 https://doi.org/10.56279/orseaj.C2024.20

# Forecasting the Sustainability of Tax Revenue in the Context of Post-ICT Adoption: The ARDL and Markov Chain Model Approach

Grayson Kisinga[1], Mashaka Mkandawile[2] and Severine Kessy[3]

**Abstract:**
*In contemporary society, Information and Communication Technology (ICT) is widely recognised as an effective tool for optimising the process of collecting tax revenue. However, the empirical literature is not conclusive, as several studies have indicated mixed effects of ICT adoption on tax revenue. Arising from this, the study investigated the cointegration relationship between ICT investment, ICT imports, internet usage, broadband penetration, Economic growth, and Tax revenue in Tanzania from 1997 to 2022. Findings from Autoregressive Distributive Lag(ARDL) Bound tests indicated all variable are cointegrated and the Error Correction Model (ECM) estimates confirmed that ICT investments, ICT imports and Broadband penetration has positive and significant influence on tax revenue in long-run. Only Internet Usage was found to have negative influence on tax revenue reflecting a warning on the effect of e-commerce on the current tax systems, specifically in developing country. With existing ICT infrastructures, Markov Chain Model exhibits a stable state after 4 years. Beyond this point, tax will continue to increase at a consistent rate. The Markov model findings suggest that there is a 0.628 probability of tax revenue being in an increased state. Conversely, the analysis reveals a 0.372 probability of a decrease in tax, suggesting probable difficulties in generating long-term revenue. These findings evidence that governments and policymakers should use technological improvements to better anticipate changes in revenue streams using the Markov chain and ARDL Model. The study recommends governments integrate electronic systems and align with tax laws to online resources to automate and modernise their tax operations.*

**Key words:** Tax revenue, ICT adoption, ARDL Model, Markov Chain model

## Introduction

Predicting tax revenue is crucial procedure in the preparation process of national budgets and financial planning (Koniagina, 2020). Each government always operationalizes pre-determined revenue collection and expenditure figures throughout a given period of time (Glenday, 2013). Mostly, tax revenues comprise 95% of government revenues, therefore; accurate tax revenue assessments are extremely essential for efficient fiscal policy (Chikwede, 2022; European Central Bank, 2014; Khahro et al., 2020). As the collection of tax increases, it allows governments to conceive maximum developmental projects for the public interest, and improve the basic infrastructure including health facilities, education, and quality of life for common citizens. Tax revenue collection is constantly influenced by a variety of factors such as changing Economic growth, taxpayer compliancy, ICT adoption, national financial policy and other qualitative factors such regime changes (Gomero, 2022). ICT adoption incorporate infrastructure projects, such as broadband expansion or digital connectivity programmes (Kim et al., 2010).

According to Mills (2017), developing countries encounter substantial administrative obstacles in generating revenue, including outdated infrastructure and the digital divide in terms of access to infrastructures, limited expertise in tax authority resulting from a shortage of qualified personnel, and challenges in gathering, organising, and interpreting data and information. While ICT investment costs may initially reduce tax revenue through deductions and incentives, they can also encourage economic

[1] University of Dar es Salaam
Email: gkisinga@gmail.com
[2] University of Dar es Salaam
[3] University of Dar es Salaam

development, productivity, and innovation, resulting in increased tax income in the long term (Mayer et al., 2020). Furthermore, ICT investments allow governments to improve tax administration and compliance, which increases tax revenue collection (Nikiforova, 2022). In reality many activities in modern economy are random and repetitive and therefore uncertain (Kupelian & Loughbridge, 2017). In this regard tax predictions are always uncertain due to incomplete information about the relations between tax revenue states and the variables considered in prediction process, and also due to information about the value of these variables.

In recent years, many scholars have used different methods to predict the trend of income change, including financial income, tax budget income, and corporate income. Streimikiene et al. (2018) used time series method to forecast tax revenue in Pakistan. Sheng et al. (2021) used neural gray model and network method to predict fiscal revenue and improved the prediction accuracy in China, thereby making up for the shortcomings of traditional methods of insufficient prediction accuracy. In their study, Li and Xing (2021) employed Markov Chain analysis to accurately estimate the income of rural inhabitants. Their research not only provides data that supports the promotion of rural revitalization strategies, but also contributes to a better understanding of the income dynamics in rural areas. Other studies applied Markov chain theory to the actual market share analysis, they established Markov forecasting model of market share, constructs the mathematical model based on Markov prediction, and using the model on the empirical research to products sales in the market (Kovacs, 2018; Nuha et al., 2022; Yan et al., 2018). In addition, some of studies proposes a very tractable approach to modeling changes in regime, in their studies the parameters of an auto-regression are viewed as the outcome of a discrete-state Markov process (Cai, 1994; Hamilton, 2010; Li et al., 2017; Zheng et al., 2021).

On the other hand, several research employed Markov analysis in studying ICT investment strategy, enabling them to discern the pertinent stages of ICT advancement, including infrastructure level, user population, and technology adoption (Abed-Allah Migdadi, 2017; Chen & Rao, 2023; Dennis & Science, 2022; Eckelman & Daigo, 2008; Kagoya et al., 2024; Ryu et al., 2024; Zatonatska et al., 2022; Zhang & Chen, 2023). Some of studies employed Markov Chain to analyse financial sustainability, conceptualising a corporation's financial condition as a probabilistic process (Haller et al., 2020; Rekova et al., 2020; Rodriguez, 2011). The approach required identifying relevant financial states such as profitability, debt-to-equity ratio, and cash flow. To facilitate the Markov chain model, the states were divided into levels such as low, medium, and high.

Several scholars as Hanousek and Palda (2022) studied the evolution of tax evasion in the Czech Republic. Zohrah et al. (2024) conducted comparative study on a more feasible model in predicting tax revenue between Fuzzy Times Series (FTS) Markov Chain order in Indonesia. Theoretically, the Technology Organizational Environmental Framework (TOE) suggests the presence of ICT infrastructure are crucial in facilitating tax authorities to embrace and utilise technology for the purpose of generating revenue. However, no study has investigated the sensitivity of tax revenue to ICT investments using a Markov Chain Model. This study aims to examine the sustainability of ICT impact on tax revenue. Among others, this study establishes the cointegration between tax revenue and ICT adoption. The research is unique in that it uses established cointegration to predict tax revenue patterns over a ten-year period using stochastic modelling.

## Literature Review

### Theoretical Framework

Theoretically, the Technology-Organization-Environment (TOE) framework offers a complete perspective for comprehending the dynamics of ICT adoption in the specific context of tax revenue generating. The presence and complexity of ICT infrastructure are crucial in facilitating tax authorities to embrace and utilise technology for the purpose of generating income. Implementing automated tax collection, processing, and enforcement systems requires sufficient infrastructure, encompassing hardware, software, and communication. According to Park and Choi (2019), the TOE framework advocates that, the reliance on digital adoption like e-filing, e-payment, and electronic fiscal devices (EFDs) in tax administration is influenced by technological factors (e.g., features of digital technologies), organisational factors (e.g., tax authorities' willingness to incorporate these technologies), and environmental factors (e.g., regulatory and market conditions).

**Hypothesis Development**

**ICT Investment, ICT imports and Tax Revenue**

ICT investments frequently lead to increased economic activity and growth, which can generate tax revenue through a variety of avenues (Sawng et al., 2021).For example, greater productivity due to ICT costs might result in higher business profits, which generates more corporate income tax revenue (Vu et al., 2020). Furthermore, rising consumer expenditure owing to economic growth may lead to greater consumption taxes, such as VAT or sales tax (Nose & Mengistu, 2023; Sawng et al., 2021; Zhang et al., 2022). Nonetheless, empirical data falls short of complete consistency. For example, Olaoye & Atilola (2018) found no significant variation in tax collection before and after e-taxation was adopted. This indicates that the E-tax has not raised tax revenue in Nigeria. Also, Mallick (2021) found that, ICT infrastructures do not have a significant positive impact on overall tax revenue collections in India. Again, the transition from physical to digital trade gives rise to concerns regarding the potential loss of revenue resulting from tariff exemptions on digital imports, which significantly impacts developing countries (Choudhury, 2020; Suominen, 2017; Teltscher, 2002). However, tax revenue generated by the sale and use of ICT items can be significant, particularly in nations where these goods are in high demand (Adedeji & Lipede, 2023). Tariffs and duties imposed on imported ICT goods can have a direct impact on tax income.Nonetheless, a detailed analysis of the specific outcomes of the research, the methodology used, and the larger consequences of ICT deployment in tax administration would be required to reach a full conclusion on the topic at hand. As has been confirmed in this conversation, ICT can help generate tax revenue, but its impact is not universally positive. As a result, a non-directional hypothesis was developedt:

