MODELLING NIGERIA'S PRESIDENTIAL ELECTION DATA USING BENFORD'S LAW AND MONTE CARLO SIMULATION

BY

SUNDAY OLAPADE TUNMIBI

Matric. No.: 147161

B. Tech. (LAUTECH), M. Inf. Sc. (Ibadan)

A Thesis in the Department of Data and Information Science,

Submitted to the Faculty of Multidisciplinary Studies

in partial fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

of the

UNIVERSITY OF IBADAN

January 2021

ABSTRACT

Allegations of fraudulent practices and bogus results led to application of election forensic tools in the analysis of election data. Previous studies examined the digital distributional patterns of electoral data using Zipfian and agent-based modelling, while neglecting sensitivity check that could reveal other anomalies. This study, therefore, was designed to analyse Nigeria's presidential election data between 2007 and 2015 by applying Benford's Law and Monte Carlo Simulation which can indicate voter distribution and reveal any anomalies in the election results, with a view to assessingthe integrity of the election process and results.

Benford's Law and Monte Carlo Simulation models were used as framework, while Modelling and Simulation, which compares the observed patterns against the expected patterns, were adopted as design. Purposive sampling was used to select 2007, 2011 and 2015 presidential election results.Data wasobtained from the website of Independent National Electoral Commission. Political parties included in the analyses were those with at least four digits vote counts: 24 parties for 2007 election; PDP, CPC, ACN and ANPP for 2011 election; and APC and PDP for 2015 election. Also included were voters' turnout for 2011 and 2015 elections (data was not available for 2007). Data were analysed using descriptive statistics and Spearman rank correlation test at 0.05 level of significance, while R programming was used for the Monte Carlo Simulation.

Whereas the 2007 election result contains only vote counts of the 24 political parties, collated at national level only, the 2011 and 2015 election results contain voters' turnout and vote counts for each political party per state.The distribution of last digits of vote counts of 2007, 2011 and 2015 elections and voters' turnouts of 2011 and 2015 elections did not follow the expected uniform distribution of last digits for fraud-free data. The distributional pattern of vote counts for 2011 and 2015 elections deviated from distributional pattern of Monte Carlo simulated vote counts. The first digits of vote counts in 2007 elections of the 24 political parties ($r=0.68$); in 2011 elections of ACN ($r=0.96$), PDP ($r=0.93$), CPC ($r=0.75$) and ANPP ($r=0.73$); and in 2015 elections of APC ($r=0.96$) and PDP ($r=0.74$) significantly correlate with Benford's Law. The occurrence of first digits in voters turnouts of 2011 ($r=0.07$) and 2015 ($r=0.37$) elections did not follow Benford's Law. The occurrence of second digits in vote counts of the 2007 elections ($r=0.36$), 2011 elections [PDP ($r=0.51$), ACN ($r=0.51$), CPC ($r=0.20$) and ANPP ($r=0.17$)] and 2015 elections [APC ($r=-0.61$) and PDP ($r=0.02$)] did not follow Benford's Law. This was also the case in voters' turnout of 2011 ($r=-0.17$) and 2015 ($r=-0.09$) elections.

The application of Benford's Law and Monte Carlo Simulation on Nigerian presidential election data of selected years reveals that the election results are not error-free. Nigeria's electoral process should apply these forensic analyses on electoral data and adjust the electoral process in line with findings.

Keywords: Benford's Law, Election forensics, Monte Carlo simulation, Vote counts, Voter turnout

Word Count: 475

ACKNOWLEDGEMENTS

I thank Almighty God for making this thesis a success. He has been faithful to me all through my academic pursuit. I could not have gone this far without His love, mercy and grace. His name is praised forever.

I am forever indebted to my supervisor Prof. W.M. Olatokun for his love, support and guidance all through the writing of this thesis. I will always cherish his counsel and fatherly advice. My appreciation also goes to Prof. M.A. Tiamiyu, a father and a great academic. I thank him for the knowledge he has imparted. I really appreciate Dr. T. A. Adegbola for making this thesis a reality. I thank him for his mentorship.I alsogreatly appreciateDr. W. E. Nwagwu forthe support, assistance and supervision.My unreserved appreciation to Dr. Funmilola O. Omotayo, Dr. Kemi Ogunsola, Dr. A.O. Opesade, Dr. J.O. Adekannbi, Mrs. F. Longe, Mr. Tayo Bamidele, Mrs. B.A. Oparachukwu, Mrs. Pauline Normeh, Mr. Samuel Aliu, Mrs. K. A.Raji and allother staffmembers of the Department of Data and InformationScience for their support and help at various points of need.

My sincere gratitude goes to Prof. A. A. B. Agbaje, Prof. E. T. Jolayemi, Dr. Adedayo A. Adepoju, Prof. Victoria Adetunji and Prof. N. M. Christopherfor theirgreat support. I also appreciateProf. I. M. Mabawonku, Prof. L. O. Aina, Prof. E. R. Aiyede, Prof. S. O. Popoola, Dr. Olusola G. Adedoja, Dr. C.O.O. Kolawole, Dr. A. O. Egunjobi, Dr. I. A. Raji, Dr. A. O. Ayeni, Dr. J. F. Ojo, Dr. S. O. Akinola, Dr. B. R. Ojebuyi, Dr. A. K. Taiwo, Dr. J. O. Adeleke, Dr. O. J. Taiwo and Dr. A. Tella for their assistance and contributions. Special thanks to the Laboratory for Interdisciplinary Statistical Analysis (UI-LISA) and Prof. O. E. Olubusoye for their support with the Monte Carlo simulation.

I greatly appreciate my lovely parents Mr. Ademuwagun A. Odetunmibi and Mrs. Grace O. Odetunmibi, my wonderful wife Farida Tunmibi, and my children (Enochand Michelle)for their understanding, encouragement and support. Thanks to Mr. Oladipo Toriola, Mr. Sayo Odetunmibi, Mr. Michael Odetunmibi, Mr. Moses Makinde and Mrs. Babakemi Makinde for their encouragement and prayers.

I sincerely appreciate Mr. David Ogunlana and Mr. Olanrewaju Dasylva for their assistance with data collection from theIndependent National Electoral Commission, Abuja. I also appreciate the support I received from Mr. Omotunde A. Muyiwa, Mr. Olayemi Odumesi, Mr. and Mrs. Kehinde Olalere, Mr. Abiodun O. Adesina, Pastor Basil Ninis and many others.

Sunday Olapade Tunmibi

January 2021

CERTIFICATION

I certify that this work was carried out by Mr. S.O. Tunmibi in the Department of Data and Information Science, University of Ibadan

....

Supervisor
W.M. Olatokun
B.A., M.Inf.Sc., Ph.D. (Ibadan)
Professor, Department of Data and Information Science
University of Ibadan, Ibadan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ACN	Action Congress of Nigeria
ANPP	All Nigeria Peoples Party
APC	All Progressives Congress
CPC	Congress for Progressive Change
FCT	Federal Capital Territory
FEDECO	Federal Electoral Commission
FSD	First Significant Digit
HTTP	Hypertext Transfer Protocol
INEC	Independent National Electoral Commission
LGA	Local Government Area
NDI	National Democratic Institute
PDP	People's Democratic Party
PERT	Program Evaluation and Review Technique
REC	Resident Electoral Commissioner
WWW	World Wide Web

CHAPTER ONE

INTRODUCTION

1.1 Background to the Study

An election is a process or a sequence of actions that results in a selection of a person for an office, dignity, or position of any kind; usually by the votes of a constituent body. This process may be as simple as counting raised hands in a room, or as complex as tallying votes across a multiplicity of jurisdictions (Simidchieva, Engle, Clifford, Jones, Peisert, Bishop, Clarke and Osterweil, 2010). The conduct of transparent and credible elections on a regular basis as established by the relevant constitutional and legal framework is a critical component of democracy. Generally, periodic and credible elections in a state or a nation are seen as a key component for enhancing the legitimacy of a government and strengthening the social contract between the government and the governed (Maendeleo Policy Forum, 2016).

The mode and period of conducting elections vary in different parts of the world. For example, based on the Fixed-term Parliaments Act 2011 in the United Kingdom, elections are held within a period of five years into the House of Commons following dissolution of Parliament by the Prime Minister. However, the five-year duration can be reduced if two-thirds majority of the members of Parliament voted for an early election. Election results from the polling stations are collated at the respective constituencies where the elected members of Parliament are announced by the local returning officers. The party with the majority of parliament members forms the government. Nevertheless, in the absence of an outright majority, parties can agree to form coalition government. However, not all countries running a parliamentary system of government in Europe adopt five-year interval between elections. While elections are held in a period of three years or five years in few countries, most countries in Europe adopt a period of four years (Hazell, 2010).

The mode of conducting election is different for countries running presidential system of government where the executive and legislature are elected separately. For example, in the

United States of America (USA), the president is elected every four years, with maximum of two terms in office through an indirect election with vote cast by electors of the Electoral College. The members of the House of Representatives are elected every two years and the members of the Senate are elected every six years into a staggered term because one-third of the Senate (known as the class) is elected every two years. The members of the House of Representatives and members of Senate are elected through a direct election without a term limit. Predominantly, the votes are recorded at the polling units with direct recording electronic machines or optical scan machines (Green, 2014).

Nigeria, like the United States of America, currently runs a presidential system of government. The Independent National Electoral Commission (INEC) is saddled with the responsibility of conducting elections every four years into offices of President and Vice President, Governors and deputy Governors, the Senate and House of Representativesand the 36 States Assemblies. The commission headed by the Chairman has 37 Resident Electoral Commissioners (RECs) for each of the 36 States and the Federal Capital Territory (FCT) of Abuja.The voting process includes accreditation of voters, balloting, counting of votes, collation and announcement of results.As a neutral, non-partisan electoralagency, INEC is expected to exhibit unalloyed impartiality and transparency in its conducts of elections inNigeria (Udu, 2015).

Electoral malpractices are committed with the aim of influencing the electoral results to favor a candidate through the adoption of bribery, cheating, illegal voting, intimidation, alteration of results and fraudulent pronouncement of the loser as winner with or without adjusting the electoral outcome (Ogbeidi, 2010). The capacity of electoral officers to rig elections cast a shadow over the electoral process of many democratic countries, including those with advanced democracies.Different reports from academics, journalists, political parties, lawyers, election monitors and observers, as well as concerned citizens have shown that many countries around the world, including the US involve in electoral malpractices (Alvarez, Hall and Hyde, 2008).

Election malpractice from 1959 till date has also become a challenge in the Nigerian political system. According to Edoh (2004), incidents of violence, and stuffing of ballot boxes as well as obstructions and intimidation of opponents were reported during the Nigeria 1959 parliamentary elections. The result of the election was a victory for the Northern People's

Congress, which won 134 of the 312 seats in the House of Representatives. Also, Awopeju (2011) noted that elections into the Western House of Assembly in 1965 ended in violence as a result of widespread rigging. The aftermath of the widespread rigging and violence was the first military coup which occurred in January 15, 1966.

According to Oromareghake (2013), election rigging was also reported during the elections organised by the military in 1979. The observed rigging during the election brought about the term "stolen Presidency", which has since become part ofNigeria's political vocabulary. The faulted election ushered in a civilian administration governed bythe National Party of Nigeria. Election rigging was also reported in the 1983 elections. Animashaun (2010) noted that there was mayhem in the two Southwest states of Oyo and Ondo as a result of the massive manipulation of votes in favor of the ruling National Party of Nigeria. The violence wasas a result of the alleged malpractices that took place during the gubernatorial elections in both states which were known to be strongholds of the major opposition party in favor of the ruling party. Apart from the loss of lives and materials experienced by opposition, the headquarters of the electoral body, Federal Electoral Commission (FEDECO), in Oyo and Ondo states were set ablase.On the 31st of December 1983, the military intervened once more and took over the government.It was not until May 1999 that democracy was restored in Nigeria.

According to Osinakachukwu and Jawan(2011), the polity had been so damaged that people no longer show interest in politics due to the prolonged reign of military dictatorship. The lackadaisical attitude shown towards the 1999 elections by Nigerians gave the military junta the free hand to manipulate the elections and handed power to Obasanjo whom the hierarchy wanted.The 2003 elections also failedto meet basic international standards. Agbaje and Adejumobi (2006) noted that the1999 and 2003 elections, like virtually all the other preceding elections in Nigeria'spost-colonial history, were classic cases of electoral malpractices. The 2007 elections are also known for massive rigging. According to Kia (2013), the domestic and international monitors, political elites and civil society groups held that more than the fraudulent occurrence in the 1999 elections,there were intense manipulations of the 2003 and 2007 general elections in a manner that the electoral outcome in some places differ from the actual voting pattern. Yagboyaju (2011) also noted that INEC showed incompetence and unfair influence in the 2007 elections which sparked heated debate and also followed by election re-runs in some states, as ordered by the court. Reports indicated that the electoral process was marred with falsification of results, ballot box snatching and other forms of fraudulent acts. Similarly, Osimen and Ologunowa (2013) reported that malpractices such as snatching of ballot boxes and underage voting were said to have occurred during the 2011 elections. According to them, the elections ignited riots in the Northern part of Nigeria and in the following months up to 1,000 people were reported to have died in post-election violence. Ladan-Baki (2016) also reported irregularities in the 2015 general elections. These irregularities include missing result sheets, fake result sheets, ballotbox snatching, Smart Card Reader snatching and connivance with INEC officials and Police to rig elections.

The commonest approaches often adopted in addressing election malpractices and associated issues are for the aggrieved parties to go to court or election tribunals (Mebane, 2004; Animashaun, 2010). Mebane reported that there were over voted ballots of more than 50,000 votes in the 2000 US presidential election in Florida that led the aggrieved parties to court. Although the Florida Supreme Court mandated a recountof votes in theweeks following the election, the U.S. Supreme Court decided to stop the recount and the winner, George W. Bush, was retained. Also, according to Animashaun, the 2007 general elections attracted the highest number of post-election litigations in Nigeria. The results of theelections declared by the electoral body were contested at the local, state and federal government levels. More so, a lot of the results declared by INEC were reversed by the election tribunals and the courts while it took nineteen months, after concluding the election before the disagreements and controversies thatarose from the presidential election were resolved.

To support the judicial approach, scholars have suggested the possibility of statistical approach which focuses on deductions that could be made from digital patterns of electoral results (Myagkov, Ordeshook, and Shaikin, 2009; Leemann and Bochsler, 2014). Myagkov, Ordeshook, and Shaikin (2009) analysed all federal elections between 1996 and2007 in Russia and between 2004 and 2007 in Ukraine. They found that errors increased inboth the 2004 and 2007 elections in Russia whereas errorsdiminished considerably since the second round of 2004 presidentialelectionin Ukraine. In their study, Leemann and Bochsler (2014) carried out statistical tests on the 2011 referendum in the Swiss canton of Berne. They compared the election returns of 342 wards in the same canton andfound empirical support for irregularities which could be due to some form of malpractices.

A typical statistical approach is election forensics (Beber and Scacco, 2008; Breunig and Goerres, 2011). Election forensics describes a nascent field of social science that uses statistical methods to verify whether election results are accurate or not (Mebane, 2010a).Election forensics techniques are based on finding patterns in the electoral returns data and the possible deviation of election datafrom the expected distribution of this pattern (Levin and Alvarez, 2013). In their study, Levin and Alvarez (2013) have developed election forensic indicators and procedures and also identified the techniques that can be applied to detect election malpractice. According to them, election forensic techniques include (i) the detection of anomalies in turnout distributions and in the relationship between turnout and vote shares (ii) detection of anomalies in the flow of votes between elections (iii) detection of deviations frompredicted vote shares (iv) and the detection of anomalies in digit distributions. Turnout distribution refers to the distribution of the percentage of eligible voters who cast a ballot in an election. Anomalies in the turnout across precincts or districts as well as anomalies in relationship between turnout and a candidate's share of the eligible electorate might be considered as evidence of malpractice. Also, a model of past elections could be used to predict the vote shares of recent elections. Deviation of the results of the recent elections from the model raises suspicion for malpractices. In addition, according to the authors, there is an expected pattern for the distribution of digits in an error-free election. Variation between the distribution of digits in the election results and the expected distribution also raises suspicion for malpractices.

Many of the common electoral forensic techniques are based on Zipfian distribution. Others focus on the application of agent-based models. In 1934, George Kingsley Zipf (1902-1950) discovered that if words are ranked in decreasing order of their frequencies in any given text corpus, the frequency of words in that text *f*is inversely proportional to its rank *r*. For example, in most English texts, "the" is most commonly used word. Zipf's statistics has shown that "the" accounts for about 7percentof the words, while the next in rank, "of", occurrs half of that number, and so on. Zipf's Law is represented thus:

$$f = k\tfrac{1}{r} \qquad (1.1)$$

where *f* represents word frequency, *r* denotes word rank and *k* denotes an arbitrary constant. The plot of *log-r* with *log-f* will produce a straight line that drops at the lowest value.

Chen and Leimkuhler (1986) observed that there are many variations of Zipf's Law such as the Lotka's Law of scientific productivity and Bradford's Law of bibliographic scattering.

Although Lotka's Law predated Zipf's Law, Chen and Leimkuhler observed that Lotka's Law, Bradford's Law and Zipf's Law are mathematically the same under some conditions and, hence, represent different perception of the same phenomenon.Lotka's law is generally used to understand the productivity patterns ofauthors in a bibliography(Adigwe, 2016). The law states that the number of authors making *n* contributions is about $1/n^2$ of those making one, and the proportion of all contributors that make a single contribution is about 60% (Lotka, 1926). Bradford's law, which is another variation of Zipf's law, was based on the observation that if a comprehensive literature search is conducted on some subject covering a specified period of time, often it will be found that the literature is scattered in a regular pattern over a very large number of sources (Bradford, 1934). Bradford found that when these sources are arranged in descending order of productivity, with the journal yielding most articles at the top of the list and the journals yielding the fewest at the bottom, the sources can be divided into a nucleus with several groups containing the same number of articles. None of these two variations, as observed from published scholarly works, have been adopted in studying election malpractices.

A variation of Zipf's Law that has been used to study election malpractice is the Benford's Law. In 1938, Frank Benford (1883-1948) observed that people use numbers that begin with digit 1 morethan those that begin with digit 9. Benford showed that, naturally, numbers from the observed datasets consistently fall into a pattern with low digits occurring more frequently in the first position than larger digits. Hence, sequences of digits are usually expected to be uniformly or randomly distributed, and statistically significant deviations from the distribution can be taken as evidence of alteration in the figures. This version of Zipf's Law has been used in image processing (Andriotis, Oikonomou and Tryfonas, 2013), detecting error in financial data (Durtschi, Hillison and Pacini, 2004) and detecting electoral irregularities (Mebane, 2006; Myagkov, Ordeshook and Shaikin, 2009; Breunig and Goerres, 2011).

Besides the Zipfian class of laws, there exist agent based modelling (Fowler and Smirnov, 2005; Benenson, Martens and Birfir, 2008). Agent based modelling is a modelling technology thatproperly fits into analysing the probable result of the continuous interaction of many boundedly rational agents in a dynamic environment (MacGregor, Edward, and Thomas, 2006).Agent based modelling has been deployed to analyse electoral outcomes (Cantu and Saiegh, 2011; Quratul-Ann, 2013). There are several approaches adopted in simulating with agent-based models. One of these methods is the application of game theory which is a decision taking tool that analyses the choice of optimum strategy in different conflicting situations for achieving individual objects out of the common goals of competitors in an election (Nagaraju, Vijender, Bikshapathi, and Lingam, 2012).According to Nagaraju etal (2012), the main feature of the game theory approach is modelling the rationality behavior of competing individuals. It focuseson optimal decision making when one individual's decisions affect the outcome of a situation for all other individuals involved. Hence, it has been applied for studying decision making for political campaign, such as whether or not to execute a door-to-door canvassing for votes or telephone solicitation(Blydenburg, 1976). The methodhasalso been used to study how party coalitions form, how voters' knowledge and emotion affect election outcomes and how political attitudes change through a campaign (Qui and Phang, 2020).

Other methods are the application of scripting techniques based on object-oriented methods (Wilenskyand Rand, 2011) and the application of the Monte Carlo analysis (Cantu and Saiegh, 2011). According to Wilensky and Rand (2011), scripting techniquesare used in the development of software for implementing agent based models. Wattenberg and Szabo (2013) adopted NetLogo, a popular software for agent based modelling, to model the USA general election in 2012. They focused on how voters base their presidential votes on issues such as immigration, tax policy and other issue spectrums. However, the major challenge of agent based modelling with scripting technique is that the models are only as good as the data that is available. This implies that detailed data on the system to be modelled is required. Also, the approach is full ofunchecked assumptions. According to Phelps (2012), agent based models are often sufficiently complex that deriving explicit solutions for quantitative aspects of their macroscopic behaviour is often impractical if not impossible, hence, they are often analysed using Monte-Carlo methods.Unlike the other methods, the Monte Carlo method can be applied to modelsystemshaving less detailed data. The Monte-Carlo approach has been specifically adopted in detecting suspicious electoral data (Kobak, Shpilkin and Pshenichinkov, 2014).

Monte Carlo analysis is a general term that refers to research that employs randomnumbers, usually in the form of a computer simulation (Johnson, 2011). Kobak, Shpilkin and Pshenichinkov (2014) analysed raw data collated from federal elections in Russia between 2000 and 2012. They used Monte Carlo simulations to confirm high statistical significance of man-made manipulationsin all elections since 2004. They discovered that several polling unitsbeyond the expected number reported turnout in integer percentages, rather than fractional values.

Empirical studies on election malpractices show that scholars focus on detecting anomalies in the distribution of turnout (Beber and Scacco, 2008; Levin, Cohn, Ordeshook and Alvarez, 2009), relationship between turnout and vote counts (Levin, Cohn, Ordeshook and Alvarez, 2009; Myagkov, Ordeshook, and Shaikin, 2009), deviation of vote counts from second-digit test of Benford's Law (Mebane, 2006; Breunig and Goerres, 2011), and deviation of electoral returns from a uniformly distributed last-digit (Beber and Scacco, 2008; Deckert, Myagkov and Ordeshook, 2011). This study attempts to apply Benford's Law and Monte Carlo simulation to model presidential election data in Nigeria.

1.2 Statement of the Problem

A free and fair election is generally believed to be crucial to the survival of democracy. A fair electoral process ensures that every vote is meticulously tallied andthe result is the careful aggregation of a society's preferences. However, different reports from scholars indicate that elections conducted in Nigeria are not fair (Edoh, 2004; Agbaje and Adejumobi, 2006; Animashaun, 2010; Awopeju, 2011; Osimen and Ologunowa, 2013). According to the scholars, elections in Nigeria are characterised with stuffing of ballot boxes, illegal voting, bribery and undue influence to manipulate results. These fraudulent acts serve as a tool for elected officials to cling to power against the will of the electorate and to undermine political opposition. Nevertheless, the common approach adopted in addressing electoral malpractice in Nigeria is for the aggrieved parties to lodge petitions in the court. Beyond the legal approach, little has been done to scientifically explain election malpractice in Nigeria.

Although, there has been an increasing emphasis on the application of election forensics to analyse election malpractice (Breunig and Gorres, 2011; Levin and Alvarez, 2013), the study carried out by Beber and Scacco (2008) on the 2003 presidential elections in Plateau State, northern Nigeria, remains the only known attempt to apply forensic tools to validate electoral returns in Nigeria. The authors adapted the last digits distribution test to detect errors in the vote counts and voters' turnout of electoral results collated at some wards in Plateau State. However, their study was too narrowpossibly because of the limited information available for the 2003 presidential elections in Nigeria.

This study attempts to explain how the study of digital distribution of vote counts and voters' turnout of electoral results can help in analysing anomalies in election data. In contrast to Beber and Scacco, this present study focuses on electoral returns collated at the state level across Nigeria, using two different but complementary methods, Benford's Law and Monte Carlo simulation. It is expected that the combination of these two methods would give a sufficiently useful insight about detecting patterns in the election data that may be useful in complementing other evidence on election anomalies.

1.3Objective of the Study

The main objective of this study is to modelvote counts and voters' turnout in Nigeria's presidential election data using Benford's Law and Monte Carlo simulation.

The specific objectives are:

1. To examine the digital distribution characteristics of vote counts and voters' turnout in Nigeria's presidential election data;
2. To determine if distribution of digits in vote counts deviates from distribution of digits as stated by Benford's Law;
3. To determine if distribution of digits in voters' turnout deviates from distribution of digits as stated by Benford's Law;
4. To examine if distribution of digits in actual vote counts deviates fromdistribution of digits in the generated surrogate vote counts of the Monte Carlo simulation; and,
5. To present a model for the application of election forensic tests.

1.4 Statement of Hypotheses

1. **H_{01}:**There is no significant correlation between the digital distribution of vote counts of presidential election results in Nigeria and the digital distribution of Benford's Law
2. **H_{02}:** There is no significant correlation between the digital distribution of voters' turnoutofpresidential election results in Nigeria and the digital distribution of Benford's Law
3. **H_{03}:**There is no significant relationship between the digital distribution of vote counts of presidential election results in Nigeria and digital distribution of Monte Carlo simulation

4. **H_{04}:** There is no significant difference in the voters' turnout across the six geo-political zones in Nigeria

1.5Scope of the Study

This study focuses on the vote counts or/and voters' turnout of the presidential election results collated across the 36 states and Federal Capital Territory (FCT), Abuja, covering the period between 2007 and 2015. The study adopted Benford's Law and Monte Carlo technique to analyse patterns in the electoral returns in Nigeria, using the results collated by the INEC at various state levels for the 2011 and 2015 presidential elections and results collated at the national level, without state-by-state breakdown, for the 2007 presidential elections. The study focused on vote counts of all parties in the 2007 presidential election result; vote counts of PDP (People's Democratic Party), CPC (Congress for Progressive Change), ACN (Action Congress of Nigeria) and ANPP (All Nigeria Peoples Party) for the 2011 presidential election result; vote counts of APC (All Progressives Congress) and PDP (People's Democratic Party) for the 2015 presidential election result; and voters' turnout in the 2011 and 2015 presidential election results.

