Comparative Analysis of The Impact of Big Data on Corruption in Selected Developing and Developed Countries

Adewale Joel Adebisi and Cherif Guermat

ABSTRACT

Big data analytics adoption is gaining momentum amongst both policy makers and business leaders. Indeed many sectors in the economy and society have been adopting recent information technology innovations to deal with issues that were previously impossible to tackle. One of these issues is corruption. This paper considers the impact of big data on corruption in developed and developing countries. Specifically, we investigate the effect of internet usage on corruption prevention and early detection, and examine the impact of investment in data-driven technology on corruption prevention and early detection. Finally, we evaluate the impact of mobile data subscription on total corruption. We use secondary data covering 1995 to 2020 for three low FinTech developing countries and three mature FinTech developed countries. Random effect regression models are employed to estimate and test for the impact of big data on corruption in developing and developed countries. We find that internet usage and fixed telephone subscription have a significant negative impact on corruption in developing countries. Investment in technology, mobile phone users subscription and gross domestic product also have a significant negative impact on corruption in developing countries. However, inflation has no significant effect on corruption in developing countries. In contrast, we find no significant impact of big data on corruption within developed countries. Big data adoption, therefore, seems to hinder corruption in developing countries, but not in developed countries.

Keywords: Big data, corruption, internet subscription, phone subscription, technology.

I. INTRODUCTION

Corruption is a global challenge. Although many countries have adopted different strategies to curtail the spread of corruption, it is still a significant problem. This is in contrast to developed countries, which have been able to curb the level of corruption within their borders to reasonable levels using institutional, technological and legal means. Not surprisingly, though, corruption has not been eliminated completely from their system.

Many developing countries have adopted anti-corruption campaigns and anti-corruption agencies to fight corruption within their systems, all to no avail. Corruption in developing countries has been perceived fastest means to wealth, while weak legal systems in developing countries are not helping either [1]. Similarly, [2] concluded that corruption in developing countries recently showed inefficiency in the existing corruption-fighting mechanism.

Given the foregoing, the economic growth and development of many nations worldwide face challenges due to corruption within their economic space. An IMF report [3] concluded a significant negative relationship between corruption and government expenditure on education and economic growth. However, many developing countries face intelligence-gathering problems in fighting corruption. Therefore, corruption must be reduced to boost economic growth and development in affected countries.

Recent developments in big data and the data-driven field have received much attention due to the potential benefits of real-time discovery of fraud and errors. Data analytics, machine learning, and artificial intelligence may also assist computers in detecting transactions out of the pattern. Big data adoption is essential considering the rate of increase in corruption and the failure of the existing mechanism to fight corruption acceptably. Some authors argued that "there is an increased opportunity to develop practical anti-Corruption tools based on big data, open data and artificial intelligence” [4]. Therefore, big data is likely to be beneficial to developing countries in fighting corruption.

Consequently, this study argues that the big data aspect of financial technology can reduce corruption drastically. Big data potential to reduce corruption may be related to monitoring data generated from mobile phones, the internet, GPRS network, financial transaction tracking, and social data...
analyses to draw insight into people’s way of life. This monitoring aims to eliminate corruption using big data and financial technology by leveraging process automation. Reference [5] utilized the web scrubbing method to analyze 370,000 online comments about China’s anti-corruption campaign between 2012 and 2015 and argued that the social media anti-corruption campaign had a more significant effect on anti-corruption awareness than the conventional approach. Thus, this study focuses on fighting corruption with big data generated from technological adoption.

The remainder of this paper is organized as follows. Section 2 discusses the literature review and hypotheses development. Section 3 describes the methodology, while section 4 covers the results and discussion. Finally, section 5 concludes.

II. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

A. Big Data

Big data is defined as the vast and diverse stream of the dataset with unlimited scope. It comprises a virtually unlimited volume of information, velocity, veracity, and variety [6]. Big data is an emerging aspect of financial technology. Its emergence is mostly due to the large amount of available data, including structured, semi-structured and unstructured data generated from the transactions through the internet of things, and processed across mobile phones, tablets, wireless, robotic solutions, and social media. Big data is the output of electronic activities, and could be in any form, including value, images, video, and text. The role of big data is to utilize vast data generated through electronic transactions to aid decision making in every facet of life while using data mining for customer risk profiling that will help in quick fraud and money laundering discovery [7].