*$H_{1a}$: ICT investment has influence on tax revenue*
*$H_{1b}$: ICT imports has influence on tax revenue*

**Broadband Penetration and Tax revenue**

Broadband penetration is the percentage of households or individuals in a country or region that can access broadband internet(Kim et al., 2010). It measures population access to high-speed internet. Governments often tax broadband services via sales, excise, or value-added taxes(Pushkareva, 2021). Thompson-Abbott et al. (2023) argue broadband internet speeds up communication, collaboration, and information access. High-speed internet helps firms optimise operations, enter new markets, and adopt new technology (Qiang et al., 2009). Broadband Penetration can create jobs by boosting internet businesses, demand ICT services, and enable remote employment (Katz, 2018). Improved internet access can also attract investment and encourage entrepreneurship, boosting employment creation and economic growth(H. M. Stephens et al., 2022). Broadband provides widespread access to information, education, and research, encouraging innovation and knowledge sharing. This supports entrepreneurship and the interchange of ideas and skills between industries, which leads to breakthrough innovations (Stephens et al., 2022). The following hypotheses was developed based on above discussion.

*$H_{2a}$: Broadband penetration has positive influence on tax revenue*
*$H_{2b}$: Internet usage has positive influence on tax revenue*

**Methodology**

**Research Philosophy and Study Design**

Because the study outcomes are observable and quantitative, the positivist worldview was employed (Saunders et al., 2009).The current study employed a deductive technique, beginning with a general overview of existing hypotheses and then testing the collected data before reaching any conclusions. The time series analysis research design was used for this study because it focuses on evaluating the impact of a change on variables across time, as well as examining insights into variable relationships. According to Hudson et al. (2019), time series analysis helps researchers foresee future data by examining historical patterns, making it useful for predicting future scenarios.

**Data Collection**

During data collection, the research objectives and specific variables including tax revenue, ICT investments, ICT imports, internet Usage and broadband penetration were carefully established to guide the study's search for relevant secondary data. Variables were operationalized as shown in **Table 1.** The researcher then chose reputable and appropriate sources for secondary time series dataset. Data collection focus on Tanzania economy and the datasets mentioned included Tanzania Revenue Authority (TRA) and World bank (WB).

**Table 1: Operationalization of Variables**

| Variable | Dimensions/ Indicators | Indicative measures elaborations | Source |
|---|---|---|---|
| Tax Revenue | Total tax revenue | Revenue generated from various types of taxes in a particular country(Hill et al., 2022) | TRA |
| ICT investment | Investment cost on ICT items | Hardware and software(Leung & Fan, 2002; Rahman et al., 2021) | TRA |
| ICT import | ICT goods imports (% total goods imports) | Hardware and software (ITU, 2016; World Bank, 2018) | WBD |
| Internet Usage | Individual using Internet (%of population) | The % of individuals within a certain population that use the internet(Dutta & Lanvin, 2021) | WBD |
| Broadband Penetration | Fixed broadband subscriptions (per 100 people) | High-speed internet connectivity enables quick data transmission, e.g., DSL, cable, fiber optics (World Bank, 2009) | WBD |
| Economic growth | Gross Domestic Product (GDP) | The aggregate earnings of a nation's inhabitants during a designated timeframe(Kohli, 2003) | NBS |

**Data Analysis**

After gathering, the numerical data underwent several statistical procedures to summarise, organise, and test the final conclusions, as hypothesized in the study. The ARDL Model were implemented using Stata 14 software and the Markov Chain model in Maple 2022.

**ARDL Model**

The Autoregressive Distributed Lag model (ARDL) is a statistical model for analysing dynamic interactions between variables in time series data. It is an econometric model that combines autoregressive and distributed lag components to capture both the short-run and long-run relationships between variables(Shrestha & Bhatta, 2018). It was necessary to test the stationarity of timeseries before using the ARDL model (Afzal et al., 2010). The statistical procedure employed to determine the stationarity of a series is called the ‘unit root test’. Widely used stationarity test methods include Augmented Dickey-Fuller, Phillips Perron, and KPSS tests. However, the current study used only the Augmented Dickey-Fuller test. The unit root testing approach gives a suggestion for a method that might be used to make the time series stationary.

**Cointegration tests**

The cointegration test is an essential procedure in the ARDL model to verify that the variables are cointegrated and do not display spurious regression (Afzal et al., 2010; Shrestha & Bhatta, 2018). This is especially beneficial when the variables have a combination of different orders of integration, where some are stationary at (I(0)) and others are non-stationary at (I(1)). Cointegration denotes the presence of a stable and enduring relationship in the long term between two or more variables. The bound test is a widely employed technique for assessing cointegration in the ARDL model (Afzal et al., 2010). This test entails calculating the Bound F-statistic, which is employed to ascertain the presence of a long-term association among the variables. If the F-statistic is statistically significant, it suggests the presence of cointegration among the variables.

In the absence of cointegration between variables, the ARDL model may fail, resulting in inaccurate outcomes (Shrestha & Bhatta, 2018). Hence, it is imperative to conduct the cointegration test to verify the presence of cointegration among the variables and establish the suitability of the ARDL model. If the variables do not exhibit cointegration, different methodologies such as the Johansen and Juselius (1990) technique can be employed (Shrestha & Bhatta, 2018). Nevertheless, this approach necessitates the integration of variables that are of the same order, which may not always be feasible. After establishing cointegration, the ARDL model can be estimated by determining the proper lag length and selecting the model based on criteria such as the Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC), or Hannan-Quinn Criterion (HQC).

**Estimation of Model Parameters**

This study, adopted the ARDL model with the functional relationship expressed as

$$TR = F(INV, BRD) \quad (1)$$

Where TR is Tax Revenue, INVand BRD represents ICT investments and Broadband penetration respectively.The ARDL Model is different from the classical cointegration model in two respects: firstly, it can handle a mix of I(0) and I(1) variables; and secondly, if the data series are small samples, in both cases, ARDL proved to be a better model. Pesaran et al. (2001) primarily developed this model and we developed our ARDL model in the following form.

$$TR_t = a_0 + \sum_{i=1}^{p} b\Delta TR_{t-i} + \sum_{i=1}^{p} c\Delta INV_{t-1} + \sum_{i=1}^{p} b\,\Delta BDR_{t-1} + \lambda_1 TR_{t-1} + \lambda_2 INV_{t-1} + \lambda_3 BRD_{t-1} + U_t \quad (2)$$

Where i are indices of lags; $i = 1,2 \ldots. p, p$ is the optimum lag length, t denotes the time periods $t = 1,2 \ldots.., T$ and $U_t$ is the error term.

λ's represents the long-run dynamics of the variables. The cointegration can be established when λ is not equal to zero for the level variable in equation 2 above and F-test is used to test the joint significance of lag level variables.

After establishing cointegration, this study utilized ARDL error correction representation to estimate long-run relationships and short-run dynamics. Cointegration raises concerns about short-run fluctuations and long-run equilibrium adjustment speed. Fluctuations are particularly noticeable as a result of specific policy changes and the time it takes for them to become effective. An error correction term (ECT) in equation (3) captures these short-run modifications and provides significant information regarding long-run equilibrium changes:

$$TR_t = a_0 + \sum_{i=1}^{p} b_i\Delta TR_{t-i} + \sum_{i=1}^{p} c_i\Delta INV_{t-i} + \sum_{i=1}^{p} d\Delta IMP_{t-1} + \sum_{i=1}^{p} e_i\,\Delta INT_{t-i} + \sum_{i=1}^{p} f_i\,\Delta BRD_{t-i} + \sum_{i=1}^{p} g_i\,\Delta GDP_{t-i} + Au_{t-i} + v_t \quad (3)$$

The coefficients $a_i$to$g_i$ capture the short-run dynamics, with Δ showing how changes in independent variables affect the dependent variable TR in the short term. p represents the optimal lag length based on the Akaike information criterion(AIC). The term $Au_{t-i}$ capture the error correction term, is typically derived from long term relationship between variables in the model. This term adjusts for any disequilibrium in the long run relationship between TR and the independent variables. The coefficient A measures the speed at which the system recovers to equilibrium after a shock.$\boldsymbol{v_t}$ is the white noise error term, representing random shocks or error not explained by the model

**Markov Chain Model**

The Markov process is a branch of modern probability theory that deals with stochastic processes. Markov analysis is a method for predicting the future state of events by analyzing the evolution trend and state of a phenomenon. Andrei Markov (1856-1922) is particularly remembered for his study of Markov chains sequences of random variables in which the future variable is determined by the present variable but is independent of the way in which the present state arose from its predecessors. Using the Markov chain Model in prediction basically require understanding of the following concept State space, Markov property, transition probabilities, the initial state distribution and prediction process.