1.6Significance of the Study

This research will provide information about whether patterns in election data could give useful insight into electoral anomaliesin Nigeria by using either the Benford number approach or the Monte Carlo method, or a combination of the two. The application of the two methods was necessitated by recent studies that have shown that Benford's Law alone is not sufficient to provide useful insight into irregularities in electoral returns (Levin and Alvarez, 2013). The adoption of Monte Carlo simulation and Benford's Law to study possible data irregularities in Nigeria election results would provide more insight for scholars and researchers who are interested in studying electoral malpractices in Nigeria. This study, therefore bridges the gap between what exists in the literature and what the future holds in election forensics.

The adoption of Monte Carlo simulation and Benford's Law to study election data pattern could also help the political parties to save the cost of litigation. It is common knowledge that the validity of all the presidential elections conducted by the Independent National Electoral Commission (INEC) since 1999 has been contested up to the Supreme Court. The theoretical insight into patterns of election malpractice could guide aggrieved parties in making better decisions on their chances of getting the desired court ruling. Hence, the time and resources spent on lengthy court rulings on election results could be minimised if not totally avoided.

The adoption of the combined methods in this study could also help the Independent National Electoral Commission (INEC) to avoid the time and resources wasted in recounting elections and/or rerunning elections in order to address election petitions. More so, studies have shown that a rerun of election does not guarantee a fair and/or acceptable result (Enabule and Ewere, 2010). With more insight into suspiciouselectoral data patterns, INEC might uncover better perception of how to curb election malpractices. This could also help to improve INEC's capacity to take preventive measures in future election cycles.

This study could also be of interest to election observers (both local and international). The combined approach adopted in this study may serve as digital tools that could enable election observers to make well-informed recommendation(s) for increased adherence to international democratic standards of free and fair elections. A proper insight into patterns formed by manipulated election results could not only help election observers to uncover discrepancy but also make the best assessment possible.

1.7Operational Definition of Key Terms

Agent-based modelis an aggregation of many interacting entities or agents in a simulation environment.

Electionis a formal process through which electorates decide and choose someone to hold public office through voting.

Election forensics is a social science field that intends to design data analysis tools which can be applied to detect discrepancy in electoral data.

Election malpracticeis the intentional or unintentional interference with the process of election which often produce a different outcome from the choices of the electorates.

Generated surrogate vote counts represent the simulated total number of votes recorded for an election.

Generated surrogate voters' turnout represents the simulated ratio of the total vote counts to the total number of registered voters.

Gerrymandering is the act of altering political boundaries so as to favour some group of people.

Modelling refers to the act or instance of making a model.

Simulation refers to use of algorithms or programming codes to replicate a system so as to predict its behaviour.

Valid votes refer to the total votes cast during an election.

Vote counts refer to the total votes recorded for an election.

Voters' preference refers to a greater liking by electorates for one political candidate or party over another or other political candidates or parties.

Voters' turnout represents the ratio of the total vote counts to the total number of registered voters.

CHAPTER TWO

LITERATURE REVIEW

2.0Introduction

This chapter presents the review of relevant literature. The first section focuses on election forensics; the second section focuses on Benford's Law and Monte Carlo analysis while the third section reviews application of Benford's Law and the last digits test. The fourth section reviews application of agent-based models and the fifth section focuses on the research framework. A summary of literature reviewed is presented in the last section.

2.1 Election Forensics

Election forensics is an emerging field that focuses on the application of data analysis tools for detecting discrepancy in election outcomes (Mebane, 2010a; Levin and Alvarez, 2013). The major challenge of investigating election malpractice is identifying the characteristicsthat separate the suspicious elections from the error-free elections. However, to draw an absolute conclusion on the integrity of an election is often impossible. An appealing approach in overcoming this challenge is to highlight anomalies from an electoral databy making deduction from the digital pattern of data distribution or by using mathematicalalgorithms(Beber and Scacco, 2008; Cantu and Saiegh, 2011; Klimek, Yegorov, Hanel and Thurner, 2012).

Klimek et al. (2012) investigated fraudulent activities in the data gathered on recent presidential elections from France, Romania, Russia and Uganda. The election data from the selected countries were gathered based on data availability. They tested for reportedstatistical features of voting results, and deviations thereof, in a cross national setting. They proposed a model to analyse and determine the extent to which the amount of malpractices might have influenced the electoral results. For simplicity, they assumed that a Gaussian distribution should represent the preferences of voters and their turnout, in every polling unit, with the average and the standard deviation selected from the actual data. The results of their analysis indicate that there is statistical evidence that malpractices occurred in the presidential elections of Russia and Uganda.

In a systematic approach to study election malpractice, Leemann and Bochsler (2014) carried out a study on the 2011 referendum in the Swiss canton of Berne. They noted that 30 out of the 383 municipalities declared that they have lost the ballots of the referendum, thereby rendering a recount of the ballots, mandated by courts, impossible. The authors were interested in knowing if this happened to cover possible fraudulent actions. Their models focused on interdependence of the referendum questions, the local and idiosyncratic characteristics that might explain parts of the results, and the assumption that referendum results can be explained with the partisan composition of the electorate. The result of their statistical test revealed that fewer empty ballots were counted for the tie-break question in the 30 municipalities which lost their ballots. The result suggests that some members of the electoral committee in the 30 municipalities might have filled in empty tie-break answer boxes.

In another study carried out in Nigeria, Beber and Scacco (2008) used deviation from uniformly distributed last digits of electoral returns to study digital pattern of election results. They analysed results sheets from wards in Plateau State, focusing on PDP (People's Democratic Party) and ANPP (All Nigeria People's Party) votes due to lack of detailed information in some polling units. They observed a lot of repeated numerals, variation in all vote countsand the accurately tabulatedtotal vote count for more than twenty percent of polling units, and a non significant relationship between turnout and suspicious distribution of digits. They also observed significantdeviations from the uniform distribution of the last digits, in particular for the numeral zero, which strongly suggests that electoral returns were indeed manipulated. Generally, in most of the studies on election malpractice, scholars have shown that many of the election forensic techniques are based on Benford's Law, the last digits test and Agent-based models (Fowler and Smirnov, 2005; Mebane, 2006; Beber and Scacco, 2008; Breunig and Goerres, 2011; Laver and Sergenti, 2012; Quratul-Ann, 2013).

2.2 Benford's Law and Monte Carlo analysis

In this section, a review of Benford's Law and Monte Carlo analysis is presented, starting with Benford's Law.

(a) Benford's Law

Simon Newcomb, in 1881, observed that numbers with 2, 3, and above occurred less frequently than numbers with a first digit of 1. Frank Benford, not knowing about Newcomb's observation, had likewise observationwhich he published in an article. This first significant digit (FSD) phenomenon was initially known as Benford's Law.Newcomb observed the probability (P) of a number having a particular nonzerofirst digit (d) as:

$$P(\text{First digit is d}) = \log_{10}[1 + (1/d)] \text{ where } d = 1, 2, \ldots, 9.$$

This formula leads to a continuous decrease in the distribution of the first digits with the assumption that FSD of numerals, in decimal, would be distributed in a skewed pattern. The formula also suggests that the probability of having a number between 1 and 9 as first digit is:

P(1) to P(9) = 0.301, 0.176, 0.125, 0.097, 0.079, 0.067, 0.058, 0.051, 0.046 respectively

This also implies that the occurrence of each second digit is given by

$$P(\text{Second digit is d}) = \Sigma^{9}_{k=1}\log_{10}(1+(10k + d)^{-1}) \text{ where } d = 0, 2, \ldots, 9.$$

That is, P(0) to P(9) = 0.120, 0.114, 0.109, 0.104, 0.100, 0.097, 0.093, 0.090, 0.088, 0.085 respectively.

Benford's Law, named after Frank Benford, originated from an observation of the first few pages of tables of common logarithms. Frank Benford noted that these first few pages were more worn than the later pages (Benford, 1938). Based on this observation, he came up with a hypothesis that people were looking up the logs of numbers with low first digits more frequently than the logs of numbers with high first digits. According to him, this is so because numbers having logarithms with low first digits (such as 1, 2, and 3) were more common in the world than those with high first digits (such as 7, 8, and 9). The first digit of a number is the leftmost digit. However, 0 is inadmissible as a first digit. Hence, 3 would be the first digits of both 3000 and 0.0037. In an empirical test carried out on the first digits of 20 diverse lists of numbers with 20,229 records, Benford observed that the graph was skewed to the left, in favour of the low digits.

Benford's Law, having no mathematical derivation, is considered as law of nature by scholars, hence, attempts at explaining the law has been focused on scale-invariance and base-invariance. The scale-invariance condition says that if there is a universal law of nature thatgoverns the distribution of leading digits then it should notdepend upon the units in which the numbers are measured. That is if data should be converted from km^2 to square miles, the same distribution of leadingdigits should result. Likewise, the base-invariance condition implies that the same distribution of leadingdigits should resultif there is an arbitrary change of the numerical basis. Benford's Law satisfies both the scale and base invariance condition (Raimi, 1976; Fewster, 2009; Judge and Schechter, 2009).

Benford's Law with its left skewed first significant digits (FSD) distribution has been proven to hold with a lot of different datasets that include the number of people in towns, the half-lives of radioactiveatoms, data used by corporations for budgets, and the frequencies of citations by papers. The range of practical use and relevance of Benford's Law is impressive. In all of these applications, the FSD distributions was adopted to represent a constantly changing blend ofdata outcomes whose resulting combination is not limited, with respect to thepossibility of extending across the nine digit space (Judge and Schechter, 2009).

A major advantage of Benford's Law is the fact that the law is not influenced by means of scale invariance (Fewster, 2009). Another benefit of Benford's Law is that it is a useful, inexpensive and easy to apply tool for detecting possible anomalies in large data set. It can also be used as a proactive approach to early detection of discrepancy (Kellerman, 2014). Benford's Lawcould also help to reduce sample size while auditing large data set by using patterns of irregularities to separate suspicious data from unsuspicious data (Petucci, 2005).

Despite its wide acceptance, studies have shown that Benford's Law has some shortcomings. According to Kellerman (2014), the law will not be successful when the data set has inner range of maximum to minimum values; and when the data set is nothaving enough digits required for conformity with Benford's distribution. Likewise, approximated data may not adhere to the Benford's distribution. In addition, Benford's distribution tests might not detect the difference between unintentional human error and intentional error. Hence, the test only focuses on interpretation of the difference between observed and expected distribution (Mebane, 2006; Breunig and Goerres, 2011; Deckert, Myagkov and Ordeshook, 2011).

According to Benford, there is 30% probability that '1' will occur as the first digit and 4.6% probability that '9' will occur as the first digit. Also, there is approximately 12% probability that '0' will occur as the second digit and 8.5% probability that '9' will occur as the second digit. The distributional pattern expected for Benford's first and second digits is shown in Figure 2.1 and Figure 2.2.

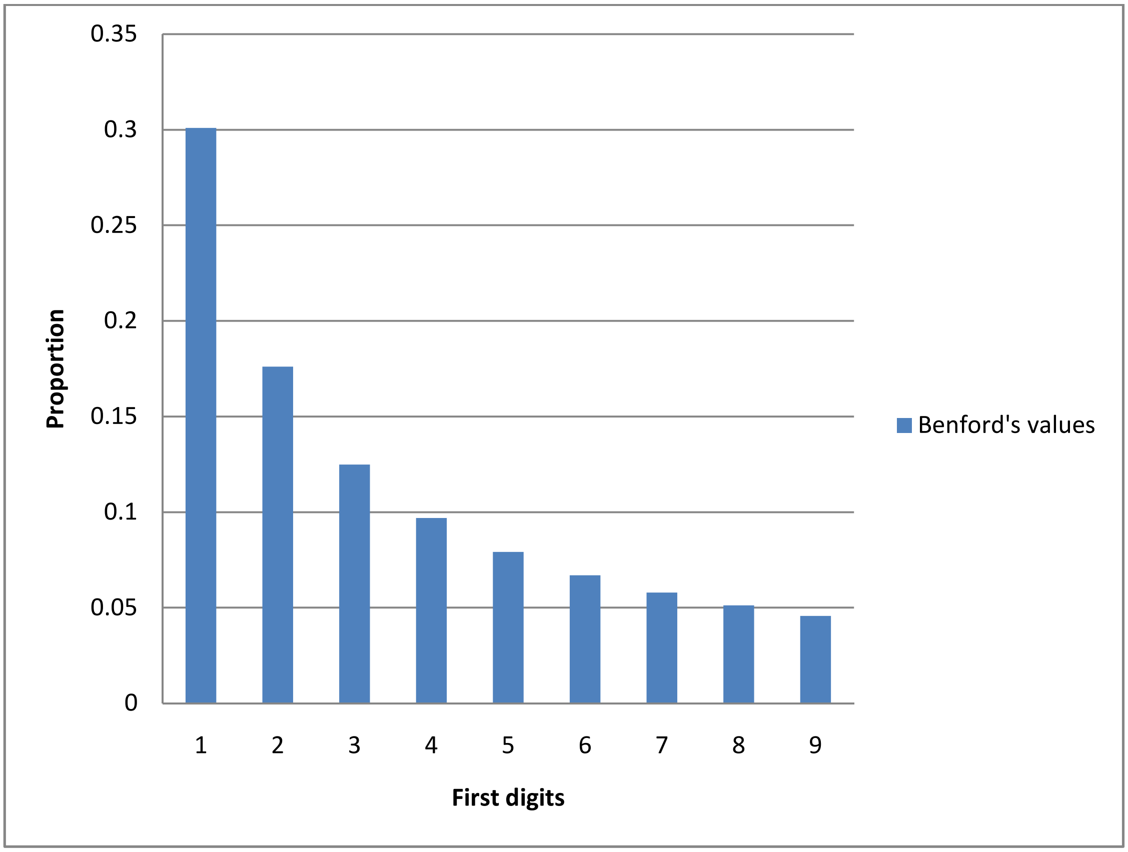

Figure 2.1: Benford's first digits distribution

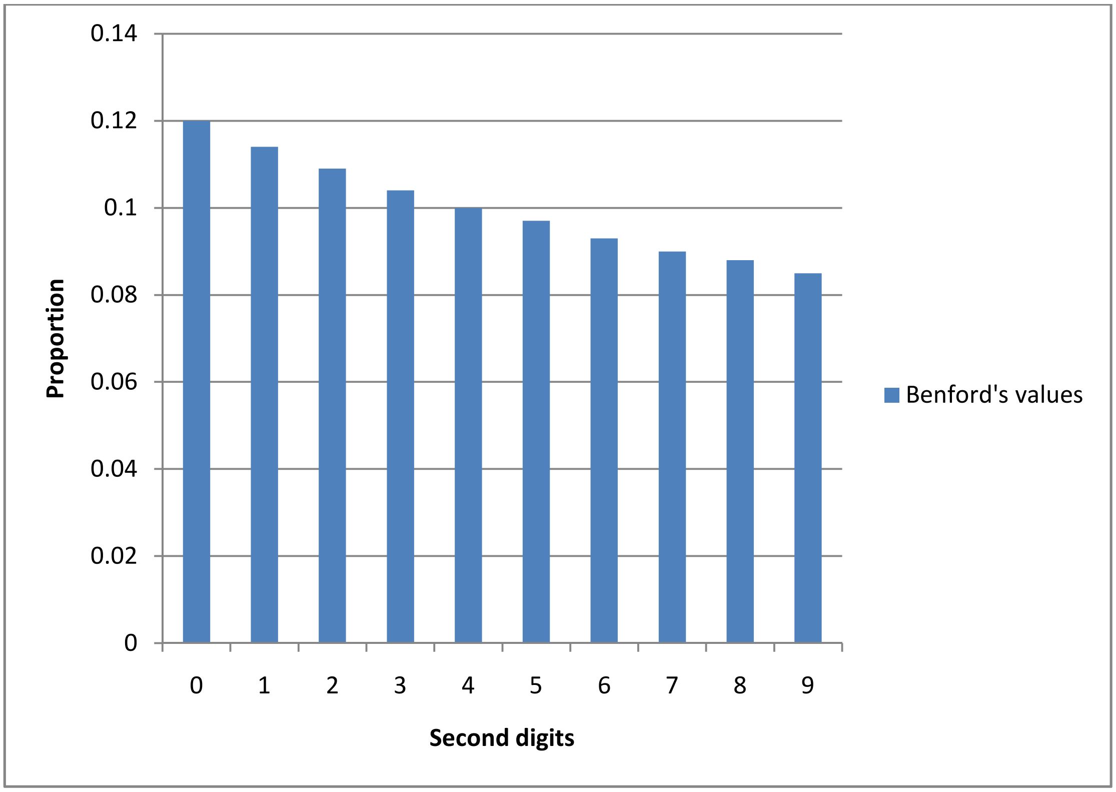

Figure 2.2: Benford's second digits distribution

The frequency distribution of the first digits is expected to be left skewed as shown in Figure 2.1. The frequency distribution of the second digits, as shown in Figure 2.2, implies that the distribution approaches a uniform distribution for higherorderdigits(Breunig and Goerres, 2011). This supports the argument of Beber and Scacco (2008) that the frequency distribution of last digits is expected to be uniformly distributed. An illustration of a uniform distribution is shown in Figure 2.3.

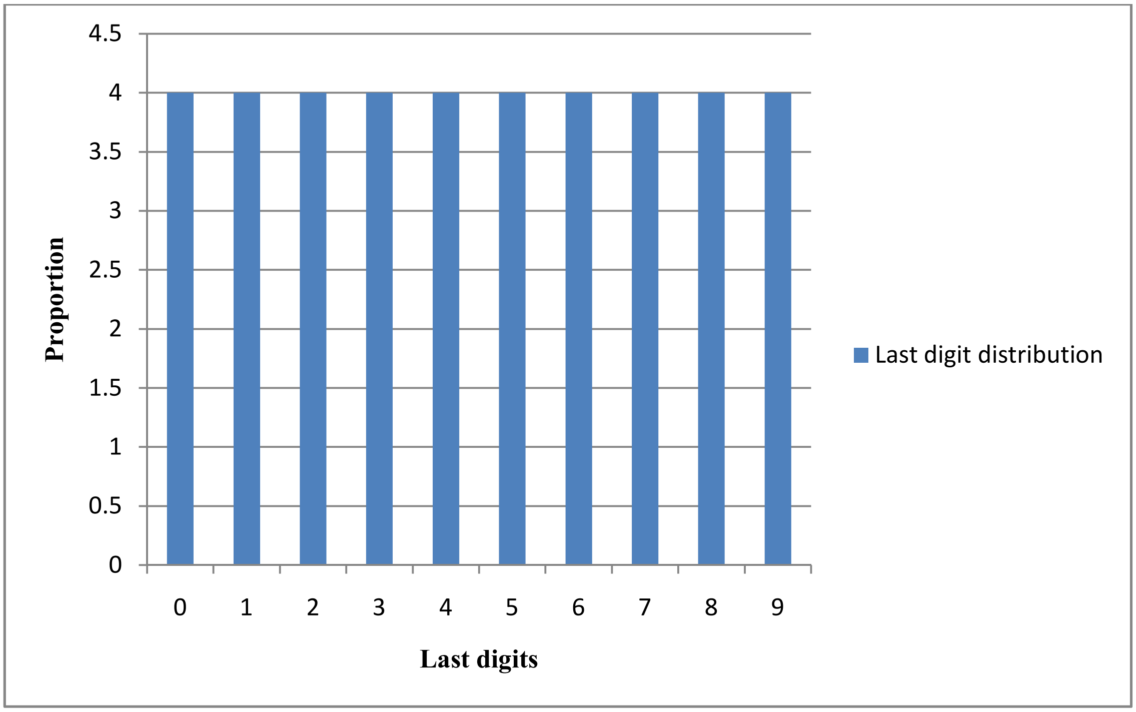

Figure 2.3: Uniform last digits distribution

(b) Monte Carlo analysis

Monte Carlo analysis is a technique in agent-based modelling. An agent-based model can be termed as a collection of multiple, interacting agents, situated within a model or simulationenvironment such as represented by the artificial world (Heppenstall, Crooks, See and Batty, 2012). Agents can be depictions of animate objects suchas human beings which can move freely within an environment or inanimate entities with static locations but having dynamic states. These inanimate or animate objectsusually operate based on some rules which determine their relationships with other entities (or agents) and their encircled environment.The rules, which are the basics of an agent's behaviour, aretypically based on if-else statements.

Although there is no general acceptability on the actual definition of the term "agent," different definitions seem to agree onmore points than they disagree. The basic feature of an agent is the capability of the different elements to make independent decisions. This requiresagents to be active rather than being passive (Macal and North, 2006). While introducing the need for agent based model, Macal and North emphasised thatthe systems to be analysed and modelled are becoming more complex in terms of their interdependencies. More so, traditional modelling tools are no longer asapplicable as they once were.

The historic root of Agent-based modelis in complex adaptive systems (CAS). Unlike System Dynamics which adopts the top-down systems approach, the assumption in CAS is that systems are built from the ground-up. The focus of CAS is on howcomplex behaviours arise in nature among different free agents. More so, agent-based model seems to be descriptive, with the aim of modelling the actual or probable behaviour of entities, rather than seeking to optimise and select optimal behaviours (Macal and North, 2006).

The firstsocial agent-based simulation was developed by Thomas Schelling(Schelling, 1978). Schelling applied cellular automatato study housing segregation patterns, in which agents representpeople and agent interactions represent a socially relevant process. The Schelling model showed that it is possible to have patterns that are not necessarilyimplied or consistent with the objectives of the individual agents. Some years later, Epstein and Axtell extended the notion of modelling people to growingentire artificial societies through agent simulation called Sugarscape model.Sugarscape model indicated that agents could emergewith a variety of characteristics and behaviours suggestive of rudimentary and abstract society. These Emergent processes could include patterns of death, conflict, war, trade, wealth, culture, and so on (Epstein and Axtell, 1996; Macal and North, 2006).

One of the greatest difficulties of utilising agent-based model concerns the issues of verification and validation. Verification is the process of ensuring that the model to be implemented agrees with the design while validation is the process of ensuring that the model to be implemented agrees with the real-world. Validation has to do withthe goodness-of-fit of the model to data or the system being modelled (Crooks, Castle and Batty 2008; North and Macal 2007; and Casti, 1997).

2.3 Application of Benford's Law and the Last Digits' Test

In a study on first digits distribution of Benford's Law, Berdufi (2014) carried out a study on Albanian parliamentary elections of 2009. He applied the first digits distribution to the party votes' results for each of the 12 districts involved in the elections. The results show that the election results were manipulated, with some parties benefiting more from the manipulations than the others. He observed that there was no correlation between the counted votes' distribution of the real data and Benford's distribution rate. In a related study, Roukema (2014) applied the first digits distribution test to the results presented by the Iranian Ministry of Interior on the 2009 presidential election. The results of the analysis show a highly significant excess of the first digit 7, for one of the four candidates involved in the election, compared to the expectation from the Benford's Law.

In a combined approach, Pericchi and Torres (2011) applied the first and second digits distribution test of Benford's Law to the 2004 US presidential elections, the Puerto Rico (1996, 2000 and 2004) governor elections, the 2004 Venezuelan presidential recall referendum and the previous 2000 Venezuelan presidential election. The results of their analysis show that, apart from the 2004 Venezuelan presidential recall referendum, the results of the others elections conform to the first and second digits Benford's distribution. The second digits distribution of Benford's Law was only rejected for electronic voting units in the Venezuelan recall referendum. The electronic results of the votes in favour of the NO violate the second digits distribution of the Benford's Law.

In a study on second digits distribution, Mebane (2006) gave reasons to support the notion that the number of occurrence of digits of election counts from the precincts is almost the same as the distribution predicted by Benford for the occurrence of second digits. This is expected because the total distribution of digits (from the first digits to the last digits of counts from general election) emerges from a combination of different intersecting electoral processes, rather than from a single process. However, Mebane also gave reasons to support the arguments that the first digits of election counts cannot have enough variation - for example, there are situations where a particular political party, due to high popularity within some areas, has similar votes from precincts of almost the same size. In such situations, the adoption of the distribution of second digits, as stated by Benford's Law, might be a better approach to study irregularities in the election process. As a fact, the numerous empirical tests, carried out by Mebane, with different types of empirical data sets from various countries supported the Benford's Law for the distribution of second digits (Mebane, 2008; Breunig and Goerres, 2011).

Breunig and Goerres (2011) investigated electoral irregularities in the 1990 to 2005 Bundestag elections ofunified Germany. The authors aggregated the candidate and party list votes for the three major political parties. They compared the number of occurrence of digits from the votes of different candidates and parties with the expected occurrence of second digits as stated by Benford's Law. The result of their study shows minimal evidence of election irregularities with respect to candidates' votes. However, they discovered fifty one violations in the party list votes for three different parties out of the several tests that were conducted.

Mebane (2010b), usingthe precinct data from the US presidential election of 2008 and the US House elections of 1984, argued that it can be necessary to associate the second digits of votes with a suitable covariate (this covariate could be the margin between the winner and the loser) so as to carry out a proper analysis on the election. This will help to determine the influence of other factors such as strategic voting patterns of the electorates and voting ruleson the distribution of digits in electoral results. However, the question is on how the observed patterns differ from other patterns with respect to the different kinds of election malpractices. Hence, several studies reveal that application of Benford's Law alone is not sufficient for the validation of data irregularities (Deckert, Myagkov and Ordeshook, 2011; Mebane, 2006; Beber and Scacco, 2008).