Big data still targets every part of human activity to generate timely and reliable information in a more effective and sophisticated form [8]. However, although big data has been acknowledged to have many benefits in all aspects of life, its adoption by developing countries is still limited due to the low level of technology diffusion within these countries.

There is empirical evidence that big data could be important to organizations and governments alike. Reference [9] identified that big data analytics adoption would help improve accuracy. Similarly, [10] adopted a prescriptive analytics method to examine big data learning analytics in higher education, to spot students at risk and provide possible interventions to help out. The study agreed that big data facilitate researchers’ real-time analysis of learning activities. Likewise, studies have discovered that neglecting the importance of big data in the educational decision-making process accounted for the failure to deliver developmental goals in Nigeria [11].

That aside, the investigation into the adoption and impact of big data in the public sector revealed that big data can address many public sector problems, even when some tension still exists between the promise of big data and reality [12]. Also, [13] believed that big data serve as the solution to security and environmental issues through analysis of big data to trace the activities of the citizens. In another perspective, [14] posited that big data increases experts' involvement in government performance appraisals. The study also asserted that citizens’ and policymakers' accessibility to performance information in the public domain would facilitate public administrators' performance and improve responsibility and efficiency. A further study by [15] concluded that big-data adoption by the government sector comes at a huge cost required by investment in new technology and human skill upgrade. Yet, despite the the cost, big data remains a valuable tool to capture security and corruption challenges and then aid economic development.

On the other hand, [16] argued that technology might facilitate transaction and activity monitoring, utilizing big data to expose corruption across entities for prosecution. Yet big data could cause false alarms and increase the cost of investigating corruption. Likewise, [17] found that big data had not succeeded in combating money laundering in the Netherlands. This failure is mostly caused by spurious results produced by innovative ideas to mine large datasets.

Undoubtedly, e-government will disrupt corruption, but to what extent may be difficult to measure without big data analytics. Moreover, technology provides a gateway for big data, and technology can integrate into every economic sector. Hence, as identified by the [42] big data can be used to fight corruption and build a culture of integrity as practiced by most developed countries. This is depicted in Fig. 1.

In the same vein, [18] concluded that open government and big data would reduce corruption in a country with an effective legal system. However, this study used aggregated measures for the panel study without considering individual country characteristics. Similarly, [12] adopted mixed data sources to conclude that big data faced the problem of data governance and data privacy. Still, it is vital to tackle public sector corruption and understand the causal relationship between developing policies and programs.

![Fig. 1. Partnering against corruption.](image_url)

B. Corruption

Corruption is a global challenge, although the developed countries curbed it to a very reasonable extent using institutions, technology and the legal system. Many studies have concluded that corruption is prevalent in developing countries to the extent that it is more lucrative than any other
type of venture [19]. Therefore, corruption increases administration costs, which has terminated most corporations globally. Reference [20] concluded that corruption per worker negatively impacted output per worker directly and indirectly on foreign private investment, expenditure on education, and capital expenditure per worker. According to the World Bank data, corruption can follow any form described in Table I.