**State Space**

The state space of a Markov chain is the set of all potential states that the chain can be in. It reflects the entire set of unique states to which the system being modelled by the Markov chain can change according to specified probabilistic principles.The state space is normally discrete, which means it contains a countable number of unique states. Each state can be represented by a distinct identification or label. The state space can be finite, which means there are only a finite number of states, or countably infinite, which means there are infinitely many states that can be indexed by integers or another countable set. The state space may change over time, based on the Markov chain's specific rules. However, once defined, a stationary Markov chain's state space remains fixed.

The state space of a Markov chain is denoted by $S$ S in mathematics and can be defined as:

$S = \{ S_1 , S_2, S_3 , \dots \}$,each state in the state space is represented by $S_i$.

The state space is crucial for determining the transition matrix $P_{ij}$ of the Markov chain, which indicates the probability of moving from state $i$ to state $j$. The size and structure of the state space affect the complexity of Markov chain analysis and modelling, including stationary distribution computations, estimated time in states, and long-term behaviour forecasts. In conclusion, the state space of a Markov chain represents all possible states the system can occupy during its evolution, based on probabilistic criteria. It is a fundamental idea for understanding and analysing Markov chains in economics, engineering, biology, and other domains.

**Transition probabilities**

The state transition probability matrix of a Markov chain gives the probabilities of transitioning from one state to another in a single time unit. The $ij^{th}$ element of the transition probability matrix represents the *conditional probability* of the chain is in state j given that it was in state i at the previous instance of time. If the transition probability matrix doesn't depend on "n" (time), then the chain is called the *Homogeneous Markov Chain*.

Again: if the probability from state $S_i$ to state $S_j$ is $S_{ij}$, then $S_{ij}$ is the first-order transition probability from the state $S_i$ transfer to state $S_j$ after a period, where $S_{ii}$, $S_{jj}$ is transition probability at the same state, it is also known as the reserve probability. **Figure 1** and **Figure 2** shows the transition probability from one state to another state as p and q respectively

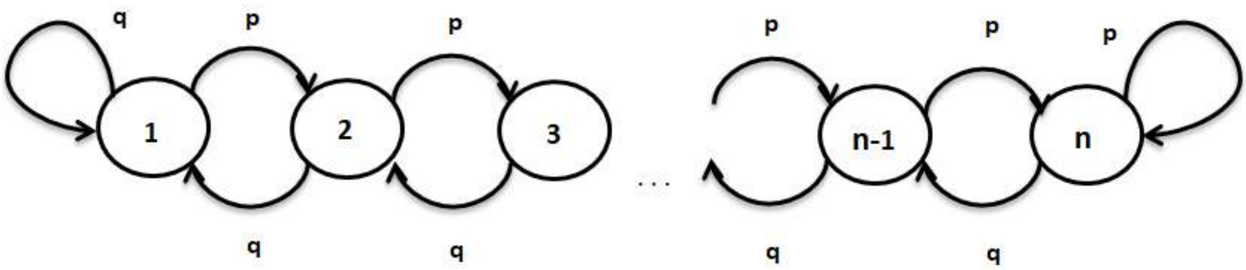


**Figure 2: Transition table**

| | | State to | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | . | . | n |
| State from | 1 | q | p | 0 | . | . | 0 |
| | 2 | q | 0 | p | . | . | 0 |
| | 3 | 0 | q | . | | . | . |
| | . | 0 | . | . | . | . | . |
| | . | . | . | . | . | . | . |
| | . | . | . | . | . | . | . |
| | n | 0 | 0 | 0 | 0 | . | p |

{ ***the element represents the probability of transition from state 2 to state 1*** }

**Figure 1: Transition diagram**

In general, transfer matrix composed of transition probability can be written as:

$$P(1) = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}, P(k) = \begin{bmatrix} P_{11}(k) & P_{12}(k) & \dots & P_{1n}(k) \\ p_{21}(k) & p_{22}(k) & \dots & p_{2n}(k) \\ \dots & \dots & \dots & \dots \\ p_{n1}(k) & p_{n2}(k) & \dots & p_{nn}(k) \end{bmatrix}$$

(1)$P \geq 0$; (2) $\sum_{j=1}^{n} P_{ij} = 1$, the sum of each columnis 1.

$P(1)$ is the first step transition probability matrix, $P(k)$ is the $k$ step transition probability matrix, it is the result of the transfer once again on the basis of step $k$–1, $P(k) = P(k-1) \times P(1) = P^K(1)$

**Markov Property**

The basic characteristic of a Markov chain is referred to as the Markov property or memoryless property. It suggests that the future behaviour of a stochastic process (or system) is determined solely by its current state, not by the sequence of events that preceded it. This property can be formally expressed as: The Markov Property states that a stochastic process $\{ X_t \}_{t \in T}$, where T represents the index set (typically discrete time steps), for any time $t$, and any sequence of values $x_0, x_1, \dots, x_t$;

$$P\left(X_{t+1} = {x_{t+1}}/{X_t} = x_t, X_{t-1} = x_{t-1}, \dots X_0 = x_0\right) = P(X_{t+1} = {x_{t+1}}/{X_t} = x_t)$$

In simpler terms, this means that given the present state $X_t$, the probabilities of transitioning to the next state $X_{t+1}$ depend only on $X_t$ and are independent of how the process arrived at $X_t$

$P(X_{t+1} = {x_{t+1}}/{X_t} = x_t)$ is called **transition probability**.

**Markov Prediction Process**

Markov chain only depends on the initial state and transition probability of the system. To predict the future states, firstly identify the current(initial) state of the system, then use the transition matrix to determine the probabilities of moving to each possible next state. For multiple steps into the future, apply the transition matrix iteratively to predict the state distribution at each future time step. The general prediction equation is given by

$S_j(k+1) = S_j(k) \times P_{ij}$ $(k = 0,1,2 \dots.)$ and in a vector form: $S(k+1) = S(k) \times P$

For prediction of dynamic evolution system (state probabilities after k+1 time steps), the model can be re-written as

$S(k+1) = S(0) \times P^{(K+1)}$, (K= 0,1,2…..)

The entry in the $i_{th}$ row and $j_{th}$ column indicates the probability of the system moving from the $i_{th}$ state to the $j_{th}$ state in k observations or trials.

In the long run a Markov Chain with transition matrix P has the property that $SP = S$. This is called stationary matrix, which means in the long-run the system will be at steady state. Later states will change very little, if at all. Not every Markov chain has a unique stationary matrix. However, if a Markov chain is regular, then it will have a unique stationary matrix and successive state matrices will always approach this stationary matrix. A transition matrix P is regular if some power of P has only positive entries.This means that after a certain number of steps, every state can be reached from any other state, making the system “irreducible” and “aperiodic”.A Markov chain is a regular Markov chain if its transition matrix is regular.

The steady state vector $\pi$ is found by solving $\pi = \pi \times P$ and $\sum_i \pi_i = 1$.The steady-state vector represents the long-term probability of being in each state, regardless of the initial state.

**Data Analysis and Interpretation**

**Descriptive Statistics**

The analysis began with common descriptive statistics used to summarize the central tendency and dispersion of a time series. The mean represents the average value of the series, while the standard deviation measures the spread of the data around the mean. From summary statistics **Table 2**, there is evidence of significant variation in the trends of the variable across a period that was taken into consideration. This is demonstrated by the significant disparity that exists between the series' minimum value and its highest value. It indicates that before further analysis of these data, it is important transform data in log form to reduce the variance and make them standard for efficient model analysis.