In a research conducted by United States Agency for International Development (2017) on the 2014 parliamentary elections in South Africa, the analysis focused on the turnout and the votes for three political parties. Using the second-digits and last digits test, the study indicated that the results of statistical analyses for one of the political party (ANC) revealed that the vote counts were error-free, but the results for the other two parties (Dem Al and EFF) showed that the votes are suspicious. They also observe that the last digits of the rounded turnout percentages have so many occurrences of zero or five (this was termed P05). In a similar research conducted by United States Agency for International Development (2017) on the 2001 Bangladesh election, where the Bangladesh National Party (BNP) won in a majority of the districts, the second-digits test suggest that both strategic voting and election malpractices occurred. The result of analysis on the last digits test of vote counts or vote proportions or of turnout also show a mix of results that are either significantly too large or significantly too small.

Beber and Scacco (2008) also developed a digit based test that exploits human biases in number generation. They analysed data from the 2002 parliamentary elections in Sweden, 2003 presidential elections in Plateau State (Nigeria), the 2000 and 2007 presidential elections in Senegal, and the 1924 and 1928 presidential elections in Chicago. They proved that last digits of electoral results should occur with equal frequencies. The assumption is that vote counts generated by corrupt officials will not have uniformly distributed last digits. They emphasised the fact that human tends to be biased in the production of random numbers (when generating figures). They seem to choose lower digits, avoid repeating numerals and also prefer adjacent numerals. The results of their analysis reveal that the elections in Sweden and the 2000 presidential elections in Senegal conform to the uniformly distributed last digits. On the contrary, the results of the 2007 presidential elections in Senegal resemble those from the 2003 presidential elections in Plateau state, Nigeria. They both show a non uniform distribution for the last digits, with excessive zeros. The results of their analysis also suggest a non uniform last digits distribution for one out of the two main candidates in both the 1924 and 1928 presidential elections in Chicago.

In a related study, Levin, Cohn, Ordeshook and Alvarez (2009) analysed the patterns of last digits distribution for the 2006 presidential election, the 2007 constitutional referendum and the 2009 constitutional referendum in Venezuela. The results of their analysis did not reveal any obvious non-random pattern such as avoidance of zeros and fives. All the last digits have approximately the same 10 percent incidence. They, however, noted that the shortcoming of the last digits test is that electronic irregularities might be designed in such a way that manipulation of the last digits is completely random. More so, the last digits distribution test requires extremely weak distributional assumptions.

2.4 Application of Agent-based Models

Laver and Sergenti (2012), on their own, started with thetwin premises that understanding multiparty competition is a core concern for everyoneinterested in representative democracy, and that we must understand multipartycompetition as an evolving dynamic system, not a stationary state. Given these premises, they investigated the dynamics of multiparty competition using computational agent-based modelling. This allowed them to modeldecision making by party leaders, in what was clearly an analytically intractable setting, interms of the informal rules of thumb that might be used by real human beings, rather thanthe formally provable best response strategies used by traditional formal theorists. Their study was fundamentally about decisions made by party leaders.

In another research, Fowler and Smirnov (2005) developed an agent-based model of dynamic parties, having itssocial turnout designed around the emergence in the social science fields. They described and analysed an agent-based model (ABM) of repeated electionsin which voters and parties behaved simultaneously. They placed voters in asocial context and allow them to mingle with each other when deciding whether or not to vote. The researchers also let parties chose the platforms they offered and these choices might change from election to election depending onfeedback from the electorate. Their model yielded significant turnout, different platforms,and various outcomes which agree withthe empirical literature on social turnout as well as the rational calculus ofvoting model.

Adopting Monte Carlo technique, Oleg (2011) simulated a total sample of 80,000 precincts in Russia and discovered that the higher the turnout, the less opportunity for falsification. Likewise, Kobak, Shpilkin and Pshenichinkov (2014) hypothesised that thefrequency of reported round percentages should be increasedif election results are manipulated or forged. They applied Monte Carlo methods to raw data from federal elections (2000 to 2012) held in Russiaand detected falsification in turnout and/or leader's result in all elections since 2004.

In another related study, Cantu and Saiegh (2011) created a set of simulated elections using Monte Carlo method as part of the tools to diagnose electoral irregularities in Argentina.

Their study indicated that Monte Carlo method can help in diagnosing electoral irregularities using recorded vote counts. Also, Rivest and Shen (2012) carried out a research on post-election auditing based on Bayes audit. Monte Carlo simulation was adopted in designing the inner loop of the Bayes audit. The Monte Carlo simulation was used to predict the possibility of a difference between the observed outcome of the election and the expected outcome from the Bayelsian model. Their study gave an experimental evidence of the effectiveness and efficiency of the Monte Carlo technique.

2.5 Research Framework

Based on the objectives of this study, and the review of theoretical and empirical literature, the adopted variables are represented in an interrelationship in the framework in Figure 2.4. The framework indicates that the digital distribution of the selected presidential election results are represented as voters' turnout and vote counts. The framework also indicates that the two approaches adopted for this study are the Monte Carlo simulation and the Benford's Law for the first and second digits.

The framework proposes a test of relationship between the digital distribution of vote counts for the selected presidential election results in Nigeria and the digital distribution of Benford's Law for the first and second digits. It also indicates a test for the relationship between the digital distribution of voters' turnoutfor the selected presidentialelection results in Nigeria and the digital distribution of Benford's Law for the first and second digits.

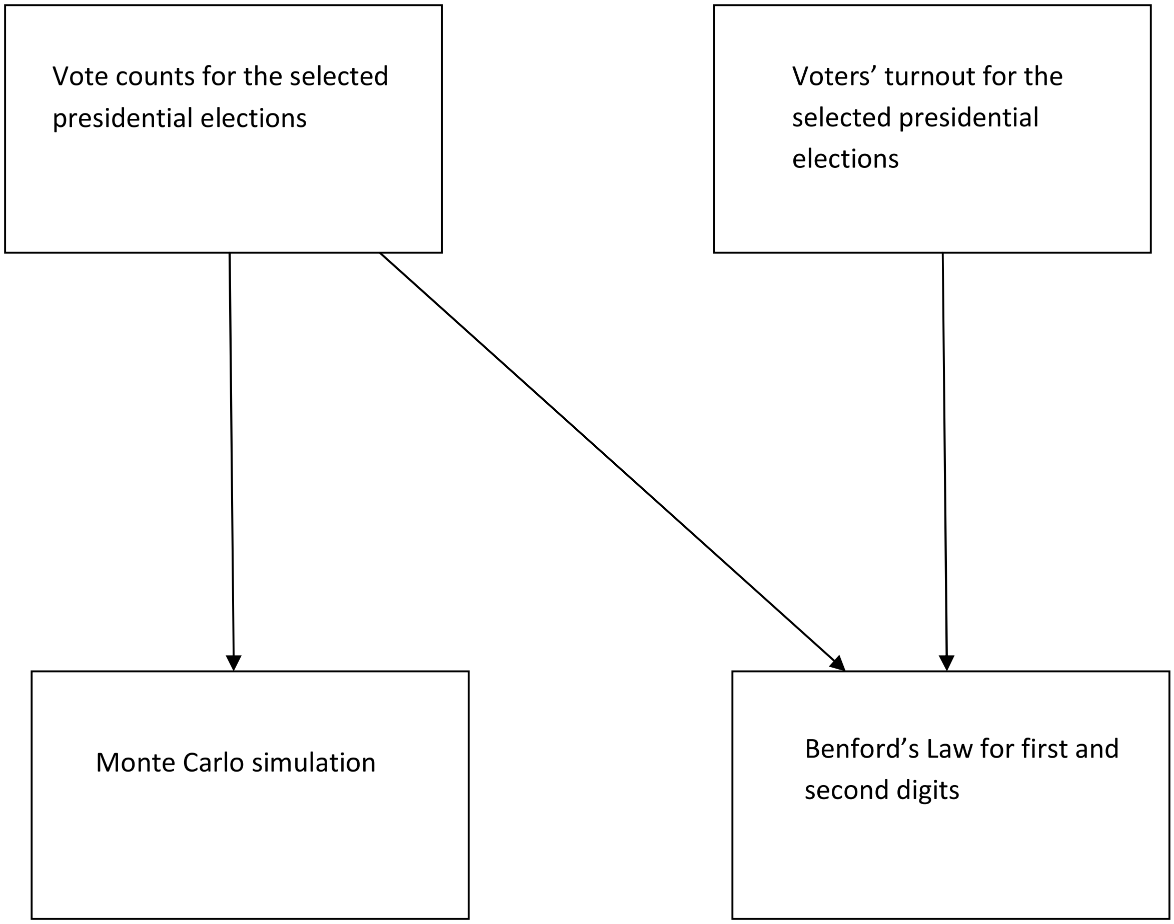

Figure 2.4: Research framework

In addition, the framework proposes a test for the relationship between the digital distribution of vote counts for the selected presidential election results in Nigeria and the results of the Monte Carlo simulation.

2.6 Summary of Literature Reviewed

So far, the review has focused on election malpractices and the application of election forensics, which include Benford's Law and Monte Carlo analysis. Scholars have focused on the application of the first and second digits of Benford's Law for detecting anomalies in election counts (vote counts and voters' turnout).Their studies interpreted deviations of the actual frequencies of numerals of the election counts from the first and second digit (Benford's Law) as indicating irregularities in vote counting. However, some argue that the second digit is more appropriate than the first digits. Their argument is based on the assumption that the first digit cannot vary enough for vote counts, especially at the precinct level where vote counts are very low.

Other scholars argue that an examination of last and next-to-last digits of election counts is a better forensic tool for detecting election malpractice. They stated that the last digits of electoral returns, in an error-free election, should be uniformly distributed. They also argue that pairs of adjacent digits are abundant on suspiciouselection return sheets while pairs of distant numerals appear with lower frequency. However, some scholars preferred to focus on the application of agent-based models to electoral data. Most of these scholars demonstrated that application of agent-based model, using Monte Carlo technique, is appropriate for detecting deviation from the expected distribution of votes and voters' turnout.

Similar to Mebane (2006), Pericchi and Torres (2011), Breunig and Goerres (2011), Berdufi (2014) and Roukema (2014), this research work will apply the first digit and second digit of the Benford's Law in analysing patterns in electoral data. This study will also focus on the application of Monte Carlo technique as demonstrated by Oleg (2011), Rivest and Shen (2012), Quratul-Ann (2013), and Kobak et al. (2014). In addition, the last digits of vote counts will be tested based on the assumptions of Beber and Scacco (2008) which has gained an increasing attention by scholars (Levin, Cohn, Ordeshook and Alvarez, 2009; Deckert, Myagkov and Ordeshook, 2011; Levin and Alvarez, 2013; United States Agency for International Development, 2017).

CHAPTER THREE

RESEARCH METHODOLOGY

3.0Introduction

This chapter presents the research methodology for the study. The first section focuses on research design, the second section presents data source; followed by the third section which focuses on the location and population of study while the fourth section focuses on sampling size and technique. The fifth section presents instruments for data collection, the sixth focuses on data analysis and the last section focuses on ethical considerations for the study.

3.1 Research Design

This study adopted modelling and simulation approach to examine patterns in electoral data gathered from the Independent National Electoral Commission of Nigeria. This study adopted Benford's Law and Monte Carlo analysis for modelling and simulation. Benford's Law is a variation of the Zipfian model while Monte Carlo analysis is one of the approaches to agent-based modelling. Both Benford's Law and Monte Carlo analysis have been applied to study irregularities in electoral data (Breunig and Goerres, 2011; Kobak, Shpilkin and Pshenichinkov, 2014).

Zipfian model is a consistent ranking pattern that is applicable to the frequency of occurrenceof different words, populations of cities within a country, the sizes of competing companies, among others (Frequencies, 2002).Benford's Law focuses on an observed pattern in the occurrence of numerals in the real world. According to Benford's Law, the occurrence of different digits in a set of numbers differ.The Law proposes a skewed distributional pattern for first digits of numerals. The distribution approaches uniform distribution as the analyst considers higher digits(Breunig and Goerres, 2011). For clarity, the first digit in 254689 is 2, the second digit is 5 and the last digit is 9. Agent based model, on the other hand, uses systematic approaches thatallows the test of "what if " and hypothetical scenarios, visualisations of patterns, and maintains a repeatable and recoverableprocess (Quratul-Ann, 2013). The adopted agent-based model approach, Monte Carlo analysis, is a research strategy that incorporates randomness, usually in the form of computer simulation, into the design, implementation or evaluation of theoretical models (Johnson, 2011).

The deviation of digital distribution of election data from the Benford's distribution and Monte Carlo Simulated data indicates the presence of anomalies in the election data. These methods (Benford's Law and Monte Carlo analysis) were applied to find patterns in the distribution of presidential election data that deviate from the expected distribution.

3.2 Data Source

Data for the study were extracted from the website of the Independent National Electoral Commission (INEC) in Nigeria, focusing on the 2007, 2011 and 2015 presidential election result sheets collated in the 36 states of the federation and Abuja - the Federal Capital Territory. Election result sheets from INEC consist of registered voters, accredited voters, valid votes, rejected votes and votes cast. This study focused on valid votes and voters' turnout. The total votes cast divided by the total number of registered voters represent the voters' turnout while the total number of valid votes for each candidate represents the vote counts. The election results used for this study consist of twenty four presidential candidates in the 2007 presidential elections; twenty presidential candidates in the 2011 presidential elections; and fourteen presidential candidates in the 2015 presidential elections.

3.3 Location and Population of Study

This study focused on the presidential election results collated across Nigerian 36 states and Federal Capital Territory (FCT), Abuja where elections took place. According to Alli (2015), INEC has a total of 119, 973 polling units across the country. However, the number of polling units varies from state to state, with Lagos State having the highest (8,462) and Nasarawa State having the lowest (1,495). Apart from the 36 states, the FCT, Abuja has 562 polling units.

The result sheets, together with all other used and unused materials,from the polling units are movedto the Registration Area Collation Centres, and thereafter to the Local Government Area (LGA) collation centres. The results are batched and transferred from the LGA collation centres to the State Collation Centres not later than 48 hours after each poll. The resultsare forwarded from the State Collation Centres to the Presidential Returning Officer who, in return, forwards the results to the Chief Returning Officer. The Chief Returning Officer is the Chairman of INEC. He collates and announces the vote counts of candidates and declares the winner of the presidential election.

The population of this study includes the 57,938,945; 60,823,022; 61,567,036; 73,528,040; and 67,422,005 registered voters for the 1999, 2003, 2007, 2011 and 2015 elections, respectively (Sheriff, Abdullahi, and Kabir, 2015).The population was collated across the 36 states and the FCT, Abuja from the inception of the Fourth Republic (1999 - 2015).

3.4 Sampling Size and Technique

Purposive sampling was used to select the2007, 2011 and 2015 presidential election results for this study because detailed election data from INEC for the 1999 and 2003 presidential elections were not available to the researcher. The target of analysis is the result collated at the state level, not the polling unit. Vote counts at the polling units, which in most cases are less than three digits for each candidate, are not sufficient for the analysis. According to Petucci (2005), a small sample cannot conform to Benford percentages, and might lead to distributional patterns that differ from Benford's Law. Data sets conform to the expected distribution of Benford's Law as they increase in size. Nigrini (2000) also observed that data sets should have numerals of up to four or more digits in order to conform to Benford's Law.

3.5 Instruments for Data Collection

This study used secondary data. An official letter of request for data collection (see Appendix 1) was collected from the office of the Director, Africa Regional Centre for Information Science, and submitted to the office of the Resident Electoral Commissioner, INEC, Lagos State, as well as the office of the Chairman, INEC, Abuja. The letter was processed and given to the appropriate Heads of Department for the release of available data on the 2007, 2011 and 2015 presidential elections. The data gathered from INEC, Lagos State was based on the results of the presidential election from the different local governments in Lagos State alone. Because of the limitation of what is available at INEC state chapter (such as INEC, Lagos State), the researcher proceeded to INEC headquarters in Abuja. The only detailed data available to the researcher at the national level was the 2015 presidential election results. This was made available to researchers on their public domain website. After two years of no response on acquisition of other election results from the INEC headquarters, the researcher resorted to other sources. The 2011 presidential election result adopted by the researcher was published by Nigerian Muse (2011) on its webpage. The result is an extract from the 2011 presidential results published on INEC webpage (http://www.inecnigeria.org/results/). The INEC webpage was updated in March 2015, and it is no longer active.

The widely criticised 2007 presidential election result was not as detailed as the 2011 and 2015 presidential results. Unlike the other results that give a state by state breakdown of the presidential election results, all the sources accessed by the researcher for the results of the 2007 presidential elections have records of only the total votes gathered by each of the presidential candidates. The only exception was the report published by the National Democratic Institute for International Affairs in April, 2008. NDI (2008) was able to publish a result table for 10 states, out of the 37 states (including the FCT) in Nigeria. The table shows the result, by state, for the 3 leading parties (PDP, ANPP and AC) and it was an extract from The Punch Newspaper of Monday April 23, 2007. Since the published result on 10 states is not sufficient for analysis, the researcher adopted the total votes gathered by each of the presidential candidate, as released by INEC and published by NDI (2008), in analysing the 2007 presidential election results.

3.6 Data Analysis

The study analysed vote counts of all political parties in the 2007 presidential election result; vote counts of PDP, CPC, ACN and ANPP for the 2011 presidential election result; and vote counts of APC and PDP for the 2015 presidential election result. The study also analysed the voters' turnout in the 2011 and 2015 presidential election results. All the political parties in 2007 presidential election results were selected for analysis because the result given by INEC is a summary of parties' votes without state-by-state breakdown or any other details. Only parties having vote counts with required digits in the 2011 and 2015 presidential elections were selected because studies show that a small sample cannot conform to Benford's percentages. Scholars suggest that numbers in data sets should have four or more digits for a good fit with Benford's Law (Nigrini, 2000; Petucci, 2005; and Kellerman, 2014). The 2007 presidential election result was not selected for analysing voters' turnout because the result lacks necessary details.

The first level of analysis is the description of digits' distribution in the 2007, 2011 and 2015 presidential election results using descriptive statistics. The adopted descriptive statistics include frequency counts of digits in the election results and measures of central tendency. The descriptive analysis of the election results focused on the first, second and last digits' distribution of vote counts in the summary of results for the 2007 presidential elections; the vote counts of PDP, CPC, ACN and ANPP for the 2011 presidential elections; the vote counts of APC and PDP for the 2015 presidential elections; as well as the first, second and last digits of the collated voters' turnout for the 2011 and 2015 presidential elections. The vote counts for each of the political parties represent the valid vote counts while the voters' turnout for each state is calculated as total vote cast in a state (total number of valid votes and total number of rejected votes) divided by total number of registered voters in the state.

The next level of the analysis focused on the application of the first and second digits' distribution of Benford's Law. First, the frequencies of all numerals for the first and second digits of the total vote counts of the 24 political parties for the 2007 presidential elections; vote counts of PDP, CPC, ACN and ANPP for the 2011 presidential elections; the vote counts ofAPC and PDP for the 2015 presidential elections; and of the voters' turnout for the 2011 and 2015 presidential elections were calculated. Thereafter, the distribution of these calculated frequencies were compared with the patterns predicted by Benford's Law for the first digit and second digits. The mean values for the first and second digits in both the vote counts for each of the political parties and the voters' turnout were also compared, respectively, to the expected mean of the first and second digits in Benford's Law.The mean is calculated as the sum of all values divided by the total number of cases.

The analysis on the application of agent-based model was based on Monte Carlo simulation. Monte Carlo simulation was not carried out on the 2007 presidential election result because the result lacks necessary details. Surrogate data were generated across the different states in Nigeria that matched the percentage votes, of the two leading parties (PDP and CPC) for the 2011 presidential elections andAPC and PDP for the 2015 presidential elections, as close as possible. For each Monte Carlo simulation, 1,000 iterations were performed for the 37 states in Nigeria, by using a programming application called R language. The average of the 1,000 iterations was selected by the R program for each surrogate data. By using the R environment for statistical computing and visualization, surrogate vote counts for each of the two leading political parties (in 2011 and 2015) were generated as a draw from a PERT distribution. The mean of the PERT distribution is the weighted average of the minimum, most likely and the maximum values that the vote counts or voters' turnout may take. Since the least voters' turnout across the state for both elections is greater than 20%, the PERT distribution for the simulated data assumed a worst case of 20% voters' turnout where all the electorates that turned out also voted for a political party as the minimum value of vote counts for the political party. The PERT distribution for the simulated data also assumed a maximum situation where a particular party won the entire vote cast. The most likely value represents the actual vote counts recorded by the political parties. The generated surrogate data from the PERT distribution was simulated to mimic the percentage votes secured by each of the two leading political parties across the state as released by INEC. Thereafter, the surrogate vote counts for each of the political parties were compared to the actual vote counts for each of the political parties. The R programming codes used for the Monte Carlo simulation are shown on Appendix 3.

Lastly, inferential statistical analysis was carried out to examine the relationship among the variables in the study. The variables were analysed using Spearman Correlation Coefficient and Kruskal Wallis test. Spearman Correlation Coefficient was adopted to test for relationship between the first and second digits' distribution of vote counts of all the 24 political parties for the 2007 presidential elections, PDP, CPC, ACN and ANPP for the 2011 presidential elections, andAPC and PDP for the 2015 presidential elections and the respective first and second digits' distribution of Benford's Law, as well as the relationship between the first and second digits' distribution of voters' turnout in the 2011 and 2015 presidential elections and the respective first and second digits' distribution of Benford's Law. Inferential statistics was not carried out on the last digits test because findings from the reviewed literatures indicate that last digits test is more appropriate for results at collated at the polling units or wards.Spearman Correlation Coefficient was also adopted to test for the relationship between the vote counts of PDP and CPC for the 2011 presidential elections as well asAPC and PDP for the 2015 presidential elections and the respective Monte Carlo's surrogate vote counts. Unlike the Spearman Correlation Coefficient tests for the Benford's distribution which tests for anomalies and conformity of the distribution of digits in the election data to the Benford's distribution, the Spearman Correlation Coefficient tests for the Monte Carlo Simulation only provide statistical evidence on whether the surrogate data represent (mimic) the actual data. In addition, Kruskal Wallis test was adopted to test for the difference in voters' turnout across the six geo-political zones in Nigeria.

Spearman Correlation Coefficient (a nonparametric alternative to linear regression) was adopted for this study because the independent variables (actual vote counts and voters' turnout) failed the normality test. The failed normality test also necessitated the adoption of Kruskal Wallis test as a nonparametric alternative to ANOVA. The tests of hypotheses were carried out using Statistical Package for the Social Sciences(SPSS).

In summary, the relationship between the research hypotheses, variables of measurement, instrument of data collection, data analysis tool and method of data analysis is presented in Table 3.1.

Table 3.1: Relationship between hypotheses, instruments and data analysis

S/N	Statement of Hypotheses	Variables of Measurement	Instrument of Data Collection	Data Analysis Tool	Method of Data Analysis
1	There is no significant correlation between the digital distribution of vote counts of presidential election results in Nigeria and the digital distribution of Benford's Law	First and second digits' distribution of vote counts, first and second digits' distribution of Benford's Law	directly from INEC office or through hyperlinks	SPSS/MS Excel	Spearman Correlation Coefficient
2	There is no significant correlation between the digital distribution of voters' turnoutofpresidential election results in Nigeria and the digital distribution of Benford's Law	First and second digits' distribution of voters' turnout, first and second digits' distribution of Benford's Law	directly from INEC office or through hyperlinks	SPSS/MS Excel	Spearman Correlation Coefficient
3	There is no significant relationship between the digital distribution of vote counts of presidential election results in Nigeria and digital distribution of Monte Carlo simulation	actual vote counts, surrogate vote counts	directly from INEC office or through hyperlinks	SPSS/MS Excel and R	Spearman Correlation Coefficient
4	There is no significant difference in the voters' turnout across the six geo-political zones in Nigeria	voters' turnout	directly from INEC office or through hyperlinks	SPSS/MS Excel	Kruskal Wallis test

3.7 Ethical Considerations for the Study

The presidential election results used for this study were adopted for research purpose only. The result of the analysis of this study was intended to give statistical evidence, and nothing more, on the applicability of Benford's Law and Monte Carlo simulation in detecting anomalies in Nigeria presidential election data. This research was not conducted to condemn the entire election process in Nigeria. It is expected that no part of this study would be shared or distributed to the public for the purpose of nullifying election results in Nigeria.

CHAPTER FOUR

DATA ANALYSIS AND PRESENTATION

4.0 Introduction

This chapter presents the findings on the result of the presidential elections held in Nigeria between 2007 and 2015. The chapter is divided into five main sections; the first section presents the digital distribution characteristics of election results, the second section presents the application of Benford's Law to election results, the third section presents the application of Monte Carlo simulation to election results, and the final section presents the test of hypotheses.

4.1 Digital Distribution Characteristics of Election Results

This section focuses on the digital distribution characteristics of the results of the presidential elections conducted in Nigeria between 2007 and 2015. The 2015 presidential election result was the most detailed of all the results released by INEC. This was followed by the 2011 presidential election result; while the 2007 presidential election result has little detail. The results are presented in the sub-sections below. Histograms were used to show the distribution of last digits. It is expected that last digits should be uniformly distributed in an error-free election (Beber and Scacco, 2008). Further analysis on the first and second digits was carried out in section 4.2.