<table>
<thead>
<tr>
<th>Categories of corruption</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bribery</td>
<td>The act of dishonesty persuading someone to act in one’s favor by a payment or other inducement. Inducements can take the form of gifts, loans, fees, rewards or other advantages (taxes, services, donations, etc.). The use of bribes can lead to collusion (e.g., inspectors under-reporting offences in exchange for bribes) and/or extortion (e.g. bribes extracted against the threat of over-reporting). To steal, misdirect or misappropriate funds or assets placed in one’s trust or under one’s control. From a legal point of view, embezzlement need not necessarily be or involve corruption.</td>
</tr>
<tr>
<td>Embezzlement</td>
<td>A small payment, also called a “speed” or “grease” payment, made to secure or expedite the performance of a routine or necessary action to which the payer has legal or other entitlement.</td>
</tr>
<tr>
<td>Fraud</td>
<td>The act of intentionally and dishonestly deceiving someone in order to gain an unfair or illegal advantage (financial, political or otherwise)</td>
</tr>
<tr>
<td>Collusion</td>
<td>An arrangement between two or more parties designed to achieve an improper purpose, including influencing improperly the actions of another party. The act of impairing or harming, or threatening to impair harm, directly or indirectly, any party or the property of the party to influence improperly the actions of a party.</td>
</tr>
<tr>
<td>Extortion</td>
<td>Patronage at its core means the support given by a patron. In government, it refers to the practice of appointing people directly.</td>
</tr>
<tr>
<td>Patronage, clientelism</td>
<td></td>
</tr>
<tr>
<td>and nepotism</td>
<td></td>
</tr>
</tbody>
</table>

There has been a consensus in the literature that, notwithstanding the level of corruption, big data analytics would help reduce corruption, provided process automation is adopted [21], [22]. Contrarily, [23] maintained that big data analytics only explain what is happening rather than why. Therefore, policy makers may be exposed to public criticism when such insights are used to justify policy decisions on corruption.

However, some of the previous studies argued that political will, low population, impartial prosecution of corrupt officers and effectiveness of data gathering by the anti-corruption agencies are more important in fighting corruption in many countries [24]-[36]. These efforts vary from one country to another. In a recent study on Asia and Africa, [19] opined that corruption reduces the tendency to adopt research, innovation, patents and development among lower-middle-income economies. Similarly, a qualitative study on strategies for combating corruption in Nigeria and Iraq established that Nigeria’s corruption-fighting institutional strategies are more effective than Iraq’s [27].

C. Internet as A Tool to Combat Corruption

The use of the internet has become one of the main platforms for social interaction among people. Therefore, it could help fight corruption when used to track corrupt offenders. A panel study by [28] posited that internet adoption has a significant negative effect on corruption but is not too substantial. In the same vein, a cross-sectional survey of 157 countries by [29] concluded that the internet and social media networking has a strong and consistent negative relationship with corruption. Thus, the study supported that internet users’ expansion will extend news to promote public officials’ accountability and reduce corruption.

In fact, the Chinese autocrats used internet data capability to suppress opposition leaders through social media propaganda. This propaganda facilitated citizens’ dominance and persuaded them to avert repression’s expensive cost [30]. Based on China’s experience, big data can be employed to solve political and social-economic conditions in other countries. Further, a study on the European Union, [27], identified that internet usage and enabling internet access were essential to promote anti-corruption effort.

Similarly, [31] concluded in a comparative study between the United States and other countries that the emergence of the internet has helped reduce the magnitude of corruption across U.S. states and the rest of the world. The above findings show that internet data can help curb corruption, through the exploitation of the data extracted from both the internet and other technology devices. A study on implementing international anti-corruption initiatives identified that parties involved in cross-border transactions are prone to money laundering. Therefore, big data should be judiciously employed to monitor such activities [32]. Similarly, combating corruption in Vietnam confirmed the failure of reinforcement, cooperation and official structure method of anti-corruption(conventional strategy), which means internet data might be an alternative [33].

However, [34] opined that corruption would continue in Nigeria until capital punishment is the sanction for corruption. The government can reduce corruption in public sector services by employing a big data analytics model to track and analyze their responsible officers’ online activities for prosecution. A similar study on Malaysia argued that e-procurement system implementation in the public sector would reduce corruption; and that internet data is an anti-corruption tool [35]. Another study on combating corruption in a microfinance institution revealed that robust data monitoring for the decision-making process would reduce corruption [36]. Thus, the following hypothesis is formulated in line with the foregoing:

H1: Greater internet usage helps the prevention and early detection of corruption.