**Table 2: Summary Statistics**

| **Variables** | **N** | **Minimum** | **Maximum** | **Mean** | **Std. Deviation** |
|---|---|---|---|---|---|
| Total tax revenue | 25 | 0.51 | 17.62 | 6.23 | 5.98 |
| ICT investment | 25 | 0.41 | 22.22 | 7.86 | 7.70 |
| ICT goods imports | 25 | 3.08 | 6.92 | 4.62 | 1.11 |
| Internet usage | 25 | 0.01 | 31.63 | 6.47 | 8.63 |
| Broadband penetration | 25 | 0.11 | 0.50 | 0.32 | 0.12 |
| GDP | 25 | 8.15 | 161.53 | 59.16 | 51.63 |

Due to a scarcity of data, the study utilised a sample(N) of 25 observations spanning the years 1999 to 2021. To ensure the model's robustness, the number of lags was limited to a maximum of 2 in order to prevent overfitting and preserve more degrees of freedom. Additionally, robustness checks were conducted, including diagnostic and stability model tests.

**Goodness of Fit Test**

R-squared is a percentage statistic that indicates how much of the dependent variable's variance is explained by the model's independent variables. In this study as shown in Table 5, an R-squared found to be 0.905 indicates that the model's variables account for approximately 90.5% of the variability in tax revenue variations. This is a rather high R-squared value, indicating that the model is fairly good at describing the variation in tax revenue variations. The adjusted R-squared value accounts for the number of independent variables in the model while controlling for degrees of freedom. It penalises the R-squared score for having variables that add little to the model's explanatory power. Again, the results show a computed adjusted R-squared value of 0.7678 implies that, after controlling for model complexity, the independent variables still explain roughly 76.78% of the variance in tax revenue variances. This implies that, while the model is effective, there may be some areas for improvement or modification in terms of variable selection or model specification. Overall, our findings indicate that the model based on the ARDL and Markov chain techniques is a reasonable fit for explaining tax revenue variations due to ICT adoption.

**ARDL Model**

To verify for unit root problem, this study subjected all six variables to the test using the Augmented Dickey-Fuller(ADF) generalized least squares test by Dickey and Fuller, (1979) as shown in **Table 3.** ADF test is a parametric test, which assumes a correlation between error terms. For these tests, the null hypothesis of a unit root cannot be rejected if the test statistic is insignificant (e.i the rule of thumb is to reject the null hypothesis "there is a unit root" when the probability value is less than or equal 5%). The linear trend in our series is then eliminated by applying the first difference transformation.

**Table 3: Unit root test**

| Variable | ADF | | | | |
|---|---|---|---|---|---|
| | At level | | At First difference | | Decision |
| | Include Trend | Include Drift | Include Trend | Include Drift | |
| Tax Revenue | 0.9952 | 0.0715 | 0.4564 | 0.0466 | (1) |
| ICT Investiment | 0.6738 | 0.2184 | 0.0016 | 0.0001 | (1) |
| ICT imports | 0.1865 | 0.1644 | 0.0126 | 0.0004 | (1) |
| Internet usage | 0.0000 | 0.0001 | 0.0002 | 0.0000' | (0) |
| Broadband penetration | 0.8575 | 0.0076 | 0.0063 | 0.0037 | (0) |
| GDP | 0.876 | 0.1127 | 0.4173 | 0.0229 | (1) |

**Table 4: ARDL Bound test**

```
H0: no levels relationship            F =  7.988
                                      t = -5.463

Critical Values (0.1-0.01), F-statistic, Case 3

     | [I_0]   [I_1] | [I_0]   [I_1] | [I_0]   [I_1] | [I_0]   [I_1]
     |   L_1     L_1 |  L_05    L_05 | L_025   L_025 |  L_01    L_01
-----+---------------+---------------+---------------+--------------
 k_5 |  2.26    3.35 |  2.62    3.79 |  2.96    4.18 |  3.41    4.68
accept if F < critical value for I(0) regressors
reject if F > critical value for I(1) regressors

Critical Values (0.1-0.01), t-statistic, Case 3

     | [I_0]   [I_1] | [I_0]   [I_1] | [I_0]   [I_1] | [I_0]   [I_1]
     |   L_1     L_1 |  L_05    L_05 | L_025   L_025 |  L_01    L_01
-----+---------------+---------------+---------------+--------------
 k_5 | -2.57   -3.86 | -2.86   -4.19 | -3.13   -4.46 | -3.43   -4.79
accept if t > critical value for I(0) regressors
reject if t < critical value for I(1) regressors

k: # of non-deterministic regressors in long-run relationship
Critical values from Pesaran/Shin/Smith (2001)
```

After confirming that the variables were integrated in various orders, the ARDL bound test was used as shown in the Table 4. Thus, the Johansen cointegration test cannot be used.That is, having a combination of series integrated at order zero $I(0)$ stationary at level requiring no differencing, and series integrated at order $I(1)$ stationary after first differencing, we proceed to analyze if there is any cointegration among the variables using the ARDL bounds test approach (based on the error correction representation) as proposed by Pesaran and Smith 2001. Results in Table 4 show the calculated F-statistics of 7.988 is greater than the upper bound and lower bound at all level of significance, therefore there is cointegration between variables. That is, there are long run relationship. We reject a null hypothesis and estimate the long run model which is the error correction model (ECM).

**Table 5: ARDL Model Estimation**

```
ARDL(1,2,2,0,2,1) regression

Sample: 1999 - 2021                          Number of obs     =         23
                                             R-squared         =     0.9050
                                             Adj R-squared     =     0.7678
Log likelihood =  55.196855                  Root MSE          =     0.0351
```

| D.lnTotalTax | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| **ADJ** | | | | | | |
| lnTotalTax | | | | | | |
| L1. | -1.058092 | .1937 | -5.46 | 0.000 | -1.496272 | -.6199127 |
| **LR** | | | | | | |
| lnICTinvest | .151229 | .0601066 | 2.52 | 0.033 | .0152585 | .2871995 |
| lnICTimport | .3558966 | .1262037 | 2.82 | 0.020 | .0704039 | .6413893 |
| lnInternetUsage | -.1187787 | .0415002 | -2.86 | 0.019 | -.2126587 | -.0248987 |
| lnBroadband | .3608458 | .1014242 | 3.56 | 0.006 | .1314082 | .5902833 |
| lnGDP | 1.421795 | .1659814 | 8.57 | 0.000 | 1.046319 | 1.797271 |
| **SR** | | | | | | |
| lnICTinvest | | | | | | |
| D1. | -.1298516 | .0505784 | -2.57 | 0.030 | -.2442679 | -.0154354 |
| LD. | -.0741788 | .0355583 | -2.09 | 0.067 | -.1546173 | .0062596 |
| lnICTimport | | | | | | |
| D1. | -.2844153 | .1151235 | -2.47 | 0.036 | -.5448427 | -.0239879 |
| LD. | -.1471554 | .0938176 | -1.57 | 0.151 | -.3593856 | .0650748 |
| lnBroadband | | | | | | |
| D1. | -.2911917 | .1136333 | -2.56 | 0.031 | -.5482482 | -.0341353 |
| LD. | -.25822 | .0776523 | -3.33 | 0.009 | -.4338817 | -.0825582 |
| lnGDP | | | | | | |
| D1. | -1.117075 | .4333778 | -2.58 | 0.030 | -2.097444 | -.1367066 |
| _cons | 1.42549 | .3993882 | 3.57 | 0.006 | .5220107 | 2.328969 |

The results in **Table 5** indicate that in long run, there's a positive and significant influence of ICT investment on tax revenue ($\beta = 0.60$, $p < 0.001$), thus, supporting hypothesis $H_1$ of this study. This finding implies, total tax revenue rises by 60% for every 1% increase investment in ICT infrastructures. However, the short run dynamics indicates that, as the ICT investment's first difference (D1) increases by 1%, the total tax revenue decreases by 8.4%, and vice versa. This may implies during the initial implementation or update of ICT infrastructures, there may be substantial upfront expenses with no immediate compensations in the short term. Both governments and enterprises may be required to allocate funds towards infrastructure construction, training, and implementation. This investment could result in a temporary decrease in tax revenue as a result of increased spending. Over time, when these infrastructures are fully functional and incorporated into the economy, they have the potential to improve productivity, efficiency, and economic growth, resulting in increased total tax revenue.

**Table 5's** results also show that ICT imports statistically raise total tax income over time, hence supporting hypothesis $H_{2a}$. The total amount of tax revenue increases by 0.36 units for every unit rise in ICT imports in long run. This outcome can be attributed, in part, to the fact that ICT imports can increase tax revenue by generating jobs and by encouraging consumer spending as a result of people investing in new services and technologies. In contrast, table 4's findings for short run dynamics showed that when the ICT import's initial difference (D1) rises by 1%, the overall tax revenue falls by 28%, and vice versa. ICT imports may have a negative short-term impact on overall tax collection, nevertheless, because it need initial investment costs to acquire and set up ICT infrastructure.