4.1.1 Digital Distribution Characteristics of the 2015 Presidential Election Result

Table 4.1 presents the 2015 presidential election results (detailed result as released by INEC is in Appendix 4). The total number of registered voters is 67,422,005 while the total number of accredited voters is 31,746,490. The total number of registered voters is the same as the total number of eligible voters.

Table 4.1: Result of the 2015 presidential election

State	Registered Voters	Accredited voters	APC	PDP	OTHER PARTIES	Valid Votes	Rejected Votes	Total Votes Cast
ABIA	1349134	442538	13394	368303	9348	391045	10004	401049
ADAMAWA	1518123	709993	374701	251664	9653	636018	25192	661210
AKWA IBOM	1644481	1074070	58411	953304	5349	1017064	11487	1028551
ANAMBRA	1963427	774430	17926	660762	9896	688584	14825	703409
BAUCHI	2053484	1094069	931598	86085	2655	1020338	19437	1039775
BAYELSA	605637	384789	5194	361209	664	367067	4672	371739
BENUE	1893596	754634	373961	303737	5566	683264	19867	703131
BORNO	1799669	544759	473543	25640	2737	501920	13088	515008
CROSS RIVER	1144288	500577	28368	414863	7283	450514	15392	465906
DELTA	2044372	1350914	48910	1211405	7458	1267773	17075	1284848
EBONYI	1071226	425301	19518	323653	20717	363888	29449	393337
EDO	1650552	599166	208469	286869	5113	500451	22334	522785
EKITI	723255	323739	120331	176466	3894	300691	8754	309445
ENUGU	1381563	616112	14157	553003	6013	573173	12459	585632
GOMBE	1110105	515828	361245	96873	2481	460599	12845	473444
IMO	1747681	801712	133253	559185	10526	702964	28957	731921
JIGAWA	1815839	1153428	885988	142904	8672	1037564	34325	1071889
KADUNA	3361793	1746031	1127760	484085	5637	1617482	32719	1650201
KANO	4943862	2364434	1903999	215779	9043	2128821	43626	2172447
KATSINA	2842741	1578646	1345441	98937	5048	1449426	32288	1481714
KEBBI	1457763	792817	567883	100972	8148	677003	38119	715122
KOGI	1350883	476839	264851	149987	6490	421328	17959	439287
KWARA	1181032	489360	302146	132602	5332	440080	21321	461401
LAGOS	5827846	1678754	792460	632327	18899	1443686	52289	1495975
NASARAWA	1222054	562959	236838	273460	1249	511547	10094	521641
NIGER	1995679	933607	657678	149222	6771	813671	31012	844683
OGUN	1709409	594975	308290	207950	16932	533172	26441	559613
ONDO	1501549	618040	299889	251368	9799	561056	21379	582435
OSUN	1378113	683169	383603	249929	9083	642615	20758	663373
OYO	2344448	1073849	528620	303376	49356	881352	47254	928606
PLATEAU	1977211	1076833	429140	549615	3633	982388	18304	1000692
RIVERS	2324300	1643409	69238	1487075	9148	1565461	19307	1584768
SOKOTO	1663127	988899	671926	152199	10134	834259	42110	876369
TARABA	1374307	638578	261326	310800	7551	579677	23039	602716

YOBE	1077942	520127	446265	25526	2005	473796	17971	491767
ZAMFARA	1484941	875049	612202	144833	3987	761022	19157	780179
FCT	886573	344056	146399	157195	3211	306805	9210	316015

The total number of votes gathered by the APC is 15,424,921; total number of votes gathered by the PDP is 12,853,162; while the total number of votes gathered by the other parties (AA, ACPN, AD, ADC, APA, CPP, HOPE, KOWA, NCP, PPN, UDP and UPP) is 309,481. The total number of valid votes gathered by all the parties is 28,587,564 while the total number of rejected votes is 844,519.

Correlation matrix on the 2015 presidential election results was generated to know how the variables relate to each other (the details of the Correlation tests are available in Appendix 2). There is a direct and strong correlation between the registered voters and the accredited voters ($r = 0.859$, $p < 0.05$). But, the correlation between the votes cast for APC and PDP is inverse and low ($r = -0.350$, $p < 0.05$). There is also a direct and moderate correlation between valid votes and rejected votes ($r = 0.560$, $p < 0.05$) and total votes and rejected votes ($r = 0.578$, $p < 0.05$). The relationship between total votes and valid votes is direct and perfect ($r = 1.000$, $p < 0.05$).

As a result of the fewer number of digits in the votes gathered by the other twelve parties, the analysis of the vote cast in the 2015 presidential elections focused on the two leading parties (APC and PDP). The breakdown of the first digits in the vote count at the state level for APC reveals that nine states (Abia, Anambra, Ebonyi, Ekiti, Enugu, Imo, Kaduna, Kano, Katsina) and FCT have digits starting with 1; six states (Cross River, Edo, Kogi, Nasarawa, Ondo and Taraba) have digits starting with 2; six states (Adamawa, Benue, Gombe, Kwara, Ogun and Osun) have digits starting with 3; four states (Borno, Delta, Plateau and Yobe) have digits starting with 4; four states (Akwa Ibom, Bayelsa, Kebbi and Oyo) have digits starting with 5; four states (Niger, Rivers, Sokoto and Zamfara) have digits starting with 6; only Lagos state has digits starting with 7; Jigawa state has digits starting with 8; and Bauchi has digit starting with 9.

The breakdown of the second digits in the vote count at the state level for APC shows that three states (Edo, Kwara and Ogun) have second digits as 0; three states (Bayelsa, Kaduna and Zamfara) have second digits as 1; three states (Ekiti, Oyo and Plateau) have second digits as 2; five states (Abia, Bauchi, Imo, Katsina and Nasarawa) have second digits as 3; two states (Enugu, Yobe) and FCT have second digits as 4; only Niger State has second digit as 5; four states (Gombe, Kebbi, Kogi and Taraba) have second digits as 6; five states (Adamawa, Anambra, Benue, Borno and Sokoto) have second digits as 7; five states (Akwa Ibom, Cross

River, Delta, Jigawa and Osun) have second digits as 8; and another five states (Ebonyi, Kano, Lagos, Ondo and Rivers) have second digits as 9.

The breakdown of the last digits in the vote count at the state level for APC reveals that six states (Delta, Kaduna, Lagos, Ogun, Oyo and Plateau) have digits ending with 0; six states (Adamawa, Akwa Ibom, Benue, Ekiti, Katsina and Kogi) have digits ending with 1; only Zamfara state has digits ending with 2; four states (Borno, Imo, Kebbi and Osun) have digits ending with 3; two states (Abia and Bayelsa) have digits ending with 4; two states (Gombe and Yobe) have digits ending with 5; four states (Anambra, Kwara, Sokoto and Taraba) have digits ending with 6; only Enugu state has digits ending with 7; seven states (Bauchi, Cross River, Ebonyi, Jigawa, Nasarawa, Niger and Rivers) have digits ending with 8; and three states (Edo, Kano and Ondo) together with the FCT have digits ending with 9. The mean for all the votes gathered by the APC is 416,889.8 while the standard deviation is 413,798.6. Figure 4.1 shows that the frequencies of the last digits in the vote counts for APC at the state level, for the 2015 presidential elections, are not uniformly distributed.

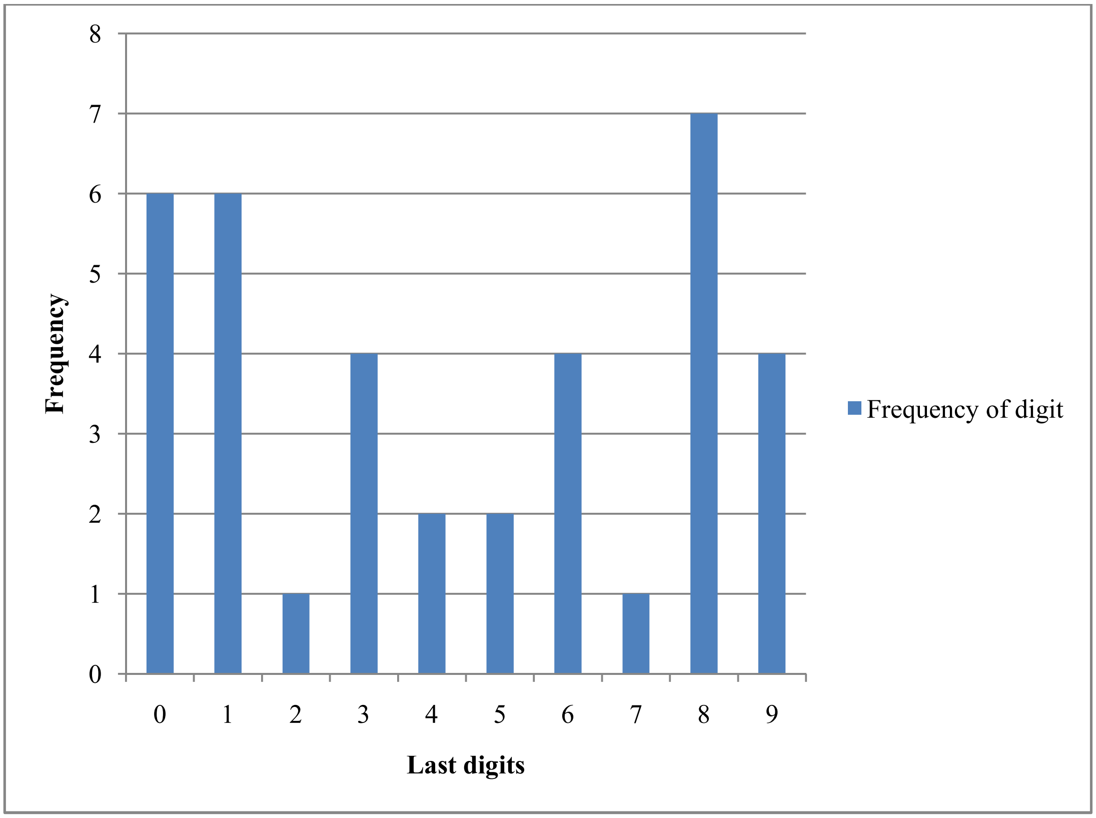

Figure 4.1: Last digits distribution for APC in 2015 presidential election

The breakdown of the first digits in the vote count at the state level for PDP shows that ten states (Delta, Ekiti, Jigawa, Kebbi, Kogi, Kwara, Niger, Rivers, Sokoto and Zamfara) and the FCT have digits starting with 1; nine states (Adamawa, Borno, Edo, Kano, Nasarawa, Ogun, Ondo, Osun and Yobe) have digits starting with 2; six states (Abia, Bayelsa, Benue, Ebonyi, Oyo and Taraba) have digits starting with 3; two states (Cross River and Kaduna) have digits starting with 4; three states (Enugu, Imo and Plateau) have digits starting with 5; two states (Anambra and Lagos) have digits starting with 6; there is no state with first digit as 7; only Bauchi state has digits starting with 8; and three states (Akwa Ibom, Gombe and Katsina) have digits starting with 9.

The breakdown of the second digits in the vote count at the state level for PDP reveals that four states (Benue, Kebbi, Ogun and Oyo) have second digits as 0; three states (Cross River, Kano and Taraba) have second digits as 1; two states (Delta and Ebonyi) have second digits as 2; two states (Kwara and Lagos) have second digits as 3; seven states (Jigawa, Kogi, Niger, Osun, Plateau, Rivers and Zamfara) have second digits as 4; eight states (Adamawa, Akwa Ibom, Borno, Enugu, Imo, Ondo, Sokoto and Yobe) and the FCT have second digits as 5; five states (Abia, Anambra, Bauchi, Bayelsa and Gombe) have second digits as 6; two states (Ekiti and Nasarawa) have second digits as 7; and three states (Edo, Kaduna and Katsina) have second digits as 8. There is no state with a second digit of 9.

The breakdown of the last digits in the vote count at the state level for PDP reveals that four states (Borno, Nasarawa, Ogun and Taraba) have digits ending with 0; there is no state with a last digit of 1; four states (Anambra, Kebbi, Kwara and Niger) have digits ending with 2; six states (Abia, Cross River, Ebonyi, Enugu, Gombe and Zamfara) have digits ending with 3; three states (Adamawa, Akwa Ibom and Jigawa) have digits ending with 4; six states (Bauchi, Delta, Imo, Kaduna, Plateau and Rivers) and the FCT have digits ending with 5; three states (Ekiti, Oyo and Yobe) have digits ending with 6; four states (Benue, Katsina, Kogi and Lagos) have digits ending with 7; only Ondo State has digits ending with 8; and five states (Bayelsa, Edo, Kano, Osun and Sokoto) have digits ending with 9. The mean for all the votes gathered by the PDP is 347382.8 while the standard deviation is 315993. Figure 4.2 shows that the distribution of the last digits in the vote counts for PDP at the state level, for the 2015 presidential elections, is not uniform.

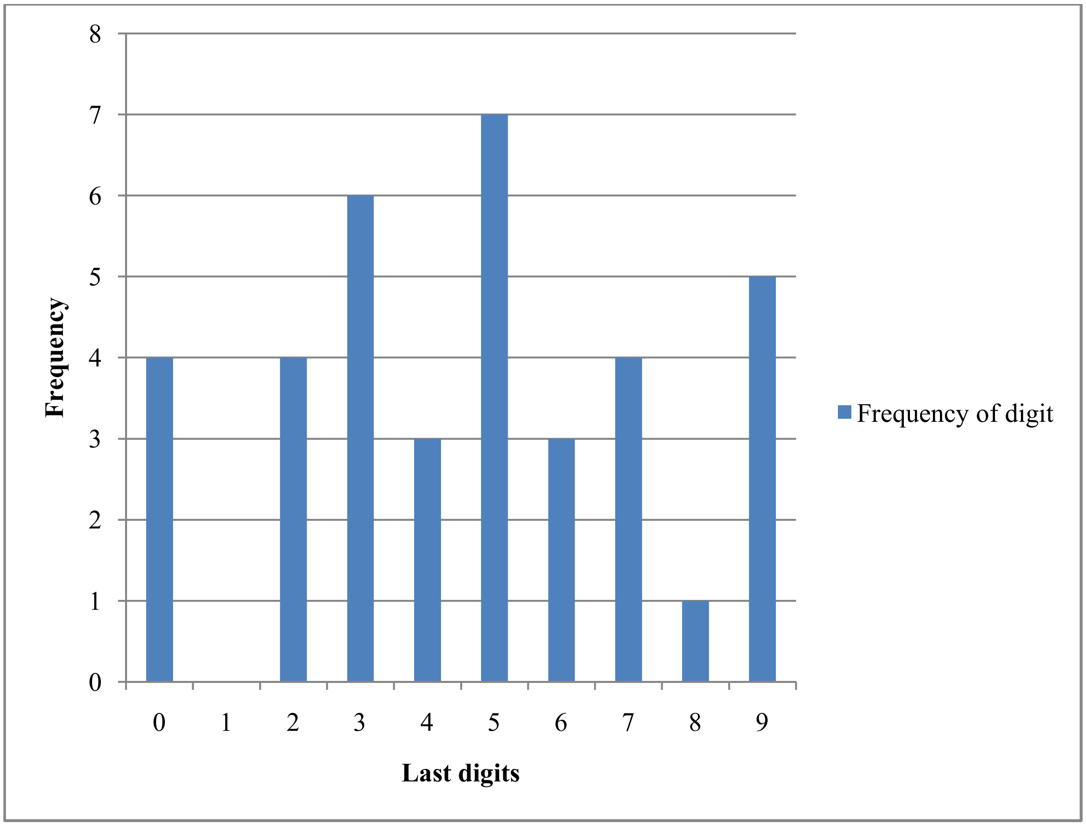

Figure 4.2: Last digits distribution for PDP in 2015 presidential election

Voters' turnout for each state was calculated as total vote cast (total number of valid votes and total number of rejected votes) per state divided by total number of registered voters per state. The breakdown of the first digits in voters' turnout at the state level shows that there is no state with a first digit of 1; three states (Abia, Borno and Lagos) have digits starting with 2; nine states (Anambra, Benue, Ebonyi, Edo, Kogi, Kwara, Ogun, Ondo and Oyo) and the FCT have digits starting with 3; fourteen states (Adamawa, Cross River, Ekiti, Enugu, Gombe, Imo, Kaduna, Kano, Kebbi, Nasarawa, Niger, Osun, Taraba and Yobe) have digits starting with 4; six states (Bauchi, Jigawa, Katsina, Plateau, Sokoto and Zamfara) have digits starting with 5; and four states (Akwa Ibom, Bayelsa, Delta and Rivers) have digits starting with 6. There is no state with a first digit of 7, 8 and 9.

The breakdown of the second digits in voters' turnout at the state level reveals that three states (Bauchi, Cross River and Plateau) have second digits as 0; three states (Bayelsa, Edo and Imo) have second digits as 1; twelve states (Akwa Ibom, Delta, Ekiti, Enugu, Gombe, Katsina, Kogi, Nasarawa, Niger, Ogun, Sokoto and Zamfara) have second digits as 2; three states (Adamawa, Kano and Taraba) have second digits as 3; there is no state with a second digit of 4; three states (Anambra, Lagos and Yobe) and the FCT have second digits as 5; only Ebonyi state and Benue state has a second digit of 6 and 7, respectively; four states (Borno, Ondo, Osun and Rivers) have second digits as 8; and six states (Abia, Jigawa, Kaduna, Kebbi, Kwara and Oyo) have second digits as 9.

The breakdown of the last digits in voters' turnout at the state level shows that there is no state with a last digit of 0; five states (Kwara, Ondo, Oyo, Kaduna and Sokoto) have digits ending with 1; four states (Lagos, Anambra, Benue and Imo) have digits ending with 2; ten states (Borno, Kogi, Ogun, Adamawa, Kano, Niger, Taraba, Yobe, Delta and Rivers) have digits ending with 3; four states (Osun, Bauchi, Zamfara and Bayelsa) have digits ending with 4; four states (Abia, Cross River, Gombe and Jigawa) have digits ending with 5; three states (Edo, Enugu and Kebbi) have digits ending with 6; only Ekiti State has last digit of 7; three states (Ebonyi, Nasarawa and Plateau) and the FCT have digits ending with 8; and two states (Katsina and Akwa Ibom) have digits ending with 9. Figure 4.3 shows that the frequencies of the last digits in voters' turnout at the state level, for the 2015 presidential elections, are not uniformly distributed.

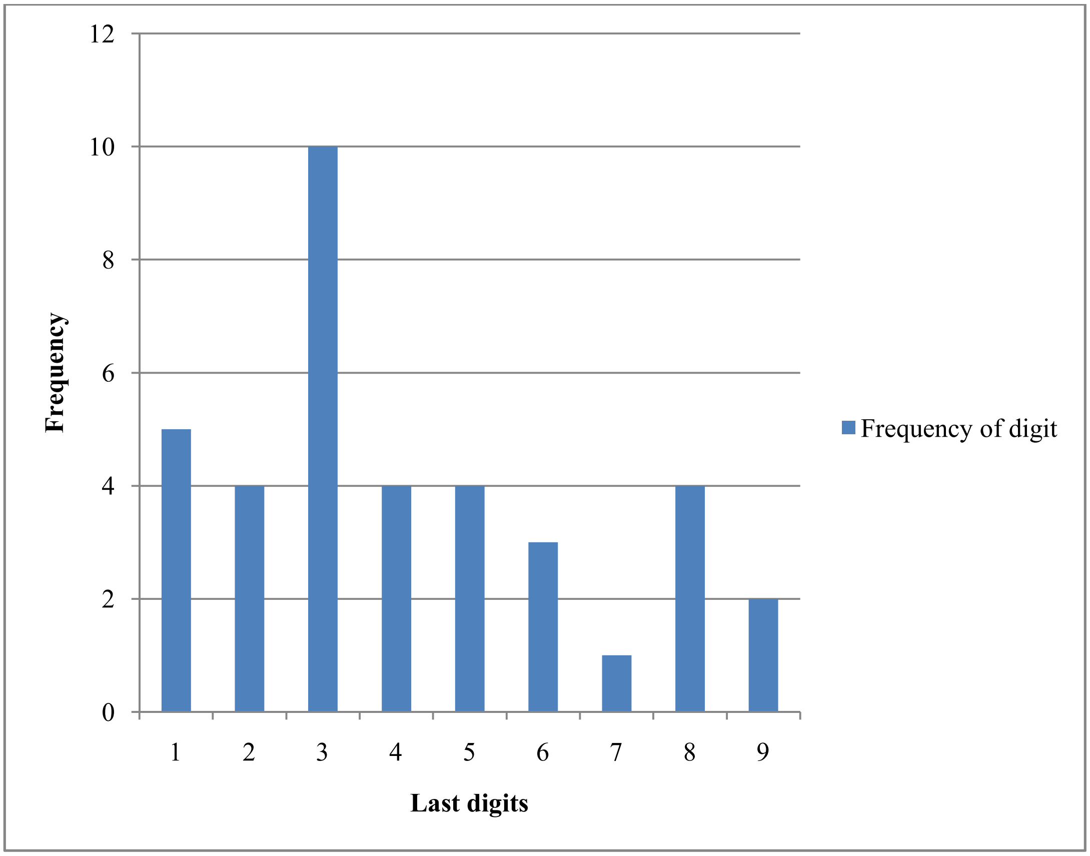

Figure 4.3: Last digits distribution for turnout in 2015 presidential election

The mean for all the total number of registered voters is 1,822,216 while the standard deviation is 1,026,219. The total vote cast (total number of valid votes and total number of rejected votes) is 29,432,083. The mean for all the total number of votes cast (valid and rejected) is 795,461.7 while the standard deviation is 431,062. The difference between the registered voters and the accredited voters at each state implies that not everyone that registered to vote came for accreditation of voters. Likewise, the difference between the accredited voters and the total vote cast (valid and rejected) for each state reveals that it was not everyone accredited that voted. The voters' turnout for each state is calculated as total vote cast divided by total number of registered voters. The percentage turnout of the 67,422,005 registered voters for the 2015 presidential election is 43.65%.

4.1.2 Digital Distribution Characteristics of the 2011 Presidential Election Result

The 2011 presidential election result is presented in Table 4.2 below. Unlike the 2015 election, there was no voters' accreditation in 2011. Likewise, there was no column for rejected votes in the 2011 result (the adopted result was extracted from http://www.inecnigeria.org/results/, as explained in section 3.5). The column for total (total votes cast) is a summation of the votes of PDP, CPC, ACN, ANPP and the other 16 parties (PDC, PMP, PPP, ADC, BNPP, FRESH, NCP, NMDP, APS, UNPD, NTP, MPPP, ARP, HDP, SDMP and LDPN). INEC made use of the final voters' registration (FINAL REG' in Table 4.3)and the total votes cast to calculate the voters' turnout. The final voters' registration has a total of 73,528,040 eligible voters.

Table 4.2: Result of the 2011 presidential election

STATES	PDP	CPC	ACN	ANPP	OTHERS	TOTAL	FINAL REG'	VOTER TURN OUT
ABIA	1175984	3743	4392	1455	2759	1188333	1524484	77.90%
ADAMAWA	508314	344526	32786	2706	19374	907706	1816094	50.00%
AKWA IBOM	1165629	5348	54148	2000	5270	1232395	1616873	76.20%
ANAMBRA	1145169	4223	3437	975	3435	1157239	2011746	57.50%
BAUCHI	258404	1315209	16674	8777	11030	1610094	2523614	63.80%
BAYELSA	504811	691	370	136	685	506693	591870	85.60%
BENUE	694776	109680	223007	8592	11654	1047709	2390884	43.80%
BORNO	207075	909763	7533	37279	15996	1177646	2380957	49.50%
C/RIVER	709382	4002	5889	2521	4547	726341	1148486	63.20%
DELTA	1378851	8960	13110	2746	6712	1410379	2032191	69.40%
EBONYI	480592	1025	1112	14296	5865	502890	1050534	47.90%
EDO	542173	17795	54242	2174	4808	621192	1655776	37.50%
EKITI	135009	2689	116981	1482	5697	261858	764726	34.20%
ENUGU	802144	3753	1755	1111	5246	814009	1303155	62.50%
FCT	253444	131576	2327	3170	7577	398094	943473	42.20%
GOMBE	290347	459898	3420	5693	10661	770019	1318377	58.40%
IMO	1381357	7591	14821	2520	3561	1409850	1687293	83.60%
JIGAWA	419252	663994	17355	7673	32492	1140766	2013974	56.60%
KADUNA	1190179	1334244	11278	17301	16961	2569963	3905387	65.80%
KANO	440666	1624543	42353	526310	39356	2673228	5027297	53.20%
KATSINA	428392	1163919	10945	6342	29934	1639532	3126898	52.40%
KEBBI	369198	501453	26171	3298	23979	924099	1638308	56.40%
KOGI	399816	132201	6516	16491	6758	561782	1316849	42.70%
KWARA	268243	83603	52432	1672	8804	414754	1152361	36.00%
LAGOS	1281688	189983	427203	8941	37229	1945044	6108069	31.80%
NASARAWA	408997	278390	1204	1047	4889	694527	1389308	50.00%
NIGER	321429	652574	13344	7138	24682	1019167	2175421	46.80%
OGUN	309177	17654	199555	2969	14360	543715	1941170	28.00%
ONDO	387376	11890	74253	6741	6577	486837	1616091	30.10%
OSUN	188409	6997	299711	3617	13980	512714	1293967	39.60%
OYO	484758	92396	252240	7156	26994	863544	2572140	33.60%
PLATEAU	1029865	356551	10181	5235	9285	1411117	2259194	62.50%
RIVERS	1817762	13182	16382	1449	5341	1854116	2429231	76.30%
SOKOTO	309057	540769	20144	5063	34775	909808	2267509	40.10%
TARABA	451354	257986	17791	1203	10731	739065	1336221	55.30%
YOBE	117128	337537	6069	143179	18202	622115	1373796	45.30%
ZAMFARA	238980	624515	17970	46554	14660	942679	1824316	51.70%

Source: INEC website (as published by Nigerian Muse, April 19, 2011)

Correlation matrix was also generated for the votes cast in the 2011 presidential election results (the details of the correlation tests are available on Appendix 2). There is no significant correlation ($p > 0.05$) between the votes cast for PDP and CPC; PDP and ACN; PDP and ANPP; and PDP and other parties. There is also no significant correlation ($p > 0.05$) between the votes cast for CPC and ACN; and ACN and ANPP. However, there is a direct and moderate correlation between the votes cast for CPC and ANPP ($r = 0.526$, $p < 0.05$); CPC and other parties ($r = 0.601$, $p < 0.05$); and ANPP and other parties ($r = 0.433$, $p < 0.05$). There is also a direct and low correlation between the votes cast for ACN and other parties ($r = 0.336$, $p < 0.05$).