D. Mobile Cellular Subscription

Due to the importance of communication between individuals, the mobile cellular phone has become part of our daily lives. Therefore, phone data obtained from communications could be used to track any form of corruption. Reference [37] argued that mobile technologies supported the anti-corruption fight globally in many places. The study also claimed that mobile technologies might be used for social accountability to report corruption committed in public service delivery and demand accountability from leaders. In the same vein, a panel study on twelve Asia-Pacific countries concluded that mobile phone penetration and internet adoption had a significant negative impact on
corruption [38]. In contrast, a study from East Africa concludes that earlier optimism around mobiles’ potential to support citizens’ anti-corruption crusade has not resulted in a significant body of research. The literature provides no substantive clues as to why this urgent topic has not been explored fully [39]. Therefore, the following hypothesis was formulated based on the above:

H2: Mobile data subscription is does not have association with corruption.

III. METHODOLOGY

This section discussed the model used in the study.

A. Methodology Specification

The basic model formulated for this study is stated below:

\[ CPI_{it} = \beta_0 + \beta_1 IIT_{it} + \beta_2 MPUS_{it} + \beta_3 FTS_{it} + \beta_4 IUIP_{it} + \beta_5 GDP_{it} + \beta_6 INF_{it} + \delta_i D_{it} + \epsilon_{it} \]  

(1)

However, given the potential difference between developing and developed countries, we extend the model by considering a dummy variable, \( D_{it} \), representing developed economies. Therefore, the full model reflects the effect of change in the slope:

\[ CPI_{it} = \beta_0 + D_{it} \delta_0 + \beta_1 IIT_{it} + \beta_2 MPUS_{it} + \beta_3 FTS_{it} + \beta_4 IUIP_{it} + \beta_5 GDP_{it} + \beta_6 INF_{it} + \delta_i D_{it} + \epsilon_{it} \]  

(2)

or

\[ CPI_{it} = \beta_0 + \beta_1 IIT_{it} + \beta_2 MPUS_{it} + \beta_3 FTS_{it} + \beta_4 IUIP_{it} + \beta_5 GDP_{it} + \beta_6 INF_{it} + \delta_i D_{it} + \epsilon_{it} \]  

(3)

The impact of independent variable \( k \) for developing countries is \( \beta_k \), whereas the impact of independent variables for developed countries is \( \beta_k + \delta_k \).

The variables are defined as follows:

- \( CPI_{it} \): Corruption perception index
- \( IIT_{it} \): Investment in technology
- \( MPUS_{it} \): Mobile phone user subscription
- \( FTS_{it} \): Fixed telephone subscriptions
- \( IUIP_{it} \): Individuals using the internet (% of the population)
- \( GDP_{it} \): Gross Domestic Product
- \( INF_{it} \): Yearly inflation
- \( D_{it} \): Dummy Variable for representing developed countries
- \( \alpha \): Constant
- \( \epsilon \): Error term

\( \beta_1 \) to \( \beta_6 \) and \( \delta_0 \) are the parameters used for measuring the impact of each independent variable used as proxies for big data on the change in the corruption coefficients as depicted in the regression models.

IV. DATA ANALYSIS AND RESULTS

A. Descriptive Statistics Summary

We use the corruption perception index to measure the countries’ corruption. The control variables are inflation, gross domestic product, and population. The independent variables are mobile phone user subscription, fixed telephone subscriptions, investment in technology, individuals using the internet, and scientific and technical journal articles to proxy for big data. Table II shows the mean, standard deviation, minimum and maximum values for each observed variable used in the study.