Long-term Internet use significantly reduces total tax revenue, as Table 5 demonstrates ($\beta$ =0.12, p =0.019). The model contradicts H3a by showing that for every 1% increase in Internet usage (%) of population, overall tax collection decreases by 12%. Again, total tax revenue rises by 77.7% for every 1% increase in Broadband Penetration. The results in table 5 also indicate that, in long run there's a positive and significant influence of broadband penetration on total tax revenue ($\beta$ = 0.776, p =0.003). This finding implies total tax revenue rises by 77.7% for every 1% increase in broadband penetration, hence supporting the hypothesis $H_4$. Some of the reason may include improvement in quality and productivity, e-commerce and digital transactions, economic progress and innovation.

**Markov Chain Model**

After establishing the cointegration between total tax revenue, ICT investment and Broadband Penetration, this study used this background to estimate tax revenue patterns over ten years. The cointegration results implies that, variables have a stable long-term relationship, which can be used to forecast and predict future tax revenue trends. Cointegration also implies that short-term deviations from the equilibrium relationship will be addressed in the long run. The Markov chain model therefore, used to predict the tax revenue dynamic based on transitional probability established from lagged tax collection values. The key of predicting tax collection lies in determining the transition distribution of tax collection by Tanzania Revenue Authority for the period from 1996 to 2022. That is, the system state with the corresponding to a conditional probability when the state at time $t$ is change into that at time $t$ +1. This study used univariate examination to analyze a single variable, the total tax revenue as shown in **Table 6.**

**Table 6: Tax State Representation**

| PERIOD | SALES | 1ST DIFFERENCE | 2nd DIFFERENCE | STATES |
|---|---|---|---|---|
| 1996/1997 | 506,630.0 | | | |
| 1997/1998 | 560,818.1 | 54,188.13 | | |
| 1998/1999 | 616,265.3 | 55,447.22 | 1,259.09 | Constant(C) |
| 1999/2000 | 686,602.3 | 70,336.97 | 14,889.75 | Increasing(I) |
| 2000/2001 | 834,764.0 | 148,161.67 | 77,824.70 | Increasing(I) |
| 2001/2002 | 941,596.5 | 106,832.52 | (41,329.15) | Decreasing(D) |

| | | | | |
|---|---|---|---|---|
| 2002/2003 | 1,107,954.1 | 166,357.64 | 59,525.13 | Increasing(I) |
| 2003/2004 | 1,339,195.1 | 231,241.04 | 64,883.40 | Increasing(I) |
| 2004/2005 | 1,609,671.9 | 270,476.78 | 39,235.74 | Increasing(I) |
| 2005/2006 | 1,930,090.7 | 320,418.78 | 49,942.00 | Increasing(I) |
| 2006/2007 | 2,511,160.0 | 581,069.30 | 260,650.52 | Increasing(I) |
| 2007/2008 | 3,342,863.2 | 831,703.19 | 250,633.89 | Increasing(I) |
| 2008/2009 | 4,019,452.9 | 676,589.70 | (155,113.49) | Decreasing(D) |
| 2009/2010 | 4,406,910.3 | 387,457.35 | (289,132.35) | Decreasing(D) |
| 2010/2011 | 5,286,013.0 | 879,102.71 | 491,645.36 | Increasing(I) |
| 2011/2012 | 6,466,013.0 | 1,180,000.02 | 300,897.31 | Increasing(I) |
| 2012/2013 | 7,654,512.1 | 1,188,499.07 | 8,499.05 | Decreasing(D) |
| 2013/2014 | 9,358,404.5 | 1,703,892.45 | 515,393.38 | Increasing(I) |
| 2014/2015 | 9,888,445.3 | 530,040.77 | (1,173,851.68) | Decreasing(D) |
| 2015/2016 | 12,499,665.8 | 2,611,220.55 | 2,081,179.78 | Increasing(I) |
| 2016/2017 | 14,126,590.3 | 1,626,924.51 | (984,296.04) | Decreasing(D) |
| 2017/2018 | 15,191,421.3 | 1,064,830.91 | (562,093.60) | Decreasing(D) |
| 2018/2019 | 15,511,330.4 | 319,909.13 | (744,921.77) | Decreasing(D) |
| 2019/2020 | 17,622,822.1 | 2,111,491.69 | 1,791,582.56 | Increasing(I) |
| 2020/2021 | 17,624,361.6 | 1,539.53 | (2,109,952.16) | Decreasing(D) |

The first difference in **table 6** indicates whether the function is increasing or decreasing, whereas the second difference indicates if the first difference is increasing or decreasing. If the second difference is positive, it indicates that the first difference is increasing. Tax revenue collection can occur in three ways: rise, maintain steady, or decrease. This system recognises three states: increasing (I), constant (C), and decreasing (D). Using the obtained data, we estimated the transition probabilities and the initial state matrix ($S_0$).

**Table 7: Transition Matrix**

| | TO | | |
|---|---|---|---|
| FROM | I | C | D |
| INCREASED (I) | 7 | 0 | 5 |
| CONSTANT(C) | 1 | 1 | 0 |
| DECREASED(D) | 5 | 0 | 2 |
| TOTAL | 13 | 1 | 7 |

Table 7 shows the number of transitions from one state to another. Through period under study 1997 to 2021 we observed 7 transitions from Increased tax collection to Increased state, zero transitions from Increased to Constant state and 5 transitions from increased state to decreased state. Again, we observed 1 transition from constant state tax collection to Increased state, 1 transition from constant-to-constant state and 0 transitions from constant state to decreased state. And lastly, we observed 5 transitions from decreased state tax collection to Increased state, 0 transitions from decreased to constant state and 2 transitions from decreased state to decreased state. The corresponding transition probability matrices are calculated as

$P_{II}(k) = {}^{7}/_{13} = 0.54$, $P_{IC}(k) = {}^{0}/_{1} = 0.00$, $P_{ID}(k) = {}^{5}/_{7} = 0.71$, $P_{CI}(k) = {}^{1}/_{13} = 0.08$, $P_{CC}(k) = {}^{1}/_{1} = 1.00$, $P_{CD}(k) = {}^{0}/_{7} = 0.00$, $P_{DI}(k) = {}^{5}/_{13} = 0.38$, $P_{DC}(k) = {}^{0}/_{1} = 0.00$ $P_{DD}(k) = {}^{2}/_{7} = 0.29$.

$$P(K) = \begin{bmatrix} P_{II}(k) & P_{IC}(k) & \dots & P_{ID}(k) \\ p_{CI}(k) & p_{CC}(k) & \dots & p_{CD}(k) \\ \dots & \dots & \dots & \dots \\ p_{DI}(k) & p_{DC}(k) & \dots & p_{DD}(k) \end{bmatrix}$$

$$P(1) = \begin{bmatrix} 0.54 & 0.00 & 0.71 \\ 0.08 & 1.00 & 0.00 \\ 0.38 & 0.00 & 0.29 \end{bmatrix}$$

The initial state probability matrices were derived from the monthly tax revenue statistics.

**Table 8: Initial State distribution**

| 2021/2022 | TAX | DIFF | STATUS |
|---|---|---|---|
| 7 | 1,509,003.70 | | |
| 8 | 1,622,783.10 | 113,779.40 | I |
| 9 | 1,960,491.20 | 337,708.10 | I |
| 10 | 1,665,290.70 | (295,200.50) | D |
| 11 | 1,713,469.10 | 48,178.40 | I |
| 12 | 2,470,303.10 | 756,834.00 | I |
| TOTAL | 10,941,340.90 | | |

From Table 8 the initial probability for Increased tax revenue $P_I = {}^{4}/_{5} = 0.80$, For Constant tax revenue $P_C = {}^{0}/_{5} = 0.00$, and for Decreased tax revenue $P_D = {}^{1}/_{5} = 0.20$

$$S(K) = [P_I \quad P_C \quad P_D]$$
$$S(0) = [0.80 \quad 0.00 \quad 0.20]$$

Finally, carry on the forecast, obtain the first state matrix $S_1 = S_0 \times P$

$$\mathrm{P}_I(1) = P_{II}S_I(0) + P_{CI}S_I(0) + P_{II}S_I(0)$$

Correspondingly, the entries in the vector add to 1 since these three states represent the entire tax revenue collection possibilities. Finally, carrying on the forecast, we answer the big question, what will be the proportion of tax revenue collection state on the next period and the next period? This process of transition to each new state creates a Markov chain.