The total number of votes gathered by the PDP is 22,495,187; total votes gathered by CPC is 12,214,853; total votes gathered by ACN is 2,079,101; total votes gathered by ANPP is 917,012; while the total votes gathered by the other 16 parties is 504,866. The total vote gathered by all the parties is 38,211,019 while the total number of registered voters is 73,528,040. The analysis of the votes cast in the 2011 presidential election focused on the four leading parties (PDP, CPC, ACN and ANPP).

The breakdown of the first digits in the vote count at the state level for PDP reveals that twelve states (Ekiti, Lagos, Osun, Abia, Anambra, Imo, Akwa Ibom, Delta, Rivers, Kaduna, Yobe and Plateau) have digits starting with 1; five states (Zamfara, Bauchi, Borno, Gombe and Kwara) and the FCT have digits starting with 2; six states (Ogun, Ondo, Kebbi, Sokoto, Kogi and Niger) have digits starting with 3; seven states (Oyo, Ebonyi, Jigawa, Kano, Katsina, Taraba and Nasarawa) have digits starting with 4; three states (Bayelsa, Edo and Adamawa) have digits starting with 5; only Benue State, Cross River State and Enugu State have digits starting with 6, 7 and 8 respectively.

The breakdown of the second digits in the vote count at the state level for PDP shows that nine states (Ogun, Enugu, Bayelsa, Cross River, Sokoto, Adamawa, Borno, Nasarawa and Plateau) have second digits as 0; six states (Abia, Anambra, Akwa Ibom, Jigawa, Kaduna and Yobe) have second digits as 1; three states (Lagos, Katsina and Niger) have second digits as 2; four states (Ekiti, Imo, Delta and Zamfara) have second digits as 3; two states (Edo and Kano) have second digits as 4; two states (Bauchi and Taraba) and the FCT have second digits as 5; two states (Kebbi and Kwara) have second digits as 6; there is no state with a second digit of 7; five states (Ondo, Osun, Oyo, Ebonyi and Rivers) have second digits as 8; and three states (Gombe, Benue and Kogi) have second digits as 9.

The breakdown of the last digits in the vote count at the state level for PDP reveals that only Zamfara State has digits ending with 0; two states (Bayelsa and Delta) have digits ending with 1; five states (Ebonyi, Cross River, Rivers, Jigawa and Katsina) have digits ending with 2; two states (Edo and Kwara) have digits ending with 3; five states (Abia, Enugu, Adamawa, Bauchi and Taraba) and the FCT have digits ending with 4; two states (Borno and Plateau) have digits ending with 5; four states (Ondo, Kano, Benue and Kogi) have digits ending with 6; five states (Ogun, Imo, Sokoto, Gombe and Nasarawa) have digits ending with 7; four states (Lagos, Oyo, Kebbi and Yobe) have digits ending with 8; and six states (Ekiti, Osun, Anambra, Akwa Ibom, Kaduna and Niger) have digits ending with 9. The mean for all the votes gathered by the PDP is 607,978 while the standard deviation is 429,220.5. Figure 4.4 implies that there is no uniform distribution in the last digits of vote counts, at the state level, for PDP in the 2011 presidential election.

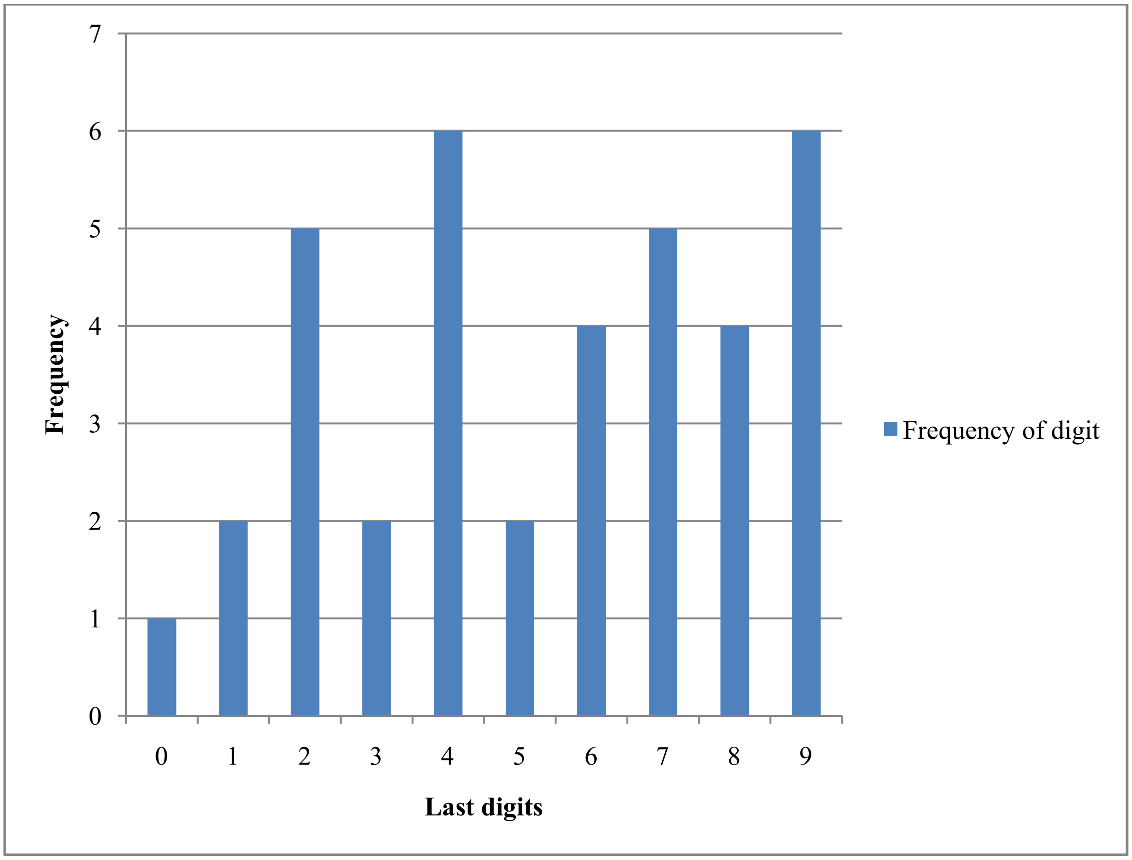

Figure 4.4: Last digits distribution for PDP in 2011 presidential election

The breakdown of the first digits in the vote count at the state level for CPC shows that twelve states (Lagos, Ogun, Ondo, Ebonyi, Edo, Rivers, Kaduna, Kano, Katsina, Bauchi, Benue and Kogi) and the FCT have digits starting with 1; three states (Ekiti, Taraba and Nasarawa) have digits starting with 2; five states (Abia, Enugu, Adamawa, Yobe and Plateau) have digits starting with 3; three states (Anambra, Cross River and Gombe) have digits starting with 4; three states (Akwa Ibom, Kebbi and Sokoto) have digits starting with 5; five states (Osun, Bayelsa, Jigawa, Zamfara and Niger) have digits starting with 6; only Imo State has digits starting with 7; two states (Delta and Kwara) have digits starting with 8; and two states (Oyo and Borno) have digits starting with 9.

The breakdown of the second digits in the vote count at the state level for CPC reveals that five states (Ebonyi, Cross River, Kebbi, Borno and Benue) have second digits as 0; two states (Ondo and Katsina) have second digits as 1; three states (Oyo, Anambra and Zamfara) have second digits as 2; seven states (Akwa Ibom, Rivers, Kaduna, Bauchi, Yobe, Kogi and Kwara) and the FCT have second digits as 3; two states (Sokoto and Adamawa) have second digits as 4; five states (Imo, Gombe, Taraba, Niger and Plateau) have second digits as 5; three states (Ekiti, Jigawa and Kano) have second digits as 6; five states (Ogun, Abia, Enugu, Edo and Nasarawa) have second digits as 7; only Lagos State has second digit as 8; and three states (Osun, Bayelsa and Delta) have second digits as 9.

The breakdown of the last digits in the vote count at the state level for CPC shows that four states (Ondo, Delta, Benue and Nasarawa) have digits ending with 0; four states (Imo, Bayelsa, Kogi and Plateau) have digits ending with 1; two states (Cross River and Rivers) have digits ending with 2; eight states (Lagos, Abia, Anambra, Enugu, Kano, Kebbi, Borno and Kwara) have digits ending with 3; four states (Ogun, Jigawa, Kaduna and Niger) have digits ending with 4; three states (Ebonyi, Edo and Zamfara) have digits ending with 5; three states (Oyo, Adamawa and Taraba) and the FCT have digits ending with 6; two states (Osun and Yobe) have digits ending with 7; two states (Akwa Ibom and Gombe) have digits ending with 8; and four states (Ekiti, Katsina, Sokoto and Bauchi) have digits ending with 9. The mean for all the votes gathered by the CPC is 330,131.2 while the standard deviation is 438,261.4. Figure 4.5 shows that the frequencies of the last digits in the vote counts for CPC at the state level, for the 2011 presidential elections, are not uniformly distributed.

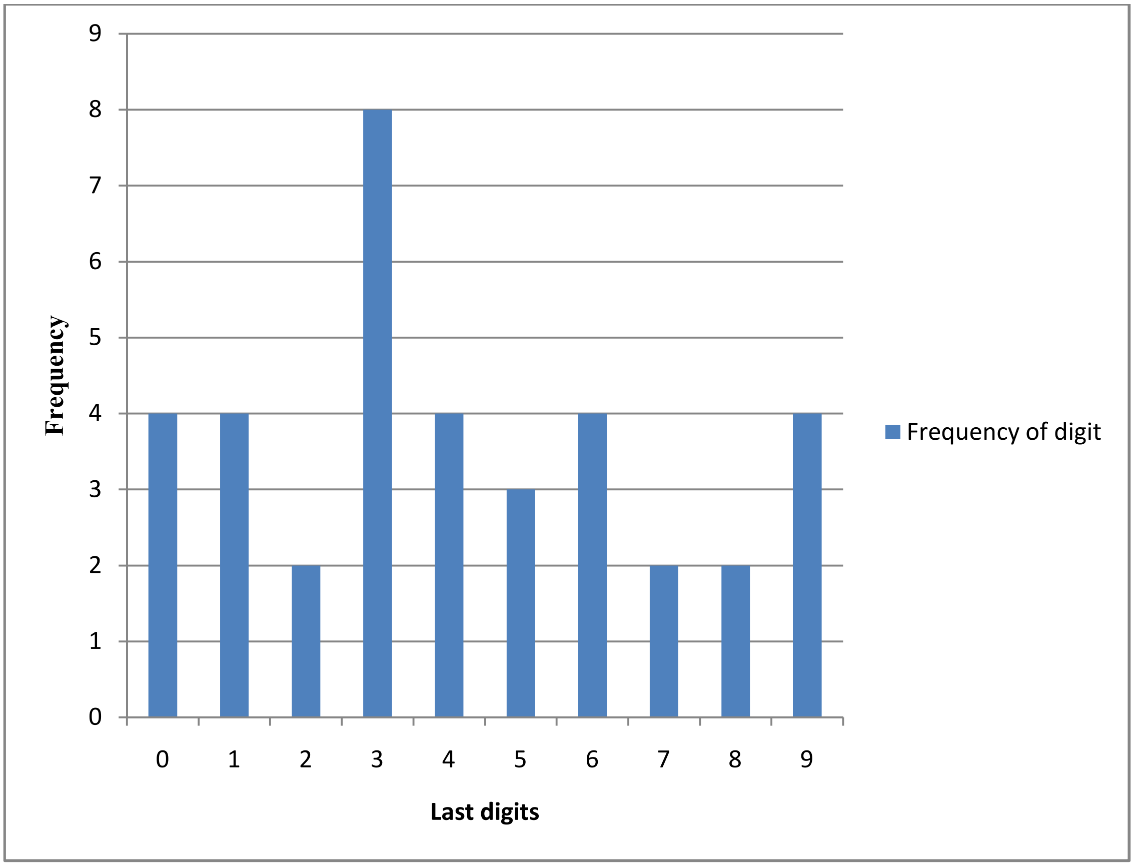

Figure 4.5: Last digits distribution for CPC in 2011 presidential election

The breakdown of the first digits in the vote count at the state level for ACN reveals that sixteen states (Ekiti, Ogun, Ebonyi, Enugu, Imo, Delta, Rivers, Jigawa, Kaduna, Katsina, Zamfara, Bauchi, Taraba, Nasarawa, Niger and Plateau) have digits starting with 1; five states (Osun, Oyo, Kebbi, Sokoto and Benue) and the FCT have digits starting with 2; four states (Anambra, Bayelsa, Adamawa and Gombe) have digits starting with 3; three states (Lagos, Abia and Kano) have digits starting with 4; four states (Akwa Ibom, Cross River, Edo and Kwara) have digits starting with 5; two states (Yobe and Kogi) have digits starting with 6; and two states (Ondo and Borno) have digits starting with 7. There is no state with a first digit of 8 and 9.

The breakdown of the second digits in the vote count at the state level for ACN shows that four states (Katsina, Sokoto, Yobe and Plateau) have second digits as 0; three states (Ekiti, Ebonyi and Kaduna) have second digits as 1; six states (Lagos, Kano, Adamawa, Benue, Kwara and Nasarawa) have second digits as 2; three states (Abia, Delta and Niger) and the FCT have second digits as 3; six states (Ondo, Anambra, Imo, Akwa Ibom, Edo and Gombe) have second digits as 4; three states (Oyo, Borno and Kogi) have second digits as 5; three states (Rivers, Kebbi and Bauchi) have second digits as 6; five states (Enugu, Bayelsa, Jigawa, Zamfara and Taraba) have second digits as 7; only Cross River State has a second digit of 8; and two states (Ogun and Osun) have second digits as 9.

The breakdown of the last digits in the vote count at the state level for ACN reveals that five states (Oyo, Bayelsa, Delta, Zamfara and Gombe) have digits ending with 0; six states (Ekiti, Osun, Imo, Kebbi, Taraba and Plateau) have digits ending with 1; five states (Abia, Ebonyi, Edo, Rivers and Kwara) have digits ending with 2; four states (Lagos, Ondo, Kano and Borno) have digits ending with 3; four states (Sokoto, Bauchi, Nasarawa and Niger) have digits ending with 4; four states (Ogun, Enugu, Jigawa and Katsina) have digits ending with 5; two states (Adamawa and Kogi) have digits ending with 6; two states (Anambra and Benue) and the FCT have digits ending with 7; two states (Akwa Ibom and Kaduna) have digits ending with 8; and two states (Cross River and Yobe) have digits ending with 9. The mean for all the votes gathered by the ACN is 56,191.92 while the standard deviation is 97,623.52. Figure 4.6 shows that the distribution of the last digits in the vote counts for ACN at the state level, for the 2011 presidential elections, is not uniform.

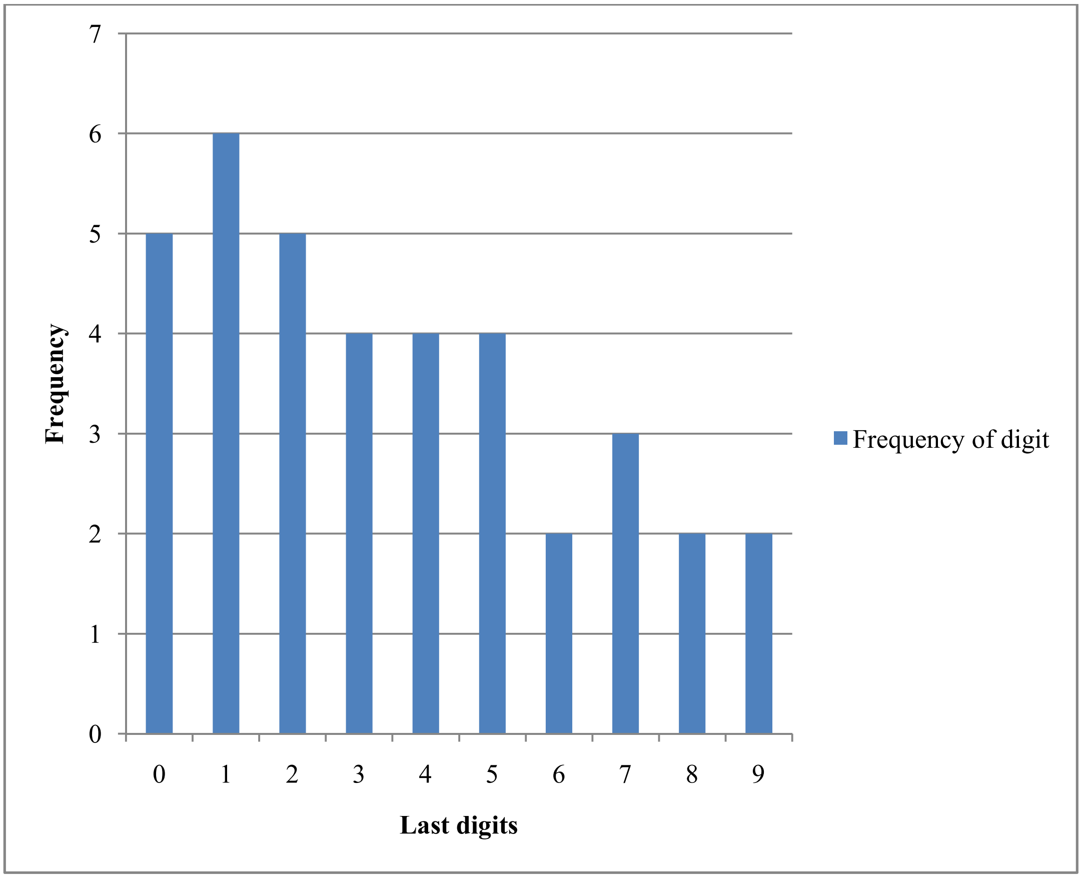

Figure 4.6: Last digits distribution for ACN in 2011 presidential election

The breakdown of the first digits in the vote count at the state level for ANPP shows that twelve states (Ekiti, Abia, Ebonyi, Enugu, Bayelsa, Rivers, Kaduna, Taraba, Yobe, Kogi, Kwara and Nasarawa) have digits starting with 1; seven states (Ogun, Imo, Akwa Ibom, Cross River, Delta, Edo and Adamawa) have digits starting with 2; three states (Osun, Kebbi and Borno) and the FCT have digits starting with 3; only Zamfara State has first digit as 4; four states (Kano, Sokoto, Gombe and Plateau) have digits starting with 5; two states (Ondo and Katsina) have digits starting with 6; three states (Oyo, Jigawa and Niger) have digits starting with 7; three states (Lagos, Bauchi and Benue) have digits starting with 8; and only Anambra State has digits starting with 9.

The breakdown of the second digits in the vote count at the state level for ANPP reveals that three states (Akwa Ibom, Sokoto and Nasarawa) have second digits as 0; four states (Oyo, Enugu, Edo and Niger) and the FCT have second digits as 1; four states (Kano, Kebbi, Taraba and Plateau) have second digits as 2; two states (Bayelsa and Katsina) have second digits as 3; five states (Ekiti, Abia, Ebonyi, Rivers and Yobe) have second digits as 4; three states (Imo, Cross River and Benue) have second digits as 5; six states (Osun, Jigawa, Zamfara, Gombe, Kogi and Kwara) have second digits as 6; seven states (Ondo, Anambra, Delta, Kaduna, Adamawa, Bauchi and Borno) have second digits as 7; there is no state with a second digit of 8; and two states (Lagos and Ogun) have second digits as 9.

The breakdown of the last digits in the vote count at the state level for ANPP shows that three states (Imo, Akwa Ibom and Kano) and the FCT have digits ending with 0; six states (Lagos, Ondo, Enugu, Cross River, Kaduna and Kogi) have digits ending with 1; four states (Ekiti, Katsina, Benue and Kwara) have digits ending with 2; four states (Jigawa, Sokoto, Gombe and Taraba) have digits ending with 3; two states (Edo and Zamfara) have digits ending with 4; three states (Abia, Anambra and Plateau) have digits ending with 5; five states (Oyo, Ebonyi, Bayelsa, Delta and Adamawa) have digits ending with 6; three states (Osun, Bauchi and Nasarawa) have digits ending with 7; two states (Kebbi and Niger) have digits ending with 8; and four states (Ogun, Rivers, Borno and Yobe) have digits ending with 9. The mean for all the votes gathered by the ANPP is 24,784.11 while the standard deviation is 88,154.50. The last digits of the vote counts for ANPP, at the state level, for the 2011 presidential election, are not uniformly distributed (Figure 4.7).

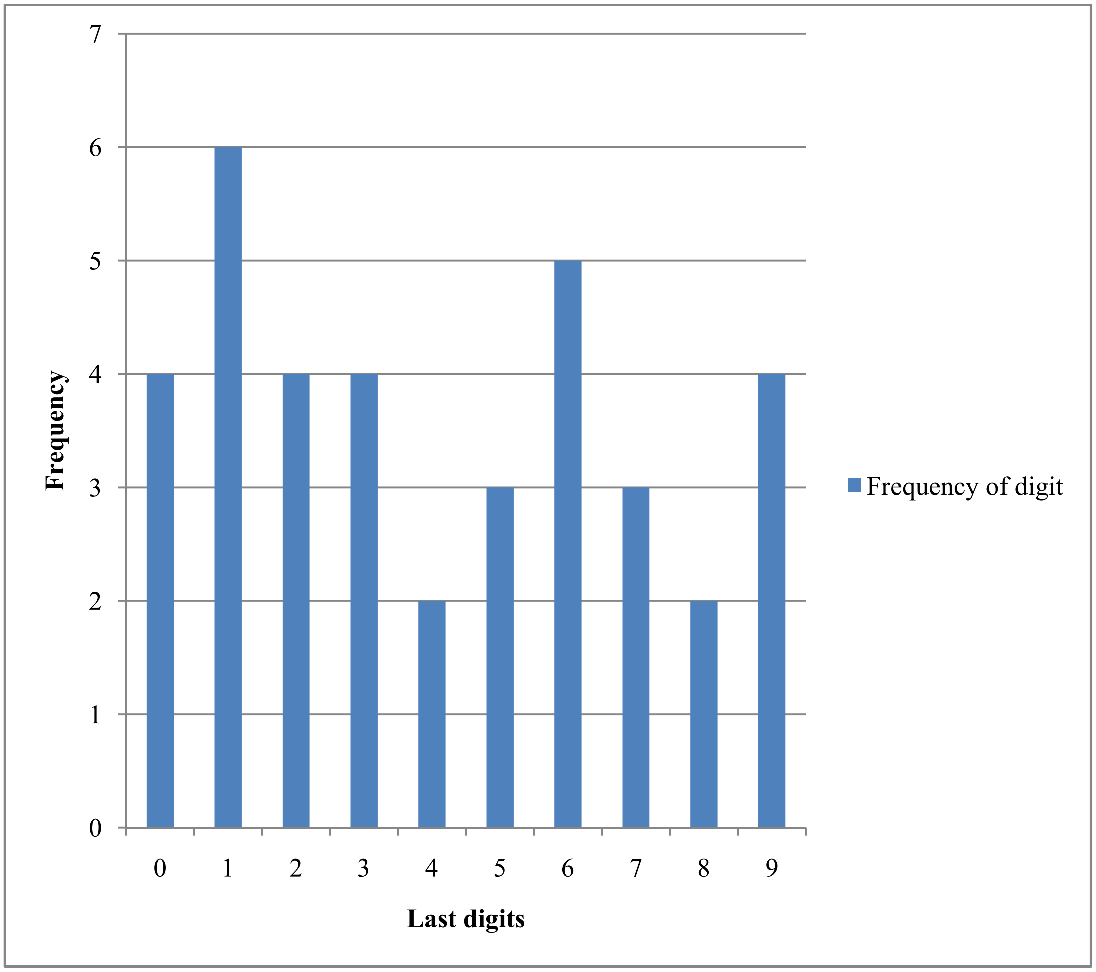

Figure 4.7: Last digits distribution for ANPP in 2011 presidential election

Voters' turnout for each state was calculated as total votes cast per state divided by final voters' registration per state. The breakdown of the first digits in the voters' turnout at the state level reveals that there is no state with a first digit of 1; only Ogun State has a first digit of 2; seven states (Ekiti, Lagos, Ondo, Osun, Oyo, Edo and Kwara) have digits starting with 3; nine states (Ebonyi, Sokoto, Adamawa, Borno, Yobe, Benue, Kogi, Nasarawa and Niger) and the FCT have digits starting with 4; eight states (Anambra, Jigawa, Kano, Katsina, Kebbi, Zamfara, Gombe and Taraba) have digits starting with 5; six states (Enugu, Cross River, Delta, Kaduna, Bauchi and Plateau) have digits starting with 6; three states (Abia, Akwa Ibom and Rivers) have digits starting with 7; and two states (Imo and Bayelsa) have digits starting with 8. There is no state with a first digit of 9.