The results presented in Table II show that our sample contain 156 observations from six countries whose corruption perception ranking varied from 7th to 154th position globally. The average mobile phone user subscription was 69.40 subscription for every 100 people, and the standard deviation was 46.56 for the sampled countries. Also, the minimum mobile phone user subscription per 100 people was 0.01, while the maximum mobile phone user subscription per 100 people had 165.6. Similarly, fixed telephone subscriptions had an average of 33.7 million units for the sampled countries, with a standard deviation of 56.3 millions during the study period. The minimum fixed telephone subscription was 65.644 thousands, while the maximum fixed telephone subscription was 193 millions for the sampled countries.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPI</td>
<td>156</td>
<td>55.686</td>
<td>52.255</td>
<td>3.00</td>
<td>154.000</td>
</tr>
<tr>
<td>MPUS</td>
<td>156</td>
<td>69.404</td>
<td>46.564</td>
<td>0.010</td>
<td>165.600</td>
</tr>
<tr>
<td>FTS (millions)</td>
<td>156</td>
<td>33.700</td>
<td>56.300</td>
<td>0.065</td>
<td>193.000</td>
</tr>
<tr>
<td>IUIP</td>
<td>156</td>
<td>37.987</td>
<td>33.215</td>
<td>0</td>
<td>94.780</td>
</tr>
<tr>
<td>GDP</td>
<td>156</td>
<td>9.206</td>
<td>4.196</td>
<td>0.940</td>
<td>18.670</td>
</tr>
<tr>
<td>INF</td>
<td>156</td>
<td>1.383</td>
<td>2.504</td>
<td>-10.300</td>
<td>12.460</td>
</tr>
</tbody>
</table>
| CPI: Corruption perception index; MPUS: mobile phone user subscription; FTS: Fixed telephone subscriptions; IUIP: Investment in Technology; GDP: Gross Domestic Product per market value; INF: Yearly inflation.

However, the individuals using the internet averaged 37.98 per cent, with a standard deviation of 33.21 per cent for the sampled countries. There is one country where internet usage by individuals is nil. The highest individual using the internet reported 94.78 per cent for the sampled countries. The result depicted the investment in technology had an average of 9.20 per cent, with a standard deviation of 4.19 per cent. Likewise, the minimum and maximum investment in technology stand at 0.94 and 18.67 per cent, respectively.

The GDP growth rate averaged during the study period was about 1.38 per cent, with an average spread of 2.50 per cent for the period under review. Also, the minimum GDP growth rate was -10.30 per cent, while the highest GDP growth rate was about 12.46 per cent for the sampled countries. The result showed that the inflation rate averaged 5.88% per cent; the standard deviation was about 7.24% for the sampled countries. The minimum inflation rate was -0.69% for the sampled period. However, the maximum inflation rate was 72.84% for the sampled countries.

B. Trend Analysis

The data shows that Nigeria and Kenya are leading on the
corruption index among the selected samples, with South Africa in the midway. In contrast, the developed countries showed a low level of corruption index. South Africa is leading mobile phone subscription from 2010 to 2020, while mobile phone subscription in Nigeria is at the lowest due to the over-saturation of the Nigerian market. The United States had the highest phone subscription, followed by the United Kingdom, while the use of fixed telephone subscriptions is very low in developing countries. The number of individuals using the internet in developed countries is more than in developing countries, implying a high generation of big data from developed countries. The gross domestic products for the sampled countries dropped in 2020, which is the effect of the COVID-19 pandemic on the global economy. The investment in technology is topped by the United States, followed by Australia, although other countries change their spending in 2020 as indicated in the graph.

Graphical illustration of the trend for the variables:
C. Regression Analysis

The correlation analysis test examined the effect of multicollinearity among explanatory variables used in the study. The pairwise correlations are presented in Table V and show that none of the independent variables had multicollinearity issues since the highest correlation results among the explanatory variables were 0.7. The result shows that none of the independent variables correlates with greater than 0.8 in absolute value.

The random-effects model results presented in Table III show big data’s effect on combating corruption in developing countries. The random-effects model shows that individuals using the internet, investment in technology, mobile phone user subscription, inflation, fixed telephone subscription, and gross domestic product growth are statistically significant at 1% significance levels. This significance implies that these variables have a statistically significant impact on corruption in developing countries. On the other hand, inflation does not significantly impact corruption. This insignificant means that these variables do not determine the corruption level of the developing countries.

Few variables, including fixed telephone subscription, and individual user internet phone, have negative coefficients, which means they negatively impact the corruption index. While inflation, mobile phone user subscription, investment in technology, and gross domestic product growth positively impact the level of corruption in developing countries.