To obtain the first state matrix $S_1 = S_0 \times P$ . Simply multiply the vector times the matrix to find the proportion of each state on the next period. The outcome from maple 22 are as follows

$$\mathrm{S}(1) = S(0) \times P(1)$$

Vector[row] (3, [0.606, 0., 0.394]) (1)
Vector[row] (3, [0.6312199999999999, 0., 0.36878]) (2)
Vector[row] (3, [0.6279413999999999, 0., 0.37205859999999996]) (3)
Vector[row] (3, [0.6283676179999999, 0., 0.37163238199999993]) (4)
Vector[row] (3, [0.6283122096599998, 0., 0.3716877903399999]) (5)
Vector[row] (3, [0.6283194127441998, 0., 0.37168058725579983]) (6)
Vector[row] (3, [0.6283184763432537, 0., 0.37168152365674584]) (7)
Vector[row] (3, [0.6283185980753767, 0., 0.37168140192462285]) (8)
Vector[row] (3, [0.6283185822502007, 0., 0.37168141774979885]) (9)
Vector[row] (3, [0.6283185843074734, 0., 0.37168141569252594]) (10)

This sequence of 10 transitions in the Markov chain exhibits the property that $\pi = \pi \times P$ and that the system is irreducible, every state can be reached from any other state. This means, the system form a regular Markov chain and in the long run, the system will be at a steady state regardless of initial state.

The steady state in form of stationary vector indicates the equilibrium distribution of the system(total tax revenue) over a long period of time following ICT adoption.

It follows in the calculations above, over an extended period of time, for example in year one, the possibility of tax revenue will move to increasing state is 0.606 and to decreasing state is 0.394. However, the equilibrium state is reached in year 4, the system is likely to be increasing with a probability of around 0.628, no possibility of decreasing, and remain constant with a probability of roughly 0.372. Judging from the risk that tax revenue chance of decreasing of 0.372, which indicates potential long-term revenue challenges that may require policy intervention.

**Diagnostic and Stability Test**

The diagnostic and stability tests for an ARDL (Autoregressive Distributed Lag) model are critical steps in ensuring the model's results are reliable and valid. These tests serve to confirm the model's suitability and robustness in capturing long-term relationships between variables. In this investigation, the Serial Correlation Test and the CUSUM (cumulative sum) and CUSUM of squares tests were considered as shown .

**Serial Correlation**

The Breusch-Godfrey test is a statistical test employed to evaluate the existence of autocorrelation in the errors of a regression model. The purpose of this evaluation is to determine if there is any serial correlation that has not been taken into consideration in the model. This correlation, if present, could result in inaccurate conclusions or less than ideal estimations of the model's parameters. The test entails utilising residuals from the regression model to calculate a test statistic, where the null hypothesis assumes the absence of serial correlation up to a specific order. Interpreting the results of the Breusch-Godfrey test in entails examining the test statistics and p-values to ascertain whether there is autocorrelation in the residuals of a regression model. P-values play a significant role in establishing the statistical significance of the results by providing information about the test statistics. A low p-value of less than the significance level 0.05, indicates the presence of autocorrelation in the residuals and suggests that the null hypothesis of no autocorrelation is to be rejected. If the p-value is higher than the significance level, it indicates that there is insufficient evidence to reject the null hypothesis. This suggests that there is no substantial indication of autocorrelation in the residuals. In this study results as shown in table 9, the p-values of 0.9404 produced from the Breusch-Godfrey test which is higher than significant level of 0.05 indicates that there is insufficient evidence to reject the null hypothesis. Therefore, there is no substantial indication of autocorrelation in the residuals. This means this study results in accurate conclusions and perfect estimations of the model's parameters.

**Recursive Estimates Stability Test**

The cumulative sum test indicates if the regression coefficients are changing systematically. The Null hypothesis is that the parameters are stable (desirable). If the dotted is between or within the upper and lower boundaries, we accept the null hypothesis. The plot of the cumulative sum of the recursive residual is shown in **Figure 3**. The plots of the CUSUM and CUSUMQ residuals represent the boundary lines. This indicates that the parameter's stability has been maintained within their critical regions. The graphs in Figure 3 show that the CUSUM and CUSUMQ tests both support the long-run stability of the tax revenue function coefficients in both ARDL and Error Correction Models.

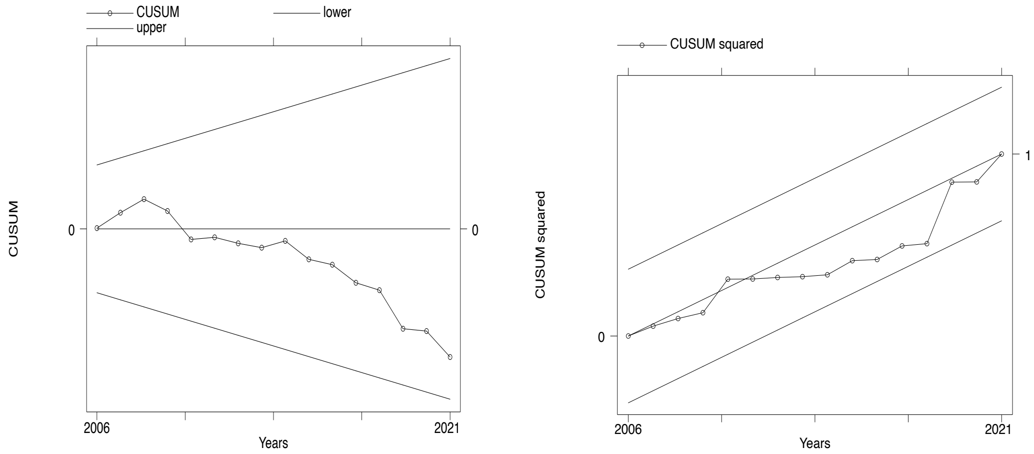


**Figure 3: CUSUM and CUSUMQ Test**

**Discussion of Findings**

The essential characteristic of sustainability, which is a broad concept with many applications in diverse contexts, is its long-term manifestation. Sustainability, in the context of tax revenue, is the capacity of governments or tax authorities to consistently sustain or enhance their tax collections by utilising improvements in ICT. Sustainability refers to the continuous evolution of economic and social systems, including economic growth and tax revenue collection. Analysing tax revenue requires considering long-term perspectives. For long-term tax collection to be sustainable, accompanying risks must not increase and become unmanageable. The findings revealed that the ARDL model is especially useful when examining the long-term interactions between ICT adoption and tax revenues, as it allows for the evaluation of both short-term and long-term dynamics. This model aids in understanding how changes in ICT adoptions correspond with tax revenue over time, while accounting for any lags in the consequences of ICT deployment with a control of economic growth.

ICT investment found to have positive and significant influence on tax revenue on long term therefore support both hypothesis **$H_{1a}$ and $H_{1b}$**. Although previous literature doesn't consider the long-term influence of ICT adoptions, the results in this study are supported by most of the studies like that of Nikiforova (2022) who revealed that, ICT investments promote the expansion of the digital economy, which includes industries like e-commerce, digital services, and online platforms. The growth of the digital economy has the potential to provide new sources of tax revenue, such as digital services taxes or online transaction fees (Lowry, 2019). Studies like that of Mallick (2021) and Olaoye & Atilola (2018) does not support this finding. However, the reasons for the inconsistency in conclusions on the impact of ICT adoption on tax revenue are related to the difficulty of aligning tax laws with the digital economy, whereby taxpayers may strategically plan their establishment in jurisdictions with a high level of tax exemption. Furthermore, most previous study did not consider the long-term impact of ICT adoption, despite the fact that ICT has a beneficial effect on tax revenue. For example, initial investment cost of ICT infrastructure may delay immediate benefit from ICT investment. On other hand, importing ICT items can boost economic activity by giving enterprises access to innovative technologies and equipment (Ozcan, 2018). This can lead to greater productivity, innovation, and competitiveness, which may benefit tax collection by increasing business profits and economic growth.