The breakdown of the second digits in the voters' turnout at the state level shows that two states (Ondo and Sokoto) have second digits as 0; two states (Lagos and Zamfara) have second digits as 1; four states (Enugu, Katsina, Kogi and Plateau) and the FCT have second digits as 2; six states (Oyo, Imo, Cross River, Kano, Bauchi and Benue) have second digits as 3; only Ekiti State has a second digit of 4; five states (Bayelsa, Kaduna, Taraba, Yobe and Kwara) have second digits as 5; five states (Akwa Ibom, Rivers, Jigawa, Kebbi and Niger) have second digits as 6; four states (Abia, Anambra, Ebonyi and Edo) have second digits as 7; two states (Ogun and Gombe) have second digits as 8; and five states (Osun, Delta, Adamawa, Borno and Nasarawa) have second digits as 9.

The breakdown of the last digits in the voters' turnout at the state level reveals that there is no state with a last digit of 0; seven states (Jigawa, Kebbi Borno, Yobe, Kogi Nasarawa and Niger) have digits ending with 1; only Lagos State has a last digit of 2; three states (Osun, Kaduna and Bauchi) have digits ending with 3; six states (Ondo, Abia, Ebonyi, Edo, Kano and Zamfara) have digits ending with 4; three states (Akwa Ibom, Cross River and Kwara) have digits ending with 5; four states (Enugu, Delta, Katsina and Plateau) and the FCT have digits ending with 6; five states (Anambra, Bayelsa, Sokoto, Taraba and Benue) have digits ending with 7; six states (Ekiti, Ogun, Oyo, Imo, Rivers and Gombe) have digits ending with 8; and only Adamawa State has a last digit of 9. Figure 4.8 shows that the frequencies of the last digits in the voters' turnoutat the state level, for the 2011 presidential elections, are not uniformly distributed.

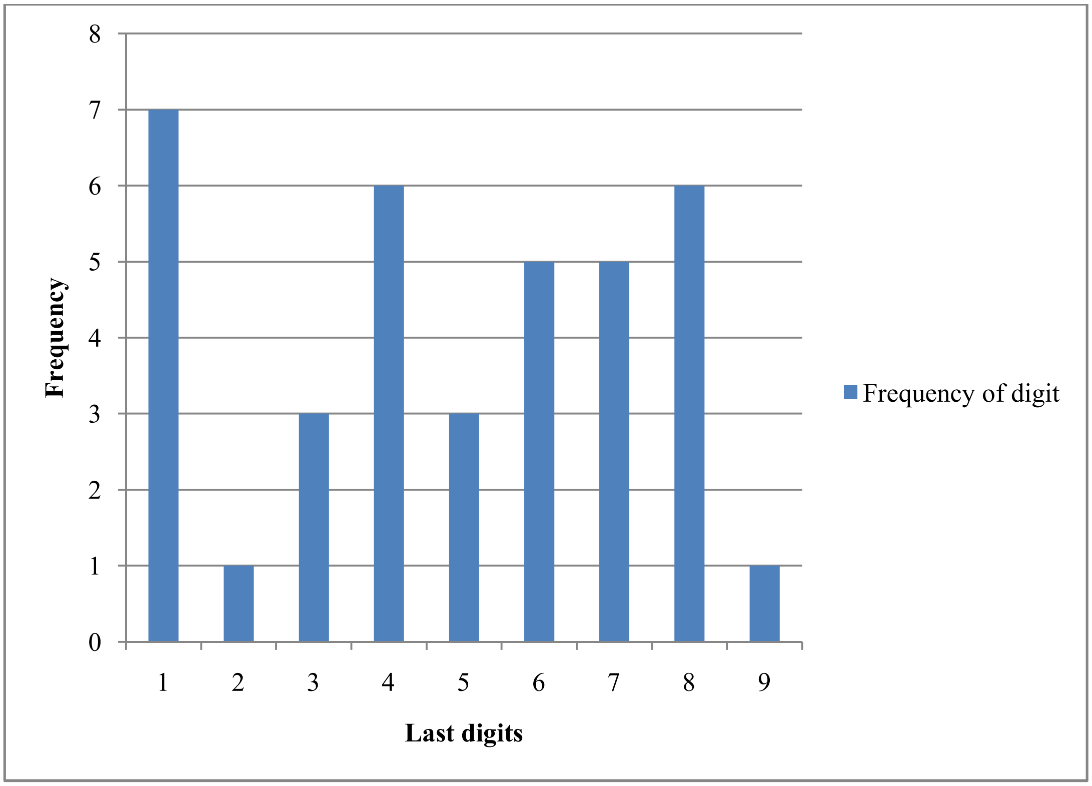

Figure 4.8: Last digits distribution for voters' turnout in 2011 presidential election

4.1.3 Digital Distribution Characteristics of the 2007 Presidential Election Result

The result of the 2007 presidential election (Table 4.3) presented by INEC Chairman neither has the figure of the total votes cast nationwide, nor the percentage scored by each of the candidates. There was also no state-by-state breakdown of the results (NDI, 2007). Therefore, the analysis only focused on the total votes cast for each of the political parties. The total number of registered voters or eligible voters is 61,567,036.

Table 4.3: Result of the 2007 presidential election

Party	Votes
People's Democratic Party (PDP)	24,784,227
All Nigeria People's Party (ANPP)	6,607,419
Action Congress (AC)	2,567,798
Progressive People's Alliance (PPA)	608,833
Democratic People's Party (DPP)	289,324
All Progressive Grand Alliance (APGA)	155,947
Alliance for Democracy (AD)	89,511
Fresh Democratic Party (FDP)	74,049
Nigeria People's Congress (NPC)	33,771
Hope Democratic Party (HDP)	28,518
Peoples Mandate Party (PMP)	24,164
African Liberation Party (ALP)	22,592
African Political System (APS)	22,459
National Democratic Party (NDP)	21,974
New Nigeria Peoples Party (NNPP)	21,665
Citizens Popular Party (CPP)	14,027
Republican Party of Nigeria (RPN)	13,566
Better Nigeria Progressive Party (BNPP)	11,705
National Conscience Party (NCP)	8,229
National Action Council (NAC)	5,692
National Majority Democratic Party (NMDP)	5,666
New Democrats (ND)	5,408
National Unity Party (NUP)	4,355
Masses Movement of Nigeria (MMN)	4,309

Source: INEC website (as published by National Democratic Institute, April 2008)

The breakdown of the first digits in the total votes count for each of the presidential candidates shows that four candidates (APGA, CPP, RPN and BNPP) have digits starting with 1; nine candidates (PDP, AC, DPP, HDP, PMP, ALP, APS, NDP and NNPP) have digits starting with 2; only NPC has first digit as 3; two candidates (NUP and MMN) have digits starting with 4; three candidates (NAC, NMDP and ND) have digits starting with 5; two candidates (ANPP and PPA) have digits starting with 6; only FDP has first digit as 7; and two candidates (AD and NCP) have digits starting with 8. There is no candidate with a first digit of 9.

The breakdown of the second digits in the total votes count for each of the presidential candidates reveals that only PPA has a second digit of 0; three candidates (NDP, NNPP and BNPP) have second digits as 1; three candidates (ALP, APS and NCP) have second digits as 2; four candidates (NPC, RPN, NUP and MMN) have second digits as 3; five candidates (PDP, FDP, PMP, CPP and ND) have second digits as 4; two candidates (AC and APGA) have second digits as 5; three candidates (ANPP, NAC and NMDP) have second digits as 6; there is no candidate with second digit of 7; two candidates (DPP and HDP) have second digits as 8; and only AD has a second digit of 9.

The breakdown of the last digits in the total votes count for each of the presidential candidates shows that there is no candidate with a last digit of 0; two candidates (AD and NPC) have digits ending with 1; two candidates (ALP and NAC) have digits ending with 2; only PPA has last digit as 3; three candidates (DPP, PMP and NDP) have digits ending with 4; three candidates (NNPP, BNPP and NUP) have digits ending with 5; two candidates (RPN and NMDP) have digits ending with 6; three candidates (PDP, APGA and CPP) have digits ending with 7; three candidates (AC, HDP and ND) have digits ending with 8; and five candidates (ANPP, FDP, APS, NCP and MMN) have digits ending with 9. The total vote gathered by all the candidates is 35,425,208.The mean value for all the votes gathered by the candidates is 1,476,050 while the standard deviation is 5,161,385. Figure 4.9 shows that the last digits of the results for the 2007 presidential elections are not uniformly distributed.

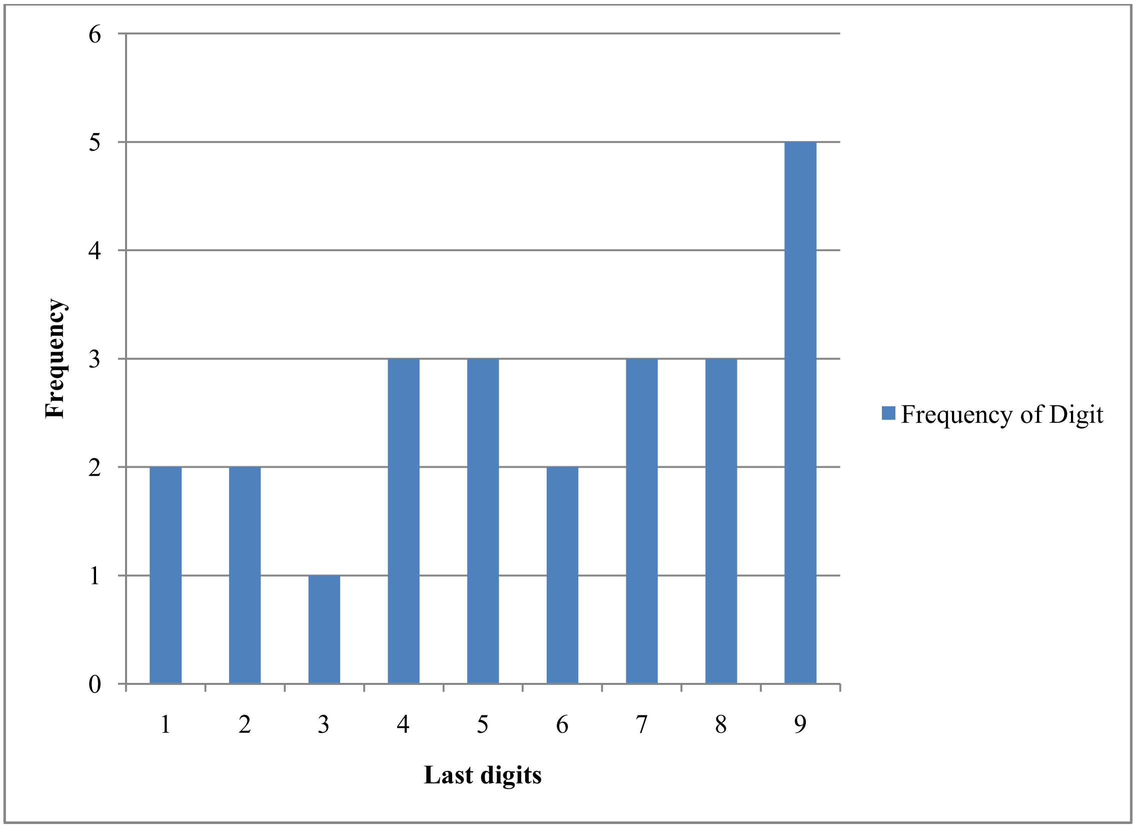

Figure 4.9: Last digits distribution for 2007 presidential election

4.2 Modelling Patterns ofthe Election Results using Benford's Law

This section focuses on modelling the distribution pattern of the first and second digits of the 2007, 2011 and 2015 presidential election results. However, the 2007 presidential election result was used to model only the total vote counts of all the 24 political parties, because the result released by INEC does not include state-by-state breakdown of the total votes or voters' turnout. The analysis was carried out, for the 2011 and 2015 presidential elections, by using the vote countsof the leading political parties of the presidential election results and the voters' turnout. The Benford's values for digits distribution were compared with the actual value of digits distribution in the election results. The actual value is calculated as frequency of each digit divided by 37 (36 states and the FCT). The Benford's values for the first digits (1 to 9) are: 0.301, 0.176, 0.125, 0.097, 0.079, 0.067, 0.058, 0.051 and 0.046 respectively. The Benford's values for the second digits (0 to 9) are: 0.120, 0.114, 0.109, 0.104, 0.100, 0.097, 0.093, 0.090, 0.088 and 0.085 respectively. More details on the Benford's distribution of the first and second digits as well as the underlying assumptions of Benford's Law are in section 2.2.

4.2.1 Benford's Law and the 2015 Presidential Elections

The focus, in this sub-section, is the application of Benford's first digit distribution and second digit distribution to the vote counts of APC and PDP in the 2015 presidential election. Benford's Lawwas also applied to the first and second digits distribution of the voters' turnout. The inferential statistics are in section 4.4 (Spearman Rank Correlation was used in comparing the actual and Benford's values).

(i) First digits distribution

Table 4.4 shows the frequency of distribution of the first digits of vote counts for APC in the 2015 presidential election. The table also shows the actual value and the respective Benford's value for the frequency distribution of each digit. Most of the digits deviate, slightly, from the Benford's value for first digit distribution. The closest to the Benford's first digits distribution is the distribution of digit 4.

Table 4.4: First digits distribution for APC in 2015

	First Digits	Frequency	Actual value	Benford's value
	1	10	0.270	0.301
	2	6	0.162	0.176
	3	6	0.162	0.125
	4	4	0.108	0.097
	5	4	0.108	0.079
	6	4	0.108	0.067
	7	1	0.027	0.058
	8	1	0.027	0.051
	9	1	0.027	0.046
Sum		**37**		
Mean			**3.348**	**3.441**

The distribution of the first digits of vote counts for PDP in the 2015 presidential election is shown in Table 4.5 (below). The mean value for the distribution of vote counts for PDP (3.159) in Table 4.5 deviates more from the Benford's value (3.441) than that of APC (3.348) in Table 4.4. Also, Table 4.5 shows that digit 1 and 5 have the closest values to the Benford's values for the distribution of first digits.

Table 4.5: First digits distribution for PDP in 2015

	First Digits	Frequency	Actual value	Benford's value
	1	11	0.297	0.301
	2	9	0.243	0.176
	3	6	0.162	0.125
	4	2	0.054	0.097
	5	3	0.081	0.079
	6	2	0.054	0.067
	7	0	0.000	0.058
	8	1	0.027	0.051
	9	3	0.081	0.046
Sum		**37**		
Mean			**3.159**	**3.441**

In summary, the distribution of first digits of the vote counts for both the APC and PDP in the 2015 presidential election result is shown in Figure 4.10. The Figure 4.10 shows a skew in favor of the low digits that approximated a geometric pattern. However, the distribution shows that none of the vote counts for both parties conform to the pattern predicted by Benford’s Law.

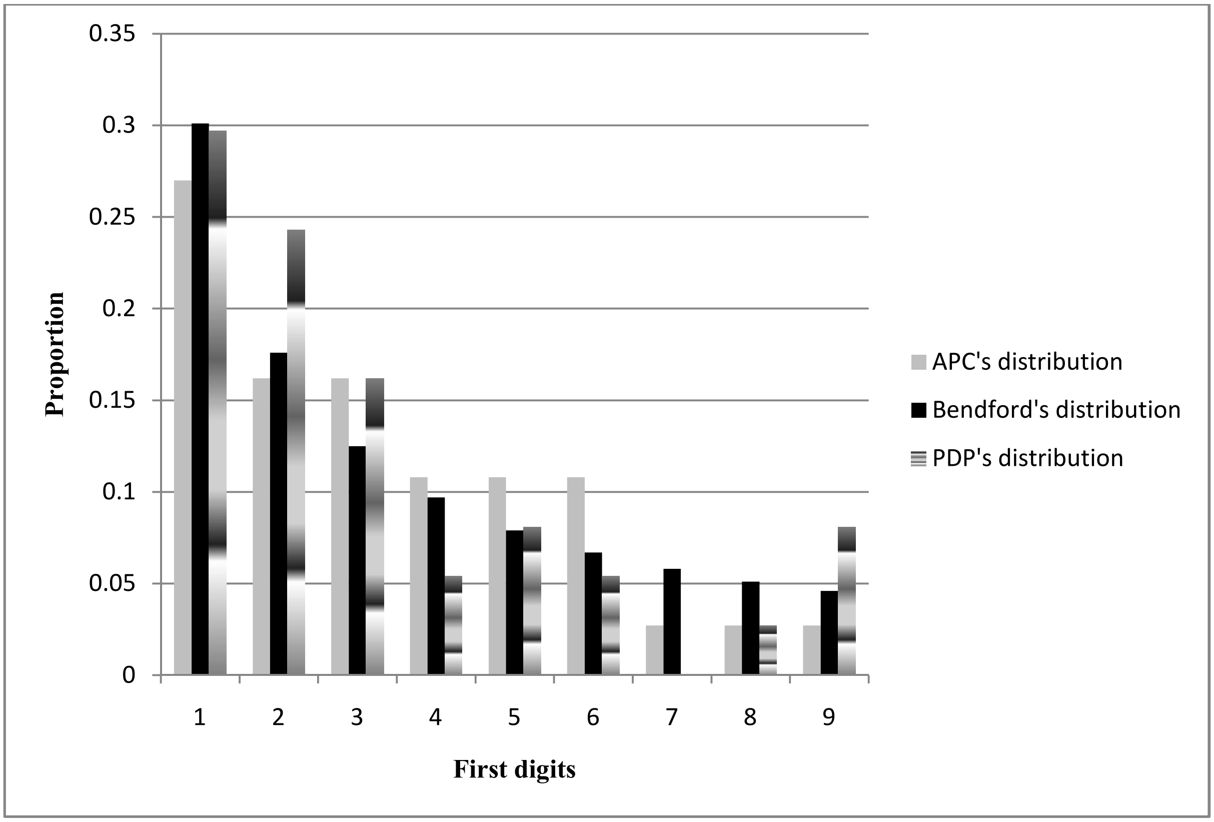

Figure 4.10: First digits distribution between Benford's value and parties' votes in 2015

Table 4.6 shows that the sequence of first digits in the voters' turnout for the 2015 presidential election is different from Benford's distribution. The absence of digits 1, 7, 8 and 9 as well as the excess distribution of digits 3 and 4 resulted in a wide deviation from Benford's values for the first digits distribution. The mean value (3.946) of voters' turnout also deviates from the expected mean of 3.441.

Table 4.6: First digits distribution for voters' turnout in 2015

	First Digits	Frequency	Actual value	Benford's value
	1	0	0.000	0.301
	2	3	0.081	0.176
	3	10	0.270	0.125
	4	14	0.378	0.097
	5	6	0.162	0.079
	6	4	0.108	0.067
	7	0	0.000	0.058
	8	0	0.000	0.051
	9	0	0.000	0.046
Sum		**37**		
Mean			**3.946**	**3.441**

(ii) Second digits distribution

Table 4.7 shows that the distribution of second digits for the vote counts of APC in the 2015 presidential elections do not match Benford's distribution. The distribution of digits 3, 7, 8 and 9 is in excess when compared to Benford's distribution while digits 0, 1, 2 and 5 have lower frequencies. The mean distribution (5.000) of the vote counts also widely deviate from the expected mean of 4.187.

Table 4.7: Second digits distribution for APC in 2015

	Second Digits	Frequency	Actual value	Benford's value
	0	3	0.081	0.120
	1	3	0.081	0.114
	2	3	0.081	0.109
	3	5	0.135	0.104
	4	3	0.081	0.100
	5	1	0.027	0.097
	6	4	0.108	0.093
	7	5	0.135	0.090
	8	5	0.135	0.088
	9	5	0.135	0.085
Sum		**37**		
Mean			**5.000**	**4.187**

Table 4.8 shows the distribution of the second digits for the vote counts of PDP in the 2015 presidential election. The excess in occurrence of digits 4 and 5, as well as the absence of digit 9 present a major deviation from the expected Benford's value. Nevertheless, the mean value of the distribution of second digits (4.162) is closer to the expected mean of 4.187 than the mean value (5.000) of the distribution of second digits in the vote count of APC.

Table 4.8: Second digits distribution for PDP in 2015

	Second Digits	Frequency	Actual value	Benford's value
	0	4	0.108	0.120
	1	3	0.081	0.114
	2	2	0.054	0.109
	3	2	0.054	0.104
	4	7	0.189	0.100
	5	9	0.243	0.097
	6	5	0.135	0.093
	7	2	0.054	0.090
	8	3	0.081	0.088
	9	0	0.000	0.085
Sum		**37**		
Mean			**4.162**	**4.187**

In summary, the distribution of second digits of the vote counts for both the APC and PDP in the 2015 presidential election result is shown in Figure 4.11. The distribution reveals that the vote counts for both parties do not conform to the pattern predicted by Benford's Law.

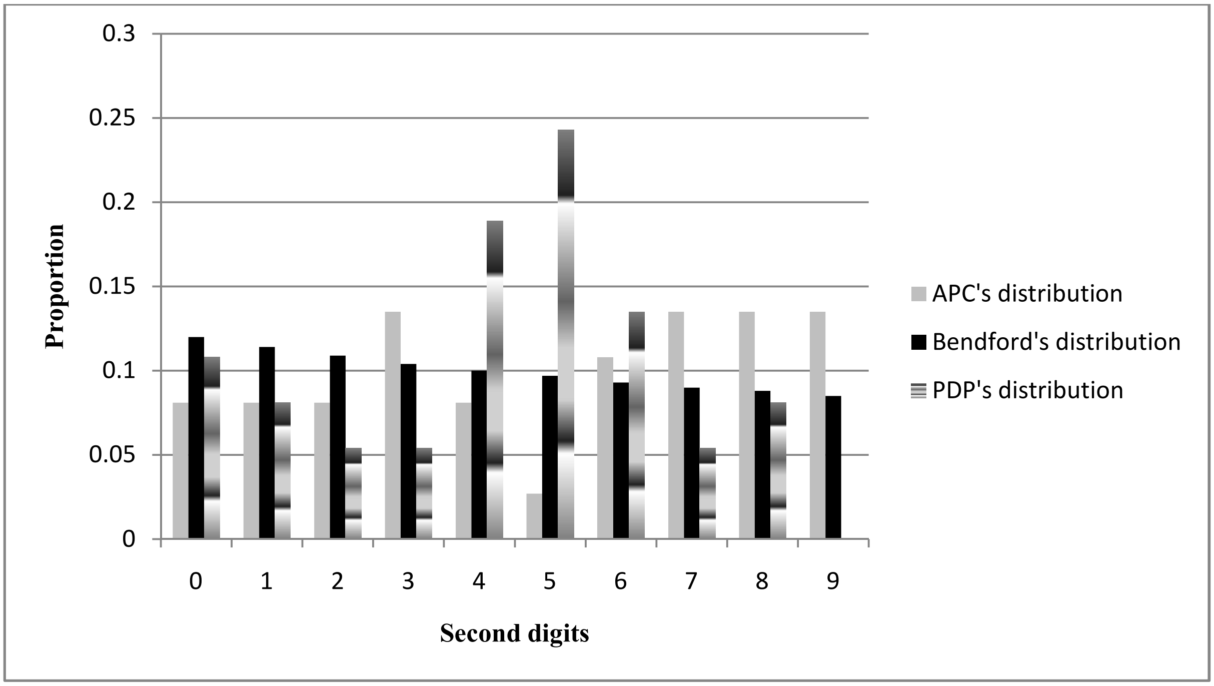

Figure 4.11: Second digits distribution for Benford's value and parties' votes in 2015

The distribution of the second digits for voters' turnout in the 2015 presidential election is presented in Table 4.9. The occurrence of digit 2 was abnormal and alarming when compared to the expected Benford's value. However, the mean value for the distribution of voters' turnout (4.189) approximates the expected mean value (4.187).

Table 4.9: Second digits distribution for voters' turnout in 2015

	Second Digits	Frequency	Actual value	Benford's value
	0	3	0.081	0.120
	1	3	0.081	0.114
	2	12	0.324	0.109
	3	3	0.081	0.104
	4	0	0.000	0.100
	5	4	0.108	0.097
	6	1	0.027	0.093
	7	1	0.027	0.090
	8	4	0.108	0.088
	9	6	0.162	0.085
Sum		**37**		
Mean			**4.189**	**4.187**

4.2.2 Benford's Law and the 2011 Presidential Elections

In this sub-section, Benford's Law for the first and second digit distribution were applied to the vote counts of PDP, CPC, ACN and ANPP in the 2011 presidential election. Benford's Law was also applied to the first and second digits distribution of the voters' turnout. The inferential statistics are in section 4.4 (Spearman Rank Correlation was used in comparing the actual and Benford's values).

(i) First digits distribution

The distribution of first digits for the vote counts of PDP in the 2011 presidential election is presented in Table 4.10. The actual values of the digits deviate from the expected Benford's distribution for first digits. The distributions of digits 1, 3 and 4 were in excess of the expected distribution. Also, there is no occurrence of digit 9. Only digit 5 approximates the expected Benford's value. The mean value for the vote counts also deviate from the expected mean.

Table 4.10: First digits distribution for PDP in 2011 presidential election

	First Digits	Frequency	Actual value	Benford's value
	1	12	0.324	0.301
	2	6	0.162	0.176
	3	6	0.162	0.125
	4	7	0.189	0.097
	5	3	0.081	0.079
	6	1	0.027	0.067
	7	1	0.027	0.058
	8	1	0.027	0.051
	9	0	0.000	0.046
Sum		**37**		
Mean			**2.865**	**3.441**

Table 4.11 shows the distribution of first digits for the vote counts of CPC in the 2011 presidential election. The actual values of vote counts deviate from the expected Benford's values. The frequencies of digit 1 and 6 were more than the expected Benford's distribution. Also, the mean of the distribution does not approximate the expected mean.

