the results with estimates and 95% confidence interval of the shape parameter, coefficients of covariates in the scale parameter as well as the acceptance rates for each parameter for the adjusted model as given in (3) and (4), and the unadjusted model in (3).

Table 2. Non-Stationary GPD parameter estimates for adjusted (Adj.) and unadjusted (UnAdj.) models

Covariates/Parameter	Estimates	95% C.I	Acceptance rate
Intercept, ϕ_0 : UnAdj.	1.6471	[1.5880, 1.7152]	-
Adj.	1.1774	[1.1426, 1.2119]	-
Balance of Payments, ϕ_1 : UnAdj.	-0.0815	[-0.1290, -0.0342]	27.40%
Adj.	-0.0344	[-0.0674, 0.0007]	24.03%
Inflation, ϕ_2 : UnAdj.	0.7671	[0.7143, 0.8236]	25.26%
Adj.	0.1864	[0.1464, 0.2255]	25.69%
Grants, ϕ_3 : UnAdj.	-0.1269	[-0.1697, 0.0830]	26.80%
Adj.	-0.0306	[-0.0669, 0.0050]	25.68%
GDP, ϕ_4 : UnAdj.	0.0135	[-0.0420, 0.0733]	27.12%
Adj.	-0.0766	[-0.1193, -0.0351]	25.27%
Budget Deficit, ϕ_5 : UnAdj.	0.0418	[0.0004, 0.0834]	26.43%
Adj.	-0.0269	[-0.0588, 0.0038]	24.97%
Shape, ξ : UnAdj.	-0.9995	[-1.2036, -0.9751]	27.29%
Adj.	-0.9928	[-1.0158, -0.9702]	26.99%

The unadjusted model yields parameter estimates for the coefficients of the covariates in the scale parameter which show that inflation, GDP, and budget deficit increase the value of the scale parameter, hence, a unit increase in these variables increases the country's debt limit. Also, the results from unadjusted model indicate that balance of payments and total grants reduce the value of the scale parameter hence, a unit increase in these variables decrease the country's debt limit. The posterior estimates of the model adjusted for budgetary leakage showed some significant changes in the effect of the debt indicators. The results indicated a notable decrease in the value and direction of the coefficient of budget deficit, such that budget deficit decreases the value of the scale parameter. This means that by adjusting for budgetary leakage, a unit increase in budget deficit causes a decrease in the country's public debt limit. Also, correcting for budgetary leakage changes the impact of GDP on public debt limit, where a unit increase in the GDP causes a decrease in public debt limit. The impact of inflation, grants, and balance of payments also change slightly when budgetary leakage is accounted for, but

the direction of impact remains the same. The posterior estimates of the shape parameter ≈ 0.99 for both adjusted and unadjusted models showed evidence of heavy tailed distribution, specifically the Pareto II type distribution. The acceptance rates for all the parameters in the two models lie between 20 – 30 % which is an indication that the Markov Chains converged.

Table 3 shows the results of the quantitative diagnostics under both models. Geweke's test statistics are outside the critical region and the p-values for all the parameters are greater than 0.05. Values in parenthesis are p-values for the z-scores. The null hypothesis that the resulting Markov chains come from the same distribution hence, fails to be rejected. Under the Heidelberg and Welch diagnostic, all the parameters under both models passed the stationarity and half-width mean test. The p-values for all the parameters are also greater than 0.05, hence the null hypothesis of stationarity is not rejected implying that the Markov chains are from a stationary distribution. These results mean that the Markov chains converged, hence accurate estimates.

Table 3. Diagnostics by Geweke and Heidelberg-Welch methods for adjusted (Adj.) and unadjusted (UnAdj.) models.

Covariates/Parameter	Geweke ($Z_{0.025} = \pm 1.96$)	Heidelberg-Welch (p-value)
Intercept, ϕ_0 : UnAdj.	-1.41 (0.08)	0.82
Adj.	-1.41 (0.21)	0.72
Balance of Payments, ϕ_1 : UnAdj.	1.43 (0.08)	0.15
Adj.	1.43 (0.14)	0.30
Inflation, ϕ_2 : UnAdj.	-0.89 (0.19)	0.59
Adj.	-0.89 (0.27)	0.30
Grants, ϕ_3 : UnAdj.	0.57 (0.29)	0.18
Adj.	0.57 (0.29)	0.87
GDP, ϕ_4 : UnAdj.	0.09 (0.47)	0.28
Adj.	0.09 (0.14)	0.77
Budget Deficit, ϕ_5 : UnAdj.	-0.59 (0.28)	0.43
Adj.	-0.59 (0.06)	0.13
Shape, ξ : UnAdj.	0.83 (0.20)	0.68
Adj.	0.83 (0.20)	0.11

3.4. Return Level Estimation

The return levels corresponding to different return periods were obtained according to (15). The return levels are computed based on specified values of the covariates and different return periods, 2-year, 5-year, 10-year and 20-year return levels. Table 4 presents the effective return levels and return periods based on some specified covariate values for the unadjusted and adjusted models. The values in parenthesis represent the 95 % confidence intervals for the return levels.

Table 4. Estimated effective return levels for adjusted (Adj.) and unadjusted (UnAdj.) models.