The negative coefficient of individuals using the internet indicates that a percentage increase in the individual using the internet will reduce corruption by about 0.83, and vice versa. Also, the negative coefficient of fixed telephone subscription implies that an additional 1000 subscription in fixed telephone leads to -0.02 decrease in corruption perception for developing countries. Contrarily, a percentage increase in investment in technology will lead to an increase in corruption by 3.3218, and vice versa. Also, a per cent increase in the gross domestic product will result in an increase in corruption by 3.06577. At the same time, a per cent increase in inflation will lead to an increase in corruption by 0.129 per cent. The positive coefficient in mobile phone user subscription implies that a per cent increase in mobile phone user subscription will increase corruption by 0.565, and vice versa.

The Wald Chi-squared statistic presented for the random-effects model shows a value of 1420.92 and a p-value of 0.000, which indicates the statistical significance of the model. This implies that the overall model is statistically significant at the 1 per cent level of significance. This significance is also suggestive that the model is a good fit to predict the change in the corruption perception index. The R-squared shows a value of 0.909, indicating that 90.9 per cent of variations in the corruption is explained by the model.

D. Economic Interpretation of the Results

Based on the result in Table III, big data adoption is required to put corruption levels in check and boost economic growth and development in developing countries. There is a need to promote individuals using the internet and fixed telephone subscribers to reduce corruption. Similarly, investment in technology, mobile phone user subscription and gross domestic product require policies to ensure they are used to reduce corruption. Policymakers are expected to implement policies and sensititize the anti-corruption agencies on the need to incorporate data extractions and analytics into their game plan for appropriate technology integration. This policy will assist every sector, including RegTech, FinTech, InsureTech and AgroTech to operate in a well-planned environment. Indeed, it is crucial to effectively improve big data adoption to curb corruption and aid economic development in developing countries.

Table III: Regression Results of Impact of Big Data on Combating Corruption in Developing Countries

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coef.</th>
<th>Std. Err.</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIT</td>
<td>3.32185</td>
<td>1.245</td>
<td>0.008</td>
</tr>
<tr>
<td>MPUS</td>
<td>0.55601</td>
<td>0.095</td>
<td>0.000</td>
</tr>
<tr>
<td>FTS</td>
<td>-0.00002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>IUUP</td>
<td>-0.83070</td>
<td>0.302</td>
<td>0.006</td>
</tr>
<tr>
<td>GDP</td>
<td>3.71894</td>
<td>0.696</td>
<td>0.000</td>
</tr>
<tr>
<td>INFL</td>
<td>0.12935</td>
<td>0.242</td>
<td>0.593</td>
</tr>
<tr>
<td>D=IIT</td>
<td>-3.06577</td>
<td>1.804</td>
<td>0.008</td>
</tr>
<tr>
<td>D=MPUS</td>
<td>-0.50304</td>
<td>0.191</td>
<td>0.008</td>
</tr>
<tr>
<td>D=FTS</td>
<td>0.00002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>D=IUUP</td>
<td>0.81476</td>
<td>0.398</td>
<td>0.041</td>
</tr>
<tr>
<td>D=GDPS</td>
<td>-3.90680</td>
<td>1.291</td>
<td>0.002</td>
</tr>
<tr>
<td>D=INFL</td>
<td>0.05613</td>
<td>2.007</td>
<td>0.978</td>
</tr>
<tr>
<td>Intercept</td>
<td>90.95898</td>
<td>7.353</td>
<td>0.000</td>
</tr>
<tr>
<td>No. of Panel id</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F/Wald Chi2</td>
<td>1420.92***</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.909</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Dependent variable: CPI.