Table 4.11: First digits distribution for CPC in 2011 presidential election

	First Digits	Frequency	Actual value	Benford's value
	1	13	0.351	0.301
	2	3	0.081	0.176
	3	5	0.135	0.125
	4	3	0.081	0.097
	5	3	0.081	0.079
	6	5	0.135	0.067
	7	1	0.027	0.058
	8	2	0.054	0.051
	9	2	0.054	0.046
Sum		**37**		
Mean			**3.568**	**3.441**

The distribution of the first digits of vote counts for ACN in the 2011 presidential election is displayed in Table 4.12. The distribution does not conform to Benford's distribution. The absence of digits 8 and 9 also led to more deviation from the expected Benford's distribution for leading digits. Likewise, the mean for the distribution widely deviates from the expected mean.

Table 4.12: First digits distribution for ACN in 2011 presidential election

	First Digits	Frequency	Actual value	Benford's value
	1	16	0.432	0.301
	2	6	0.162	0.176
	3	4	0.108	0.125
	4	3	0.081	0.097
	5	4	0.108	0.079
	6	2	0.054	0.067
	7	2	0.054	0.058
	8	0	0.000	0.051
	9	0	0.000	0.046
Sum		**37**		
Mean			**2.649**	**3.441**

Table 4.13 presents the first digits distribution of vote counts for ANPP in the 2011 presidential election. The distribution shows that the first digits do not conform to Benford's distribution. However, the mean of the distribution is closer to the expected mean when compared to the vote counts of the other three leading parties (PDP, CPC and ACN).

Table 4.13: First digits distribution for ANPP in 2011 presidential election

	First Digits	Frequency	Actual value	Benford's value
	1	12	0.324	0.301
	2	7	0.189	0.176
	3	4	0.108	0.125
	4	1	0.027	0.097
	5	4	0.108	0.079
	6	2	0.054	0.067
	7	3	0.081	0.058
	8	3	0.081	0.051
	9	1	0.027	0.046
Sum		**37**		
Mean			**3.459**	**3.441**

In summary, the distribution of the first digits of vote counts for PDP, CPC, ACN and ANPP in the 2011 presidential election result is shown in Figure 4.12. The distribution shows that none of the vote counts for all the parties conform to the pattern predicted by Benford's Law for the first digits.

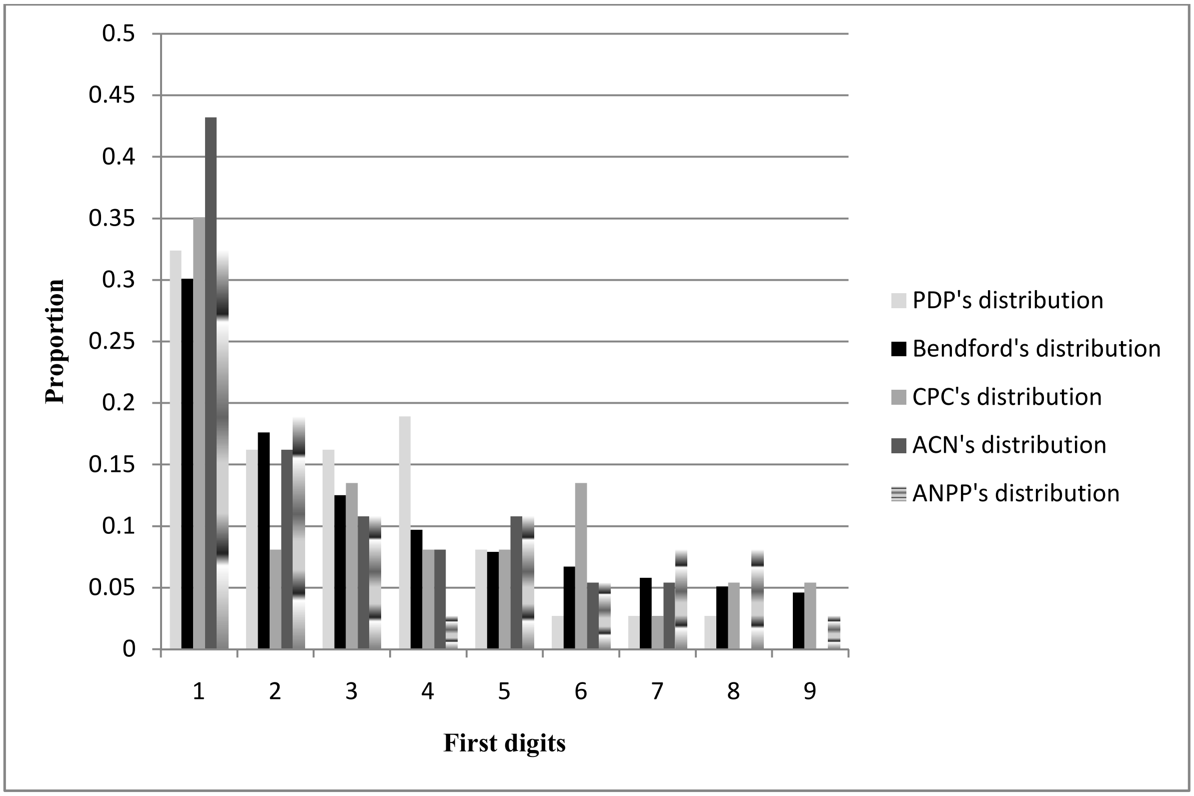

Figure 4.12: First digits distribution between Benford's value and parties' votes in 2011

The first digits distribution of the voters' turnout for the 2011 presidential election is displayed in Table 4.14. The distribution shows a very high distribution of digits 4, 5 and 6. The digit 1 that is expected to have a high frequency is absent in the distribution of the leading digits. The distribution does not, in any way, conform to Benford's distribution. Also, the mean of the distribution widely differs from the expected mean.

Table 4.14: First digits distribution for voters' turnout in 2011

	First Digits	Frequency	Actual value	Benford's value
	1	0	0.000	0.301
	2	1	0.027	0.176
	3	7	0.189	0.125
	4	10	0.270	0.097
	5	8	0.216	0.079
	6	6	0.162	0.067
	7	3	0.081	0.058
	8	2	0.054	0.051
	9	0	0.000	0.046
Sum		**37**		
Mean			**4.757**	**3.441**

In Figure 4.13, the first digits distribution of the voters' turnout in 2011 was compared with the first digits distribution of the voters' turnout in 2015. Figure 4.13 shows that digits 1 and 9 do not occur in the voters' turnout for both elections while digits 3, 4, 5 and 6 ocurred in excess when compared to the expected Benford's distribution. In summary, Figure 4.13 shows that voters' turnout for both presidential elections do not conform to the Benford's Law.

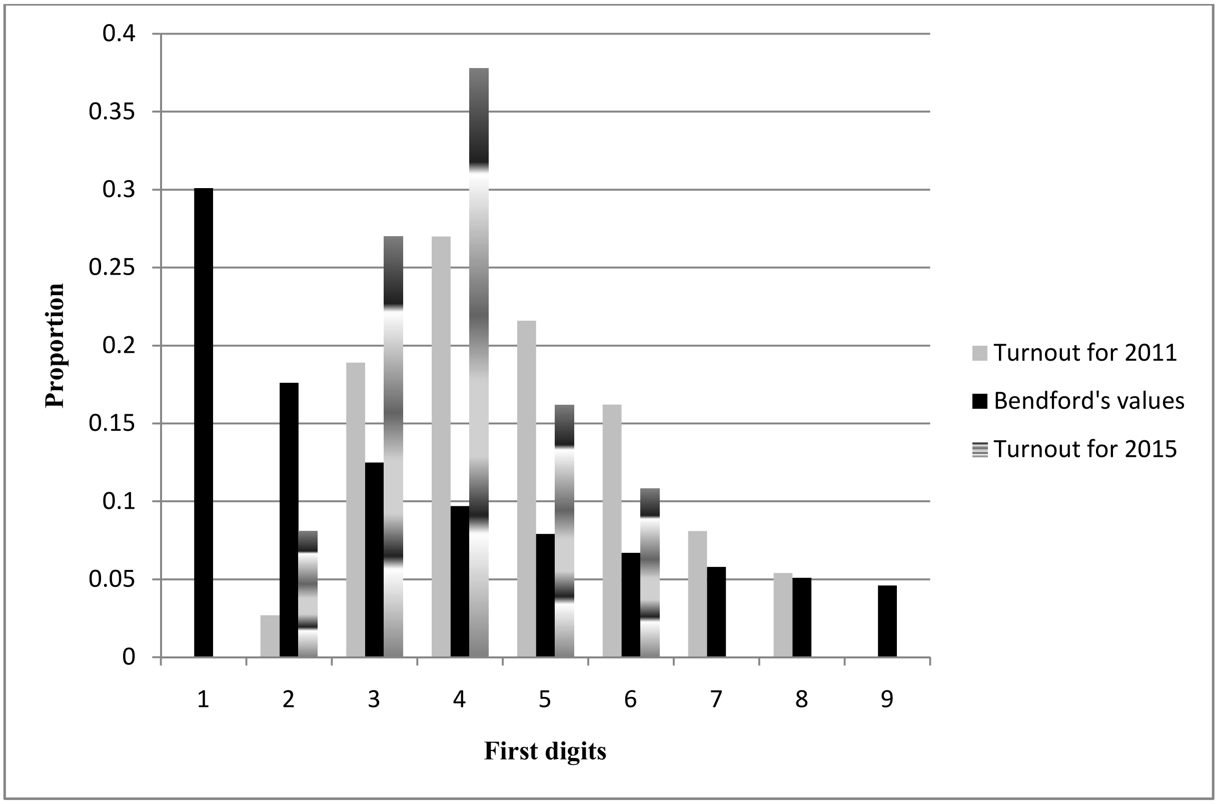

Figure 4.13: First digits distribution for Benford's value and turnout in 2011 and 2015

(ii) Second digits distribution

The distribution of the second digits of vote counts for PDP in the 2011 presidential election is presented in Table 4.15.The distribution shows that there was a high occurrence of digits 0, 1 and 8; as well as a missing digit 7; and a very low occurrence of digits 2, 4 and 6. The mean of the distribution also deviates from the expected mean.

Table 4.15: Second digits distribution for PDP in 2011 presidential election

	Second Digits	Frequency	Actual value	Benford's value
	0	9	0.243	0.120
	1	6	0.162	0.114
	2	3	0.081	0.109
	3	4	0.108	0.104
	4	2	0.054	0.100
	5	3	0.081	0.097
	6	2	0.054	0.093
	7	0	0.000	0.090
	8	5	0.135	0.088
	9	3	0.081	0.085
Sum		**37**		
Mean			**3.405**	**4.187**

Table 4.16 shows the distribution of the second digits of vote counts for CPC in the 2011 presidential election. Though the mean of the distribution is close to the expeted mean, Table 4.16 shows that the distribution does not conform to Benford's distribution. The table also shows the excessive occurrence of digits 3, 5 and 7.

Table 4.16: Second digits distribution for CPC in 2011 presidential election

	Second Digits	Frequency	Actual value	Benford's value
	0	5	0.135	0.120
	1	2	0.054	0.114
	2	3	0.081	0.109
	3	8	0.216	0.104
	4	2	0.054	0.100
	5	5	0.135	0.097
	6	3	0.081	0.093
	7	5	0.135	0.090
	8	1	0.027	0.088
	9	3	0.081	0.085
Sum		**37**		
Mean			**4.135**	**4.187**

The distribution of the second digits of vote counts for ACN in the 2011 presidential election is shown in Table 4.17. The distribution shows a very high occurrence of digits 2, 4 and 7; and a very low ocurrence of digits 1, 8 and 9. The distribution does not conform to Benfrod's law. The mean of the distribution also deviates from the expected mean.

Table 4.17: Second digits distribution for ACN in 2011 presidential election

	Second Digits	Frequency	Actual value	Benford's value
	0	4	0.108	0.120
	1	3	0.081	0.114
	2	6	0.162	0.109
	3	4	0.108	0.104
	4	6	0.162	0.100
	5	3	0.081	0.097
	6	3	0.081	0.093
	7	5	0.135	0.090
	8	1	0.027	0.088
	9	2	0.054	0.085
Sum		**37**		
Mean			**3.919**	**4.187**

Table 4.18 shows the distribution of the second digits of vote counts for ANPP in the 2011 presidential election. Apart from the absence of digit 8, Table 4.18 shows that the distribution does not conform to Benford’s Law. However, the mean value of the distribution (4.243) is close to the expected mean (4.187).

Table 4.18: Second digits distribution for ANPP in 2011 presidential election

	Second Digits	Frequency	Actual value	Benford's value
	0	3	0.081	0.120
	1	5	0.135	0.114
	2	4	0.108	0.109
	3	2	0.054	0.104
	4	5	0.135	0.100
	5	3	0.081	0.097
	6	6	0.162	0.093
	7	7	0.189	0.090
	8	0	0.000	0.088
	9	2	0.054	0.085
Sum		**37**		
Mean			**4.243**	**4.187**

In summary, the distribution of the second digits of vote counts for PDP, CPC, ACN and ANPP in the 2011 presidential election result is shown in Figure 4.14. The distribution shows that none of the vote counts for all the parties conform to the pattern predicted by Benford's Law.

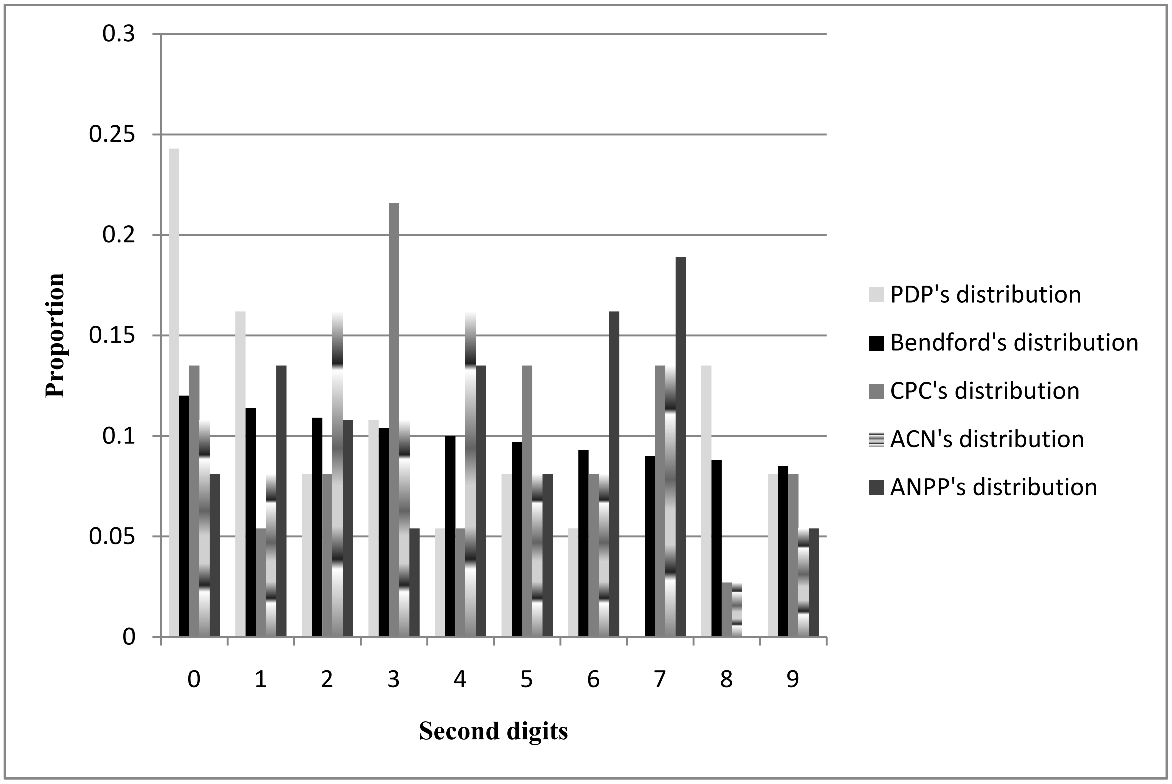

Figure 4.14: Second digits distribution for Benford's value and parties' votes in 2011

The distribution of the second digits for voters' turnout in the 2011 presidential election is presented in Table 4.19. The distribution shows an excessive ocurrence of digits 2, 3, 5, 6, 7 and 9 and a low occurrence of digits 0, 1, 4 and 8. The distribution, which does not conform to Benford's distribution for second digits, also has a higher mean value (4.811) than expected (4.187).

Table 4.19: Second digits distribution for voters' turnout in 2011

	Second Digits	Frequency	Actual value	Benford's value
	0	2	0.054	0.120
	1	2	0.054	0.114
	2	5	0.135	0.109
	3	6	0.162	0.104
	4	1	0.027	0.100
	5	5	0.135	0.097
	6	5	0.135	0.093
	7	4	0.108	0.090
	8	2	0.054	0.088
	9	5	0.135	0.085
Sum		**37**		
Mean			**4.811**	**4.187**

The second digits distribution of the voters' trunout in 2011 was compared with the second digits distribution of the voters' turnout in 2015 in Figure 4.15. The figure shows that digits 2, 5 and 9 occurred in excess for both elections. In summary, Figure 4.15 reveals that voters' turnout for both presidential elections do not conform to the Benford's Law.

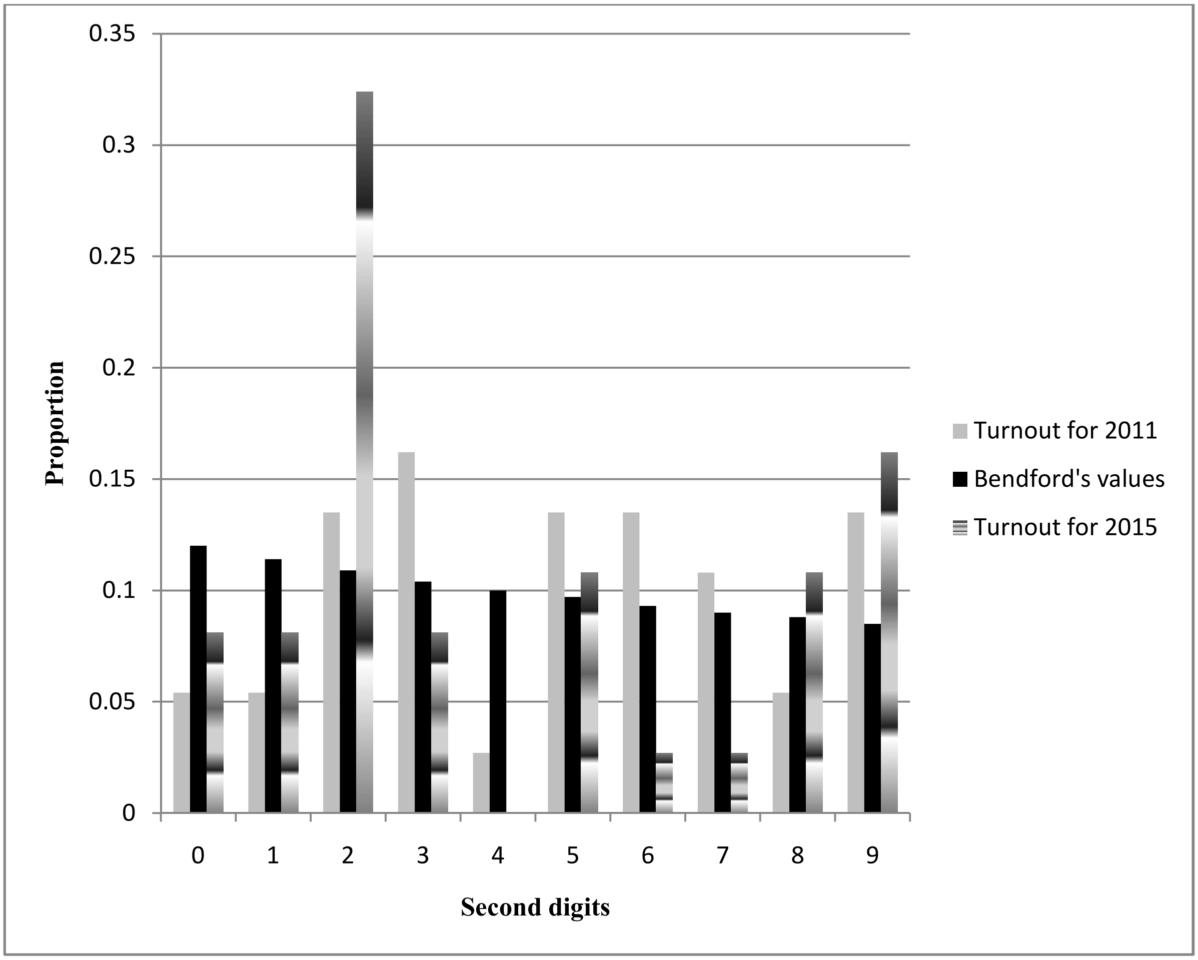

Figure 4.15: Second digits distribution for Benford's value and turnout in 2011 and 2015

4.2.3 Benford's Law and the 2007 Presidential Elections

In this sub-section, Benford's Law for the first and second digit distribution were applied to the total vote counts for all parties, as collated at the national level without state-by-state breakdown, in the 2007 presidential election. The inferential statistics are in section 4.4 (Spearman Rank Correlation was used in comparing the actual and Benford's values).

(i) First digits distribution

The distribution of first digits for the total vote counts in the 2007 presidential election is presented in Table 4.20. The actual values of the digits deviate from the expected Benford's distribution for first digits. The distributions of digits 2, 5, 6 and 8 were in excess of the expected distribution. Also, there is no occurrence of digit 9. The mean value for the total vote counts also slightly deviate from the expected mean.

Table 4.20: First digits distribution for votes in 2007 presidential election

	First Digits	Frequency	Actual value	Benford's value
	1	4	0.167	0.301
	2	9	0.375	0.176
	3	1	0.042	0.125
	4	2	0.083	0.097
	5	3	0.125	0.079
	6	2	0.083	0.067
	7	1	0.042	0.058
	8	2	0.083	0.051
	9	0	0.000	0.046
Sum		**24**		
Mean			**3.458**	**3.441**

(ii) Second digits distribution

Table 4.21 shows that the distribution of second digits for the total vote counts in the 2007 presidential elections do not match Benford's distribution. The distribution of digits 1, 2, 3, 4 and 6 is in excess when compared to Benford's distribution while digits 0, 5, and 9 have lower frequencies. There is no occurrence of digit 7. The mean distribution (3.917) of the vote counts also widely deviate from the expected mean of 4.187.

Table 4.21: Second digits distribution for votes in 2007 presidential election

	Second Digits	Frequency	Actual value	Benford's value
	0	1	0.042	0.120
	1	3	0.125	0.114
	2	3	0.125	0.109
	3	4	0.167	0.104
	4	5	0.208	0.100
	5	2	0.083	0.097
	6	3	0.125	0.093
	7	0	0.000	0.090
	8	2	0.083	0.088
	9	1	0.042	0.085
Sum		**24**		
Mean			**3.917**	**4.187**

4.3 Simulation of the Election Results using Monte Carlo Method

This section focuses on the simulation of the vote counts in the 2011 and 2015 presidential election results using Monte Carlo approach. The 2007 presidential election was not considered because of the absence of key variables (vote counts for parties at the state level) in the available result. Voters' turnout from the election results was also not considered for simulation of the 2011 and 2015 presidential election results because of lack of necessary details on the factors leading to the total number of registered voters. A total of 37 different surrogates of vote counts for each political party per state were generated, for each of the 2011 and 2015 presidential election, to represent the 37 different states (including the FCT) by simulating with a PERT distribution. The PERT distribution uses the minimum, most likely and maximum values of the observed election data as explained in chapter three (section 3.6). The analysis was carried out on the two leading political parties of the presidential election results.

4.3.1 Monte Carlo Simulation ofthe 2015 Presidential Election Result

This sub-section presents the Monte Carlo Simulation of the vote counts for APC and PDP in each of the geopolitical zones and all the states in the 2015 presidential election. Figure 4.16 presents the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the South-west geo-political zone. The figure shows that the vote counts for APC and PDP differ from the simulated votes in all the states in the South-west geo-political zone. Figure 4.16 also reveal that the actual votes for APC and PDP in Lagos State are far beyond the simulated votes for both parties.