Return period/level	Covariate values ≈ 1	Covariate values ≈ -1
2-year level: UnAdj.	9.63 [8.79, 10.67]	4.78 [4.62, 5.04]
Adj.	5.08 [4.79, 5.38]	4.60 [4.49, 4.77]
5-year level: UnAdj.	10.48 [9.55, 11.64]	5.10 [4.92, 5.39]
Adj.	5.43 [5.11, 5.77]	4.91 [4.78, 5.10]
10-year level : UnAdj.	10.77 [9.80, 11.96]	5.21 [5.03, 5.50]
Adj.	5.55 [5.22, 5.90]	5.01 [4.88, 5.20]
20-year level : UnAdj.	10.90 [9.93, 12.12]	5.26 [5.08, 5.56]
Adj.	5.61 [5.27, 5.97]	5.05 [4.93, 5.26]

The 2-year, 5-year, 10-year and 20-year return levels are different because of the influence of the scale parameter due to of presence of covariate information. The choice of 1 and -1 is used to show how the return levels change based on

different values of the covariates, but any other values can be selected as the results are approximately the same. The non-stationary model for both the adjusted and unadjusted case therefore, indicate the presence of a trend in the scale

parameter. The estimates of the return level also increases when p decreases or when the return periods are increased. The four return periods 2-year, 5-year, 10-year and 20-year correspond to probabilities 0.5, 0.2, 0.1, and 0.05 respectively. The return levels for model adjusted for budgetary leakage are significantly lower than those in the unadjusted model implying that budgetary leakage has a notable negative impact on public debt limit. For instance, the 2-year return levels when all the covariates are positive are 9.63 and 5.08 for unadjusted and adjusted models respectively.

4. Conclusion and Recommendation

In conclusion, the debt threshold of KShs. 1.8948 trillion established in the study is the standard amount of debt that should be borrowed by Kenya. Any values beyond this threshold are considered as extreme debt levels that may threaten the country's sustainability. Based on the results, it can be concluded that an increase in inflation increases the country's public debt limit or tolerance but the effect is more significant in the presence of budgetary leakage. Similarly, it can be concluded that increase in balance of payments (more exports less imports) and total grants decreases the country's annual debt limit and the effect is more significant in the presence of budgetary leakage. The effects of inflation, grants and balance of payments are, however, not of significant in examining the impact of budgetary leakage on the country's debt as the results showed a negligible change in the estimates of their coefficients after adjusting for budgetary leakage.

However, the study findings show a significant change in the estimates of budget deficit and GDP after correcting for budgetary leakage. Moreover, the much lower return level estimates under the adjusted model show the negative impact that budgetary leakage has on the public debt tolerance but the 2-year, 5-year, 10-year and 20-year return level estimates assuming an increasing trend, that is when covariate values ≈ 1 for all the debt indicators are greater than KShs. 5 trillion. It follows that at the current debt level of approximately KShs. 5 trillion, Kenya can still tolerate more debt as long as the country's GDP and budget deficit also increase simultaneously but may become unsustainable if budgetary leakage through tax evasion, corruption, money laundering and other financial fraud are not addressed.

For further work, it could be possible to use a time-varying threshold and perform the analysis on the whole time series to account for the effect of seasonality and help in examining the behavior of public debt series in different regimes. It is also recommended that a further study be conducted that computes the Value at Risk and Expected Shortfall for public debt series in a non-stationary setting to obtain different risk measures based on changing covariate values.

References

- [1] Ngunjiri, N., "What level of public debt is desirable for Kenya?," 26 February 2018. [Online]. Available: <https://www.businessdailyafrica.com/analysis/letters/What-level-of-public-debt-is-desirable-for-Kenya-/4307714-4320606-1axwhxz/inde>. [Accessed 2018 April 2018].
- [2] M. W. Nandelenga, "Debt sustainability and the optimal debt in Kenya," 2010.
- [3] C. K. Kiptoo, "The determinants of Kenya external debt sustainability," *Doctoral dissertation, school of Business, University of Nairobi*, 2012.
- [4] C. M. Matiti, "The effect of selected determinants on public debt in Kenya," 2013.
- [5] J. W. Mathenge, "Determine the breaking point of Kenya debt an application of extreme value theory," 2017.
- [6] L. Ochieng', "Kenya loses over Sh600bn every year in tax evasion," Nation Media Group, 21 May 2015. [Online]. Available: <https://www.nation.co.ke/business/Kenya-loses-over-Sh600bn-every-year-in-tax-evasion/996-2724818-5jswbtz/index.html>. [Accessed 25 April 2018].
- [7] E. Smith, "Bayesian modeling of extreme rainfall data," *Doctoral dissertation, Univeristy of Newcastle upon Tyne*, 2005.
- [8] F. De Paola, M. Giugni, F. Pugliese, A. Annis and F. Nardi., "GEV Parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in african test cities," *Hydrology*, vol. 5, no. 2, p. 28, 2018.
- [9] L. A. A. Cheng Linyin, E. Gilleland and R. W. Katz, "Non-stationary extreme value analysis in a changing climate.," *Climatic change*, vol. 127, no. 2, pp. 353-369.
- [10] P. Jonathan, D. Randell, Y. Wu and K. Ewans, "Return level estimation from non-stationary spatial data exhibiting multidimensional covariate effects," *Ocean Engineering*, vol. 88, pp. 520-532, 2014.
- [11] A. C. Davison and R. L. Smith, "Models for exceedances over high thresholds," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 393-442, 1990.
- [12] J. Tawn, "An extreme-value theory model for dependent observations.," *Journal of Hydrology*, vol. 101, no. 1-4, pp. 227-250, 1988.
- [13] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 2, pp. 180-190, 1928.
- [14] L. De Haan and A. Ferreira, *Extreme value theory: an introduction*, New York: Springer Science & Business Media, 2007.
- [15] P. Northrop, N. Attalides and P. Jonathan, "Cross - validatory extreme value threshold selection and uncertainty with application to ocean storm severity.," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 66, no. 1, pp. 93-120, 2017.