Table IV: Results of Impact of Big Data on Combating Corruption in Developed Countries

<table>
<thead>
<tr>
<th>Variables</th>
<th>Chi 2(2)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment in Technology</td>
<td>0.04</td>
<td>0.845</td>
</tr>
<tr>
<td>Mobile Phone user</td>
<td>0.10</td>
<td>0.749</td>
</tr>
<tr>
<td>Subscription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Telephone</td>
<td>1.16</td>
<td>0.282</td>
</tr>
<tr>
<td>subscription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals using</td>
<td>0.00</td>
<td>0.951</td>
</tr>
<tr>
<td>the internet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Domestic Product</td>
<td>0.03</td>
<td>0.863</td>
</tr>
<tr>
<td>Inflation</td>
<td>0.01</td>
<td>0.926</td>
</tr>
</tbody>
</table>
CPI\_it = \beta_0 + [\beta_1 I\!T\!T\_it + \beta_2 MPUS\_it + \beta_3 FTS\_it + \beta_4 IUIP\_it + \\
\beta_5 GDP\_it + \beta_6 INFL\_it] + [\delta_1 D\!_1 I\!T\!T\_it + \delta_2 D\!_2 MPUS\_it + \\
\delta_3 D\!_3 FTS\_it + \delta_4 D\!_4 IUIP\_it + \delta_5 D\!_5 GDP\_it + \\
\delta_6 D\!_6 INFL\_it]

E. Result of the Impact of Big Data on Corruption in Developed Countries

The Wald test results presented in Table IV show the effect of big data on combating corruption in developed countries. For example, testing the investment in technology in developed countries involves testing the hypothesis that \( \beta_1 + \delta_1 = 0 \) in equation 2.

The Wald test shows that individuals using the internet, mobile phone user subscription, inflation, fixed telephone subscription, investment in technology, and gross domestic product do not significantly impact corruption in developed countries. The insignificance implies that these variables do not determine the corruption level of the developed countries. The p-values of all the variables are all substantially greater than 10%.

F. Economic Interpretation of The Results

Table IV shows the results of testing the hypotheses that the impact of independent variable k on the CPI of developed countries is insignificant. In other words, the hypothesis \( \beta_k + \delta_k = 0 \) is tested using the Wald test.

The results in Table IV mean that big data does not affect corruption in developed countries. This absence of effect is possibly due to big data being optimally used to reduce corruption in these countries. Therefore, policymakers in the developed economies should focus more on moral persuasion or public enlightenment to eliminate the little corruption left in their countries. The results also suggest that developed countries have already adopted big data to wipe out corruption in their countries. This is possible through the digitalization of most activities, which closes the net on most corruption activities. This is perhaps good news for developing countries as developed countries may help them fight corruption through training and intelligence gathering and sharing.

To sum up, the results confirm that individuals using the internet, investment in technology, Gross Domestic Product, fixed telephone subscription and mobile phone user subscription are significant factors that affect corruption in developing countries. The situation is different in developed countries where none of the variables affects corruption. The result shows that corruption in developed countries has reached a minimal level, so big data may maintain it at low level but cannot reduce it further. Therefore, policymakers of developing countries must consider adopting big data to mitigate corruption.

V. Discussion

A. Hypothesis One

Hypothesis one states that internet usage affects corruption prevention and early detection in the selected countries. The hypothesis was tested using random effect panel regression methods. The regression results are presented in Table III, and the Wald test results are presented in Table IV. The results indicate a statistically significant impact of the number of individuals using the internet on corruption in developing countries (p-value<0.01) for developing countries. The significance implies the acceptance of the first study hypothesis, namely that internet usage affects corruption prevention and early detection in developing countries. In contrast, in the case of developed countries, the p-value equals 0.951, way above the 5% or 1% threshold for significance. This insignificance implies the rejection of the first hypothesis for developed countries. This finding suggests that internet usage significantly impacts corruption in developing, while it has no impact on corruption in developed countries.

B. Hypothesis Two

This hypothesis states that mobile data subscriptions have a negative relationship with corruption level in the selected countries. The assumption was tested using the random effect regression method and Wald test. The analysis results are presented in Table III and Table IV. The results of developing countries indicate a statistically significant but positive impact of mobile phone users on corruption (p-value<0.01). The significance implies the rejection of the second hypothesis that mobile data subscriptions do not have relationship with total corruption in the selected developing countries.