While Internet usage and broadband penetration are both related to the adoption of ICT infrastructure, the current study findings suggests that only broadband penetration has a long term beneficial impact on total tax.Thus supporting hypothesis **$H_{2a}$** but rejects hypothesis **$H_{2b}$** of the current study. This may be attributed to improvements in productivity, the expansion of the tax base, economic growth, and the innovation that comes with high-speed internet infrastructure. On the other hand, the adverse impact of internet usage on tax collection could be due to the substitution effects and fluctuations in consumer behaviour that impact traditional tax sources especially if governments fail to adjust their tax rules to accommodate digital transactions. This finding is supported by Pushkareva (2021) that broadband expansion improves productivity and economic growth by strengthening information access, e-commerce, and company efficiency. Increased economic activity raises corporation, income, and consumption taxes. Also, it is supported by other studies like that of Stephens et al. (2022) that access to broadband encourages entrepreneurial activity and facilitates the growth of businesses, especially in industries that rely on digital technology and knowledge-based industries. Countries that have a high degree of broadband penetration are generally considered more appealing to foreign investors. Nevertheless, the utilisation of the internet for shopping purposes may lead to a decline in traditional retail tax revenues if consumers opt to exclusively make purchases online or from foreign sources.

Again, this study showed that the Markov Chain model can be used to predict the sustainability of tax revenues by modelling the transitions between various phases of revenue performance. This approach provides insight into the likelihood of sustaining or increasing tax revenue levels following ICT adoption, informing policymakers about the possible risks and advantages of such investments. The Markov model findings suggest that there is a 0.628 probability of tax revenue being in an increased state. This indicates tax authority is likely to be in a favourable position of collecting more tax revenue in the long term. These findings are supported by the result from ARDL model in the current study and other previous studies, such as the one conducted by Mapesa et al. (2020), which suggests tax revenue may increase as a result of ICT adoption in Tanzania. Conversely, the Markov chain analysis reveals a 0.372 probability of a decrease in tax collection, suggesting probable difficulties in generating revenue in long term. This aligns with research indicating that tax revenue tends to decline following ICT adoption (Olaoye & Atilola, 2018).

The findings demonstrate that the Markov Chain is in a steady state, indicating that it is regular and may be utilised to predict future tax revenue based on present tax revenue conditions. This is possible because any state within the system can be reached from any other state, making the system "irreducible". This is supported by studies like that of Junquera-Varela et al. (2024) analysing the sustainability of EU economies. Junquera-Varela et al. (2024) utilizing the Markov chains, demonstrates that a process of convergence will occur inside each country. The EU28 Member States will advance at varying rates, resulting in either convergent or divergent trends in regard to developed economies, the European average, and each other. Findings also show different tax revenue scenarios and their probabilities of occurrence overtime can be analysed. This means the probability of being in an increasing state of 0.628 is answering very crucial question for a tax authority, that is, how likely is that the tax revenue will move from decreased to increased state in the next years. On the other hand, showing the probability of being in decreasing of 0.372 identify the possibility of extreme revenue state which request the authority to prepare mitigation strategies. By adjusting the transition probabilities to reflect different policy scenarios, the Markov chain can be used to assess the impact of potential changes in tax policy or economic condition on tax revenue predictions.

**Conclusion, implication and Further studies**

The integration of ICT in tax administration has potential for improving revenue sustainability, although its efficacy is impacted by various contextual circumstances. The ARDL and Markov Chain models offer useful frameworks for examining the relationship between ICT use and the sustainability of tax revenue collection. By integrating economic growth as a mediating variable, the analysis isolates the specific contribution of ICT adoption to changes in tax revenue, strengthening the conclusions. Difference from previous studies, the use of Markov chain model in this context demonstrates a probabilistic approach to evaluating changes in ICT adoption levels and their consequences on tax revenue regimes. Markov chains used to provide insights into the stability and predictability of tax revenue patterns after changes in ICT adoption. Previous studies did not address the inconsistence conclusions on the influence of ICT on tax revenue. In line with this argument, first, the study adds to the literature by diving deeper to investigate the influence of ICT adoption on tax revenue in short time and long-term. Secondly, the finding of this study contributes to advancing TOE framework serving as a robust theoretical lens to analyses the adoption of ICT in tax administration. The study is valuable, in particular to governments and policymakers might use technological improvements to better anticipate changes in revenue streams.Furthermore, the findings highlight the necessity of incorporating ICT strategies into fiscal planning and management processes to improve revenue forecasting accuracy and optimise fiscal policy decision-making. Further investigation should be conducted to examine these processes, with a specific emphasis on the interaction of technology, governance, and human capability in attaining sustainable tax revenue results

**References**

Abed-Allah Migdadi, Y. (2017). The Adopting of Markov Analysis to Forecast the Operations Competitive Advantages of Mobile Phone Service Providers: The Case of Jordan. *Operations and Supply Chain Management: An International Journal*, *10*(4), 214–225.

Adedeji, A., & Lipede, O. (2023). Can Information Communication Technology Unlock Tax Revenue Mobilization in Sub Saharan Africa. *DBN Journal of Economics and Sustainable Growth*, *5*(3), 1–27.

Afzal, M., Farooq, M. S., Ahmad, H. K., Begum, I., Quddus, M. A., Ahmad, H. K., Begum, I., & Quddus, M. A. (2010). *Relationship between school education and economic growth in pakistan : ARDL Bounds Testing Approach to Cointegration Stable URL : http://www.jstor.org/stable/41762413 Linked references are available on JSTOR for this article : And economic growth in paki*. *48*(1), 39–60.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, *12*(3), 309–316.

Chen, J., & Rao, V. R. (2023). Evaluating strategies for promoting retail mobile channel using a hidden Markov model. *Journal of Retailing*, *99*(1), 66–84.

Chikwede, K. (2022). *An Assessment of Tax Revenue Performance Using Tax Capacity and Tax Effort: The Case of Southern African Development Community (2000-2020)*. 1–102.

Choudhury, R. N. (2020). Assessing the trade of digitisable goods: implications for South Asia. *Journal of Economic Studies*, *48*(1), 63–78.

Dennis, L. A., & Science, C. (2022). *Markov chain model representation of*. *32*(6).

Dutta, S., & Lanvin, B. (2021). *The Network Readiness Index 2021 Shaping the Global Recovery How digital technologies can make the post-COVID world more equal*.

Eckelman, M. J., & Daigo, I. (2008). Markov chain modeling of the global technological lifetime of copper. *Ecological Economics*, *67*(2), 265–273.

European Central Bank. (2014). *The assessment of fiscal effort*. *October*, 69–82.

Glenday, G. (2013). Revenue Forecasting. *The International Handbook of Public Financial Management*, *October 1975*. https://doi.org/10.1057/9781137315304.0033

Gomero, G. D. (2022). The factors that influence revenue collection success in the Beneshangul Gumzu region: the case of Assosa town. *Journal of Positive School Psychology*, *6*(2), 5788–5797. http://journalppw.com

Haller, A., Gherasim, O., & Bălan, M. (2020). Medium-term forecast of European economic sustainable growth using Markov chains. *Alina-Petronela Haller et al.• Medium-Term Forecast of European Economic... Zb. Rad. Ekon. Fak. Rij*, *38*(2), 585–618.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis* (pp. 202–209). Springer.

Hanousek, J., & Palda, F. (2022). The evolution of tax evasion in the Czech Republic: a Markov chain analysis. *Journal of Public Economic*, *36*, 57–72. https://www.academia.edu/download/39306391/0912f514208f58addc000000.pdf

Hill, S., Jinjarak, Y., & Park, D. (2022). How do Tax Revenues Respond to GDP Growth? Evidence from Developing Asia, 1998–2020. *Asian Development Bank*.

Hudson, J., Fielding, S., & Ramsay, C. R. (2019). Methodology and reporting characteristics of studies using interrupted time series design in healthcare. *BMC Medical Research Methodology*, *19*(1), 137. https://doi.org/10.1186/s12874-019-0777-x

ITU. (2016). Core list of ICT indicators. *Itu*, *March*, 34. http://www.itu.int/en/ITU-D/Statistics/Pages/coreindicators/default.aspx%0Ahttps://www.itu.int/en/ITU-D/Statistics/Documents/coreindicators/Core-List-of-Indicators_March2016.pdf%0Ahttp://www.itu.int/en/ITU-D/Statistics/Documents/coreindicators/Core-List-o

Junquera-Varela, R. F., & Lucas-Mas, C. Ó. (2024). Digital transformation of tax and customs administrations. *Revenue Administration Handbook*, 129–140. https://doi.org/10.1596/978-1-4648-2053-3_ch6

Kagoya, S. M., Mashaka, M., & Caine, J. (2024). The Application of Markov Chains Model in Determining the Switching Behaviour of Customers in Phone Companies in Developing World: A Study of Tanzania. *Rethinking ICT Adoption Theories in the Developing World: Information and Communication Technologies*, 269–283.