On the other hand, the result for developed countries showed an insignificant effect on corruption with p-value of 0.748. This insignificance also implies accepting the second hypothesis that mobile data subscriptions does not have relationship with corruption level in developed countries.
C. Discussion of Findings

Big data deserves greater attention globally. Data nowadays has reached all corners of business and social interactions. With appropriate analytical methods and optimal use of big data for decision-making the potential is limitless. The technological penetration into daily human endeavors across the globe means that big data are generated from every digital contact point globally. The impact of big data on many variables was considered in the studies of many researchers, including; [7]-[17]. However, only a few of these studies examined the impact of big data on corruption, while the concerned studies are limited in scope and had contradictory conclusions. Most previous studies failed to examine how big data can solve social and economic issues like corruption in comparative studies.

D. The Effect of Internet Usage on Corruption Prevention and Early Detection

As revealed in the regression result in Table III and Table IV, greater numbers of internet users significantly reduces the level of corruption in developing countries. The effect, however, is statistically insignificant in developed countries. The results imply that internet usage reduces corruption in developing countries, while internet usage does not affect corruption in developed countries. Hence, this finding on developing countries agrees with the study of [25]-[30], [40] as the studies concluded the significant negative relationship between internet usage and corruption. However, the results on developed countries indicate no impact of internet usage on corruption. Corruption levels in developed countries are generally low, and thus big data differential cannot explain variation in corruption levels.

Consequently, the adoption of internet for government activities and transactions should reduce corruption levels in the public sector of developing countries.

E. The Impact of Mobile Data Subscription on Total Corruption in Selected Countries

Based on the results in Tables III and Table IV, the study discovered that the impact of mobile phone subscription on corruption was significantly positive for developing countries. This result implies that mobile data subscription increases corruption in developing countries. This increase is because policies like whistleblowing and freedom of information to expose perpetrated corruption in developing countries were not adequately investigated. This finding agrees with the study of [22, 40] that mobile phone usage has a significant positive effect on corruption in developing countries. On the other hand, mobile phone subscriptions in developed countries had insignificant positive effects on corruption, which means mobile phone subscriptions do not affect the corruption of developed countries, which is at a minimal level already.

VI. Conclusion

This study provides evidence that big data can be employed to reduce corruption in developing countries, while corruption in developed countries seems to be at a floor and cannot be reduced further with big data technology. Nevertheless, developing countries need to activate big data implementation in line with investment in technology to reduce corruption. It has been shown that developing countries with greater internet usage have lower levels of corruption. In contrast, big data does not seem to have further role in reducing the (already lower) levels of corruption in developed countries. Their early technology adoption across all sectors seems to have already born fruit. Therefore, the result showed that big data could not further reduce corruption in developed countries.

This study agrees with the investigations of [12, 13, 17, 21-23], [41] that big data could be used to solve many economic problems. This study implies that policymakers in developing countries interested in fighting corruption must implement policies that take technological adoption in all sectors seriously. Consequently, this will assist them in maximizing the potential of big data in solving crime and corruption faced in developing countries.

The result of this study is of interest to the developing countries to utilize data generated from internet usage to reduce corruption through big data analytics and intelligence gathering

The developing countries are advised to incorporate big data planning into the investment in technology to optimize the benefit of big data and reduce corruption levels. However, developing countries do not seem to be exploiting the potential information available across mobile networks. This is why policymakers in developing countries should pay attention to big data generated by mobile phone users for possible tracking of corruption and crime cases. Strangely, the use of fixed telephone subscriptions for homes and offices does negatively impact corruption levels in developing countries. Perhaps, developing countries have appropriate technologies to tap into the fixed telephone network but not the mobile networks. On the other hand, it could simply be the fact that the fixed telephone subscription is linked with internet usage, such that the fixed telephone subscription is a proxy for internet use. We leave the above two issues for future studies.

Overall, the developing countries should learn from the experience of developed economies and encourage training in information technology and big data analytics.

REFERENCES


Soudijn MR, de Been WH. Law enforcement and money laundering: Big data is coming. *Criminal Defiance in Europe and Beyond*. 2020: 399-426.