Katz, R. L. (2018). The impact of the broadband internet on employment. *Digitized Labor: The Impact of the Internet on Employment*, 95–108.

Khahro, S. H., Farooq, S., Ali, T. H., Memon, N. A., & Khahro, H. (2020). An Assessment of Research Trends on fiscal policy. *Library Philosophy and Practice*, *2020*, 1–11.

Kim, Y., Kelly, T., & Raja, S. (2010). *Building broadband: Strategies and policies for the developing world*. World Bank Publications.

Kohli, U. (2003). GDP growth accounting: a national income function approach. *Review of Income and Wealth*, *49*(1), 23–34.

Koniagina, M. N. (2020). Forecast of budget revenues from taxes in the context of economy digitalization. *IOP Conference Series: Materials Science and Engineering*, *940*(1). https://doi.org/10.1088/1757-899X/940/1/012040

Kovacs, M. (2018). Market share modelling and forecasting using markov chains in the case of romanian banking institutions. *Quality-Access to Success*, *19*.

Kupelian, B., & Loughbridge, J. (2017). How does uncertainty impact economic activity ? *PWC Global Economy Watch*, *6*, 1–4.

Leung, H., & Fan, Z. (2002). Software cost estimation. In *Handbook of Software Engineering and Knowledge Engineering: Volume II: Emerging Technologies* (pp. 307–324). World Scientific.

Li, Y., Cui, L., & Lin, C. (2017). Modeling and analysis for multi-state systems with discrete-time Markov regime-switching. *Reliability Engineering & System Safety*, *166*, 41–49.

Lowry, S. (2019). Digital services taxes (DSTs): Policy and economic analysis. *CRS Report*, *45532*.

Mallick, H. (2021). Do governance quality and ICT infrastructure influence the tax revenue mobilisation? An empirical analysis for India. *Economic Change and Restructuring*, *54*(2), 371–415.

Mayer, W., Madden, G., & Wu, C. (2020). Broadband and economic growth: a reassessment. *Information Technology for Development*, *26*(1), 128–145.

Mills, L. (2017). Barriers to increasing tax revenue in developing countries. *Knowledge, Evidence and Learning for Development*, 1–15. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/13053/K4D_HDR_Tax capacity barriers .pdf?sequence=126&isAllowed=y

Nikiforova, L. (2022). Use of innovative information technology in e-commerce and digital economy. *Innovation and Sustainability.№ 1: 65–71.*

Nose, M., & Mengistu, A. (2023). Exploring the Adoption of Selected Digital Technologies in Tax Administration: A Cross-Country Perspective. *IMF Notes*, *2023*(008).

Nuha, H., Widiasih, W., & Romadon, M. W. (2022). The implementation of Markov chain to predict market share smartphone customers in Surabaya during pandemic COVID-19. *Voice of Academia (VOA)*, *18*(1), 211–225.

Olaoye, C. O., & Atilola, O. O. (2018). Effect of e-tax payment on revenue generation in Nigeria. *Journal of Accounting, Business and Finance Research*, *4*(2), 56–65.

Ozcan, B. (2018). Information and communications technology (ICT) and international trade: evidence from Turkey. *Eurasian Economic Review*, *8*, 93–113.

Park, H. J., & Choi, S. O. (2019). Digital innovation adoption and its economic impact focused on path analysis at national level. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(3), 56. https://doi.org/10.3390/joitmc5030056

Pushkareva, N. (2021). Taxing times for development: Tax and digital financial services in Sub-Saharan Africa. *Financing for Development*, *1*(3), 33–64.

Qiang, C. Z.-W., Rossotto, C. M., & Kimura, K. (2009). Economic impacts of broadband. *Information and Communications for Development 2009: Extending Reach and Increasing Impact*, *3*, 35–50.

Rahman, H. U., Ali, G., Zaman, U., & Pugnetti, C. (2021). Role of ICT investment and diffusion in the economic growth: A threshold approach for the empirical evidence from Pakistan. *International Journal of Financial Studies*, *9*(1), 14.

Rekova, N., Telnova, H., Kachur, O., Golubkova, I., Baležentis, T., & Streimikiene, D. (2020). Financial sustainability evaluation and forecasting using the Markov chain: The case of the wine business. *Sustainability*, *12*(15), 6150.

Rodriguez, A. C. (2011). Prediction of currency crises using a fiscal sustainability indicator. *Economic Analysis Review*, *26*(2), 38–60.

Ryu, S., Kwak, K., & Lee, S. (2024). The growth patterns of firms in the industrial value chain using the hidden markov model: Evidence from the Korean semiconductor industry. *Computers & Industrial Engineering*, *192*, 110187.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students*. Pearson education.

Sawng, Y., Kim, P., & Park, J. (2021). ICT investment and GDP growth: Causality analysis for the case of Korea. *Telecommunications Policy*, *45*(7), 102157.

Sheng, Y., Zhang, J., Tan, W., Wu, J., Lin, H., Sun, G., & Guo, P. (2021). Application of grey model and neural network in financial revenue forecast. *Computers, Materials & Continua*, *69*(3), 4043–4059.

Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *The Journal of Finance and Data Science*, *4*(2), 71–89. https://doi.org/https://doi.org/10.1016/j.jfds.2017.11.001

Stephens, H. M., Mack, E. A., & Mann, J. (2022). Broadband and entrepreneurship: An empirical assessment of the connection between broadband availability and new business activity across the United States. *Telematics and Informatics*, *74*, 101873.

Stephens, H., Mack, E., & Mann, J. (2022). How the emergence of broadband internet affected new business activity. *Available at SSRN 4046439*.

Streimikiene, D., Rizwan Raheem, A., Vveinhardt, J., Pervaiz Ghauri, S., & Zahid, S. (2018). Forecasting tax revenues using time series techniques–a case of Pakistan. *Economic Research-Ekonomska Istraživanja*, *31*(1), 722–754.

Suominen, K. (2017). Fuelling trade in the digital era: Policy roadmap for developing countries. *Geneva, Switzerland: International Centre for Trade and Sustainable Development (ICTSD)*.

Teltscher, S. (2002). Electronic commerce and development: Fiscal implications of digitized goods trading. *World Development*, *30*(7), 1137–1158.

Thompson-Abbott, A. K., Norton, T. A., & Keener, R. (2023). The Tax Implications of Remote Work. *Journal of state taxation*.

Vu, K., Hanafizadeh, P., & Bohlin, E. (2020). ICT as a driver of economic growth: A survey of the literature and directions for future research. *Telecommunications Policy*, *44*(2), 101922.

World Bank. (2009). *ICT Performance Measures: Methodology and Findings*. 133–143.

World Bank, W. B. (2018). *Data for development;An Evaluation of World Bank Support for Data and Statistical Capacity*. 6–7. https://doi.org/10.18356/26ee3f2b-en

Yan, Q., Qin, C., Nie, M., & Yang, L. (2018). Forecasting the electricity demand and market shares in retail electricity market based on system dynamics and Markov chain. *Mathematical Problems in Engineering*, *2018*(1), 4671850.

Zatonatska, T., Klapkiv, Y., Dluhopolskyi, O., & Fedirko, O. (2022). Forecasting of the Employment Rate in the EU ICT Field. *Comparative Economic Research. Central and Eastern Europe*, *25*(3), 7–25.

Zhang, J., Zhao, W., Cheng, B., Li, A., Wang, Y., Yang, N., & Tian, Y. (2022). The Impact of Digital Economy on the Economic Growth and the Development Strategies in the post-COVID-19 Era: Evidence From Countries Along the "Belt and Road". *Frontiers in Public Health*, *10*(May), 1–17. https://doi.org/10.3389/fpubh.2022.856142

Zhang, X., & Chen, N. (2023). Does financial institutions assure financial support in a digital economy for energy transition? Empirical evidences from Markov chain and DEA technique. *Environmental Science and Pollution Research*, *30*(23), 63825–63838.

Zheng, K., Li, Y., & Xu, W. (2021). Regime switching model estimation: spectral clustering hidden Markov model. *Annals of Operations Research*, *303*, 297–319.

Zohrah, B. S. P., Bahri, S., & Baskara, Z. W. (2024). Forecasting Non-Metal and Rock Mineral (MBLB) Tax Revenue Using the Fuzzy Time Series Markov Chain Method in East Lombok Regency. *Eigen Mathematics Journal*, *7*(1), 104–111. https://doi.org/10.29303/emj.v7i1.171