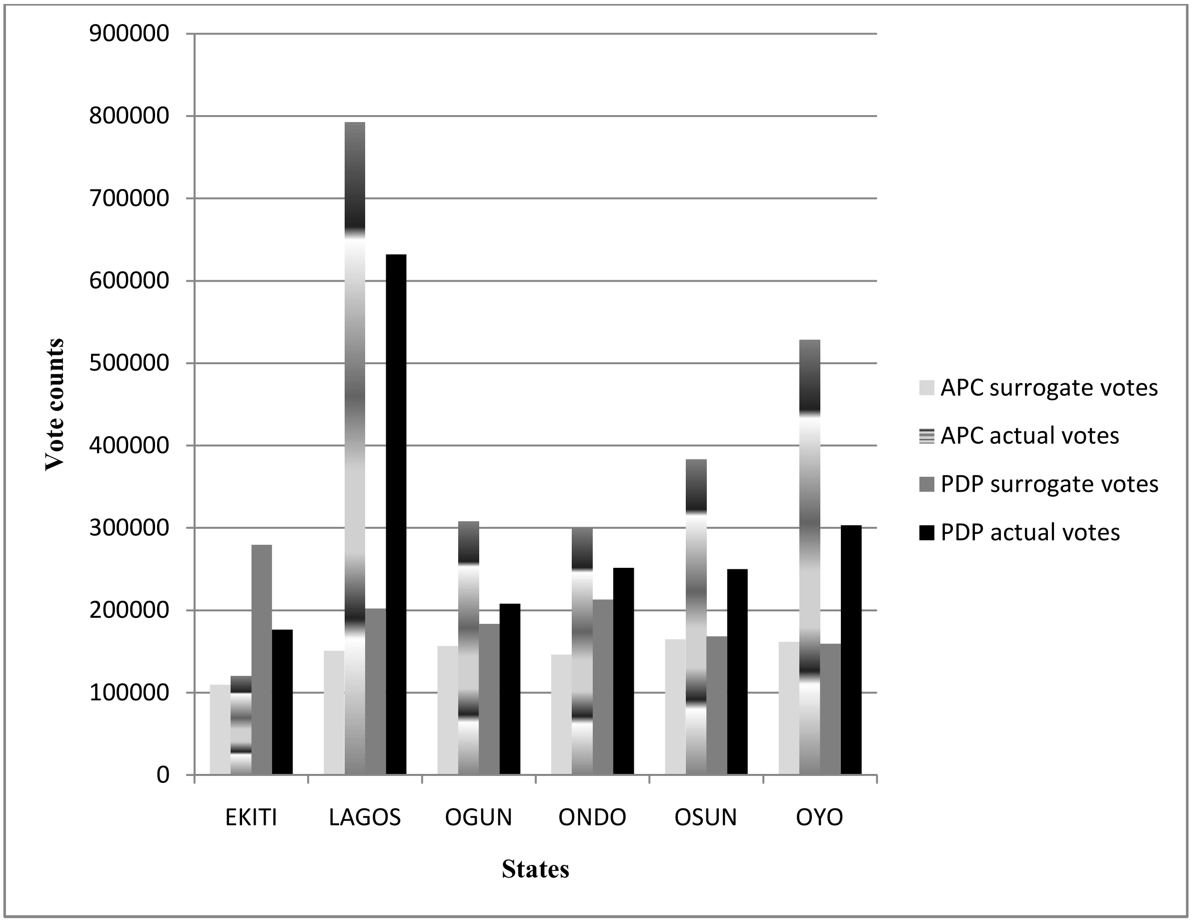

Figure 4.16: Monte Carlo surrogate and actual votes for South-west zone in 2015

Figure 4.17 presents the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the South-east geo-political zone for the presidential election in 2015. Figure 4.17 shows that the simulated votes and actual vote counts for APC are low and close in Abia State, Anambra State, Ebonyi State and Enugu State. However, the simulated votes and actual vote counts differ for APC in Imo State and for PDP in all the states in the South-east zone.

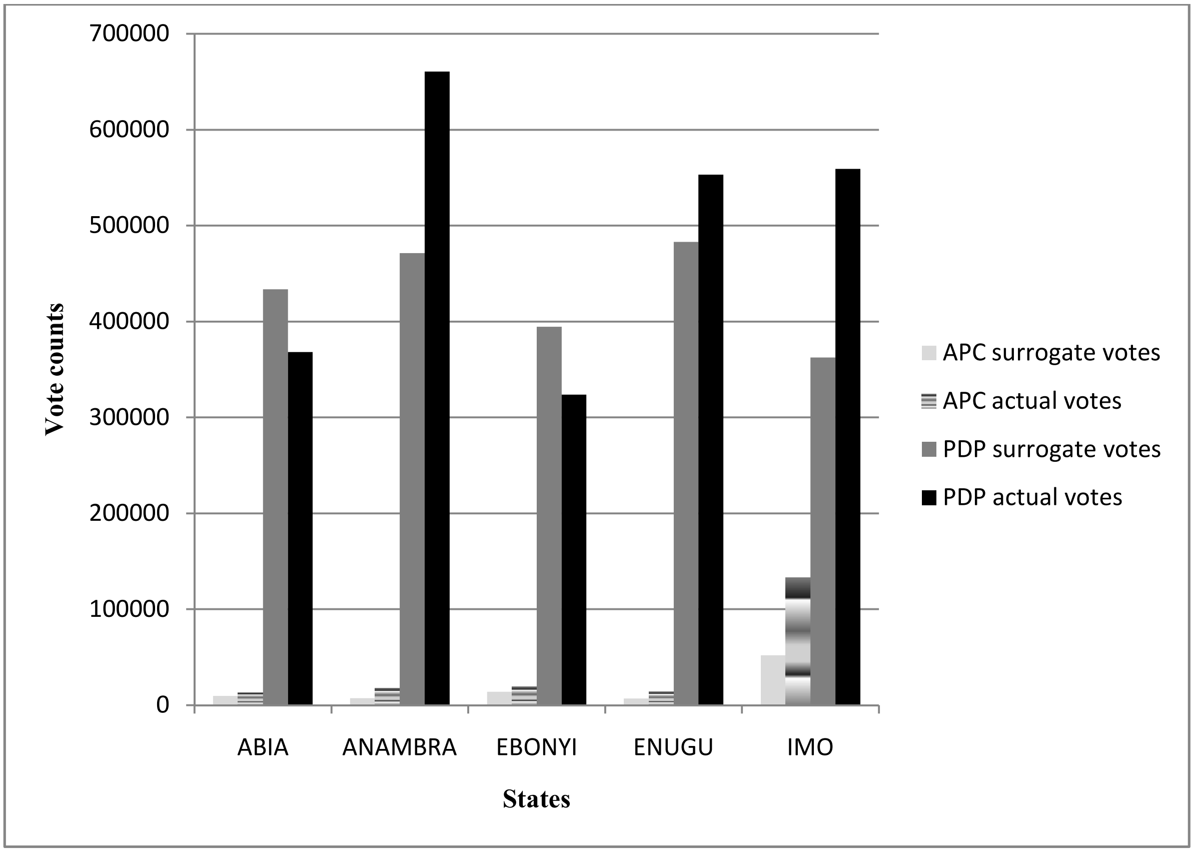

Figure 4.17: Monte Carlo surrogate and actual votes for South-east zone in 2015

Figure 4.18 presents the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the South-south geo-political zone for the presidential election in 2015. The figure shows that APC recorded very low votes in Bayelsa State. Figure 4.18 also reveals that the simulated votes were higher for PDP in all the states in the South-south Zone. However, the actual vote counts for PDPare alarming and highly exceed the simulated votes in Akwa Ibom State, Delta State and Rivers State.

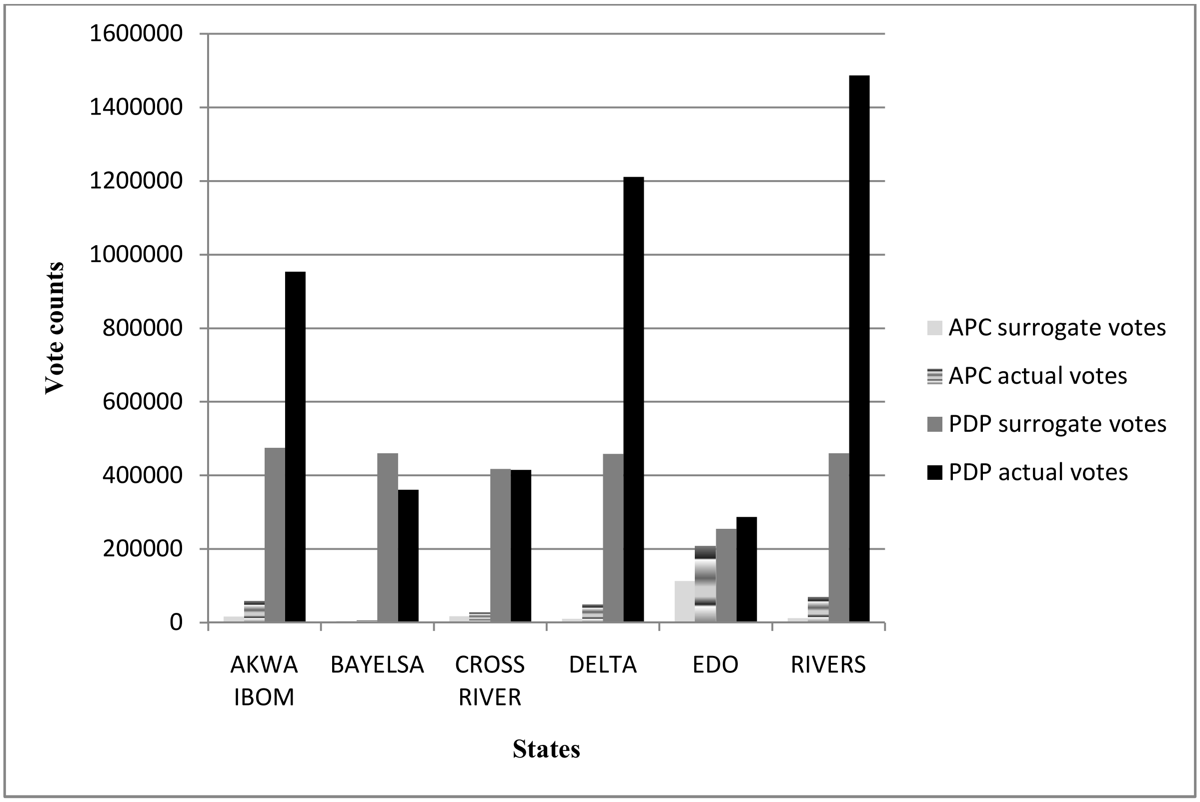

Figure 4.18: Monte Carlo surrogate and actual votes for South-south zone in 2015

Figure 4.19 shows the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the North-west geo-political zone for the presidential election in 2015. Although the simulated votes were higher for APC in all the states in the North-west geo-political zone, the figure shows an extremely higher actual vote counts for APC in Kano State when compared to the simulated votes.

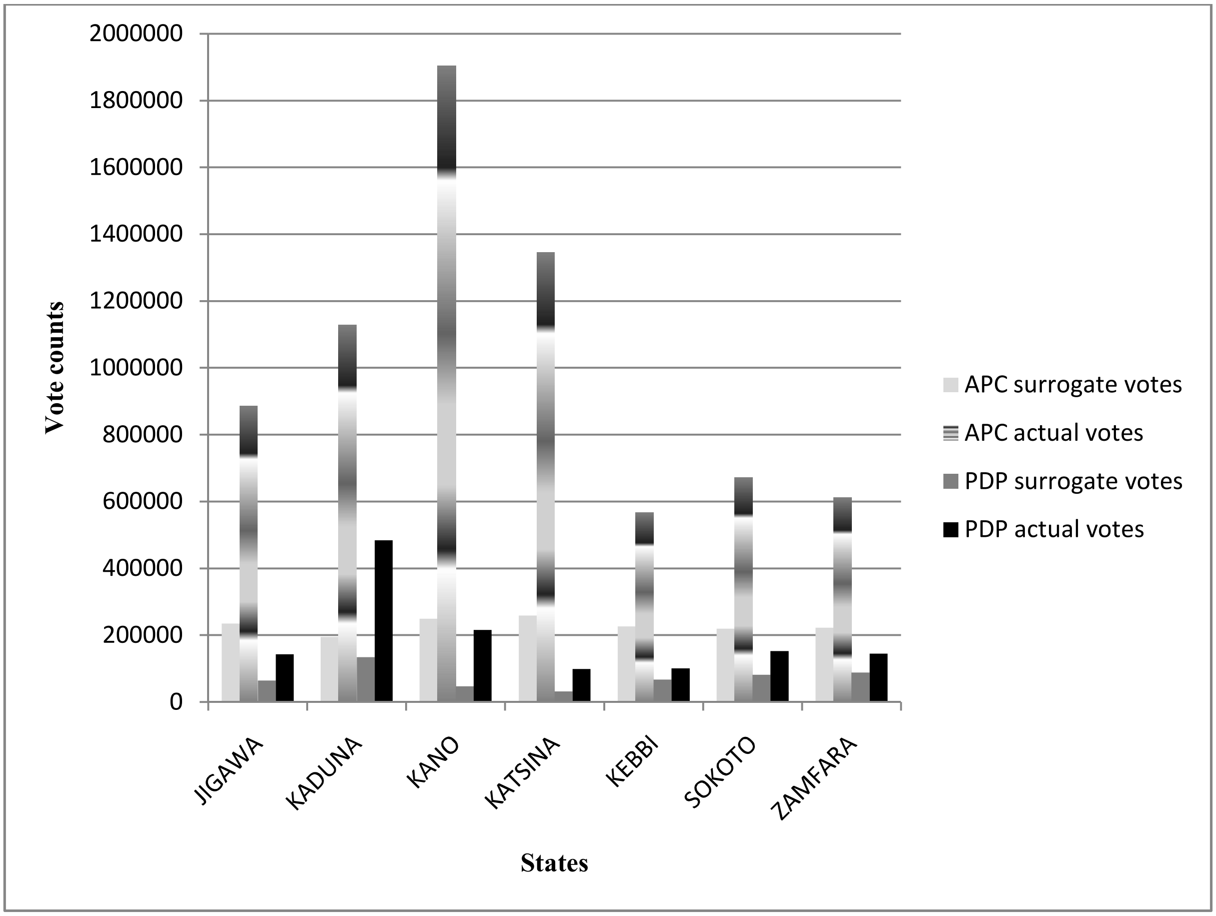

Figure 4.19: Monte Carlo surrogate and actual votes for North-west zone in 2015

Figure 4.20 presents the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the North-east geo-political zone for the presidential election in 2015. Figure 4.20 shows that the actual vote counts and the simulated votes are close for PDP in Borno State, Gombe State and Yobe State. On the contrary, the actual vote counts were extremely high for APC in Bauchi State when compared with the simulated votes.

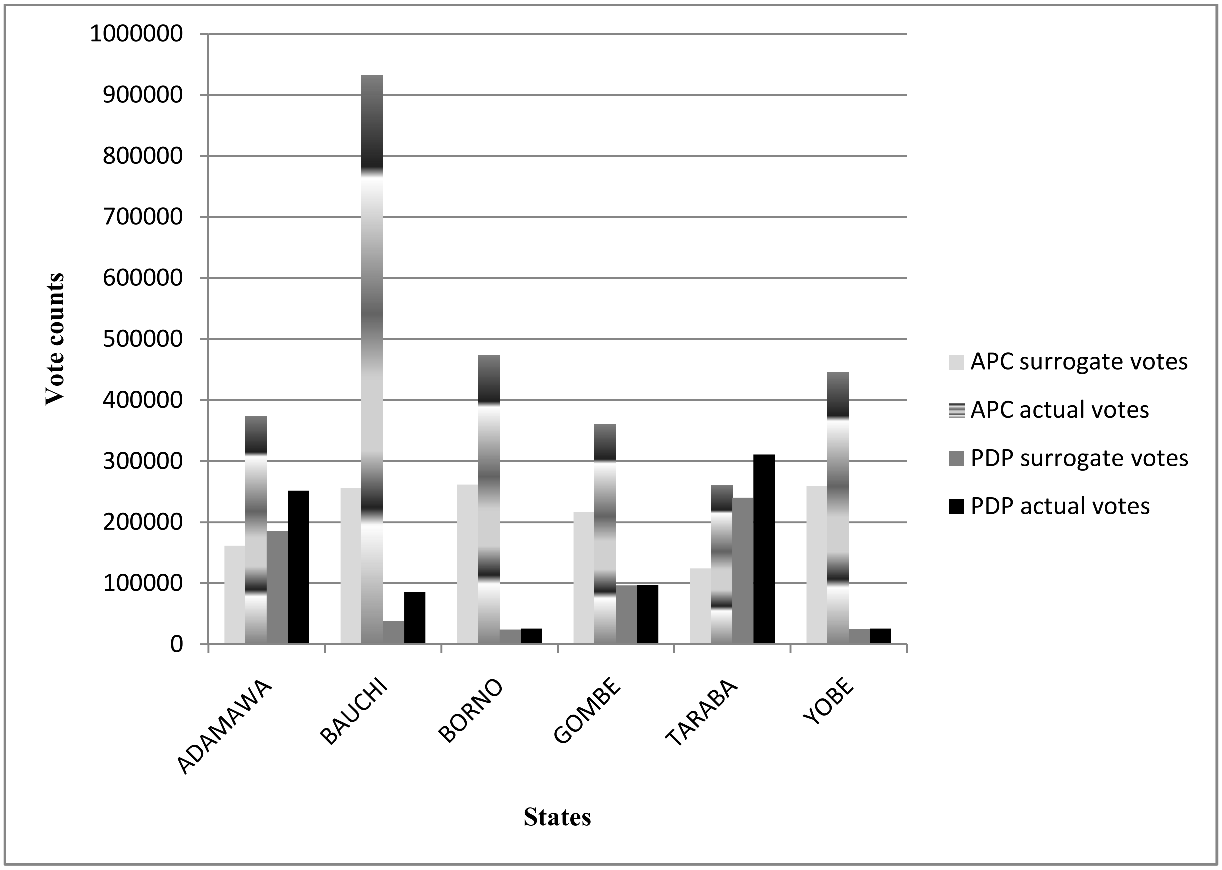

Figure 4.20: Monte Carlo surrogate and actual votes for North-east zone in 2015

Figure 4.21 shows the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in the North-central geo-political zone for the presidential election in 2015. The actual votes and the simulated votes are close for PDP in Kogi State and Kwara State, as well as for APC in the FCT. However, there is an extremely high vote counts for APC in Niger State when compared with the simulated votes.

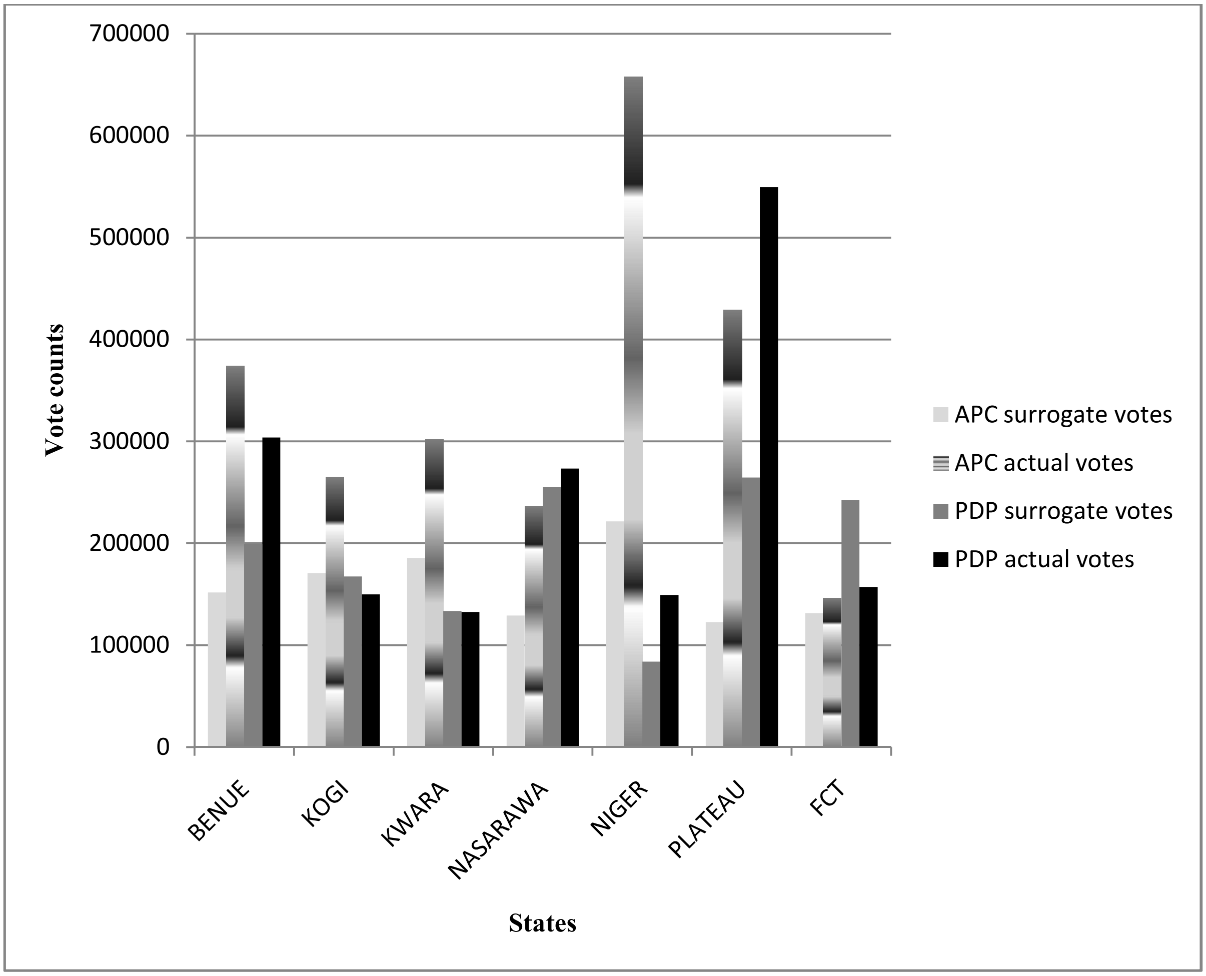

Figure 4.21: Monte Carlo surrogate and actual votes for North-central zone in 2015

Figure 4.22 shows the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in all the states for the presidential elections in 2015. For clarity, the distribution of the actual vote counts and the simulated votes have been analysed from Figure 4.16 to Figure 4.21. The mean (416889.8) and the standard deviation (413798.6) of the vote counts for APC in all the states differ from the mean (139155.8) and the standard deviation (87957.74) of the simulated votes. Likewise, the mean (347382.8) and the standard deviation (315993.0) of the vote counts for PDPin all the states differ from the mean (225596.4) and the standard deviation (152096.4) of the simulated votes.

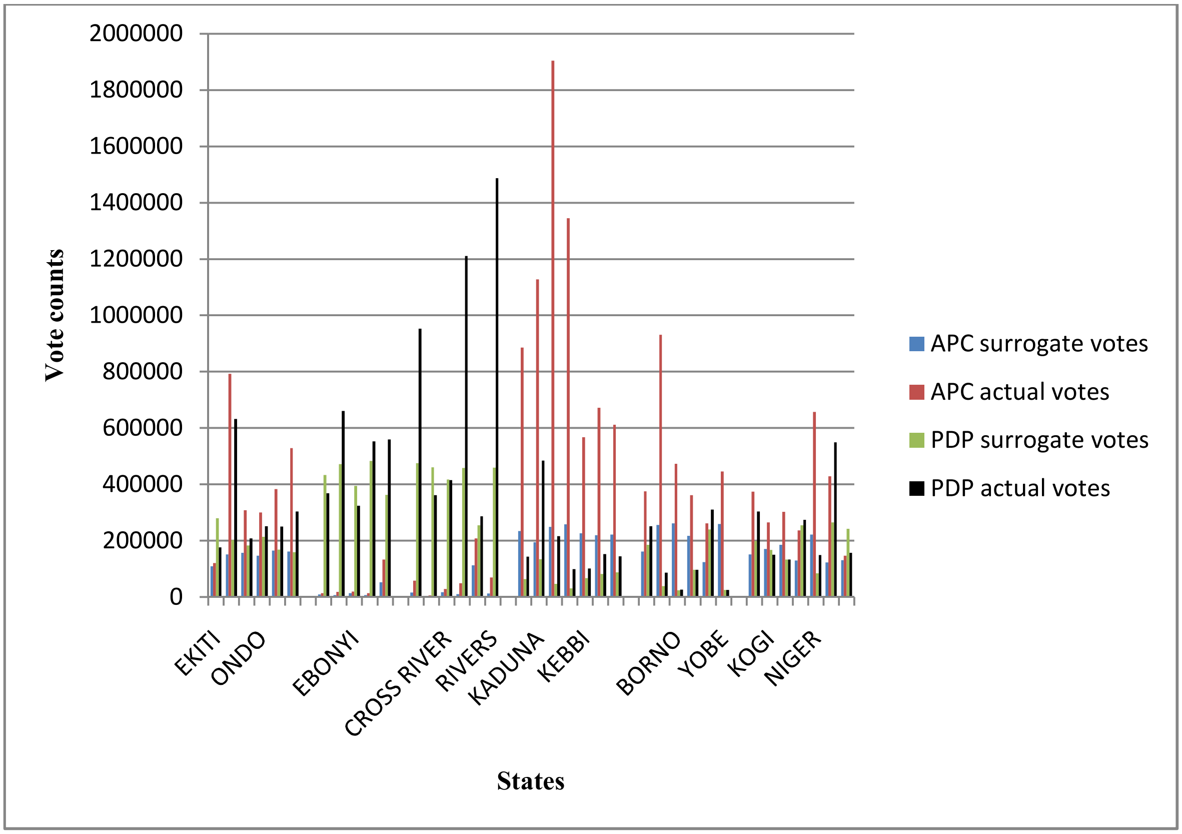

Figure 4.22: Monte Carlo surrogate and actual votes for all states in 2015

4.3.2 Monte Carlo Simulation ofthe 2011 Presidential Election Result

This sub-section presents the Monte Carlo Simulation of the vote counts for PDP and CPC in each of the geopolitical zones and all the states in the 2011 presidential election. Figure 4.23 presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the South-west geo-political zone. Figure 4.23 shows that the actual vote counts and the surrogate votes are higher for PDP in all the states in the South-west zone. Figure 4.23 also reveals that PDP recorded more votes than the expected simulated votes in the South-west, with the exception of Ekiti State. Also, the actual vote counts and the simulated votes are very low for CPC in Ekiti State, Ogun State, Ondo State and Osun State.

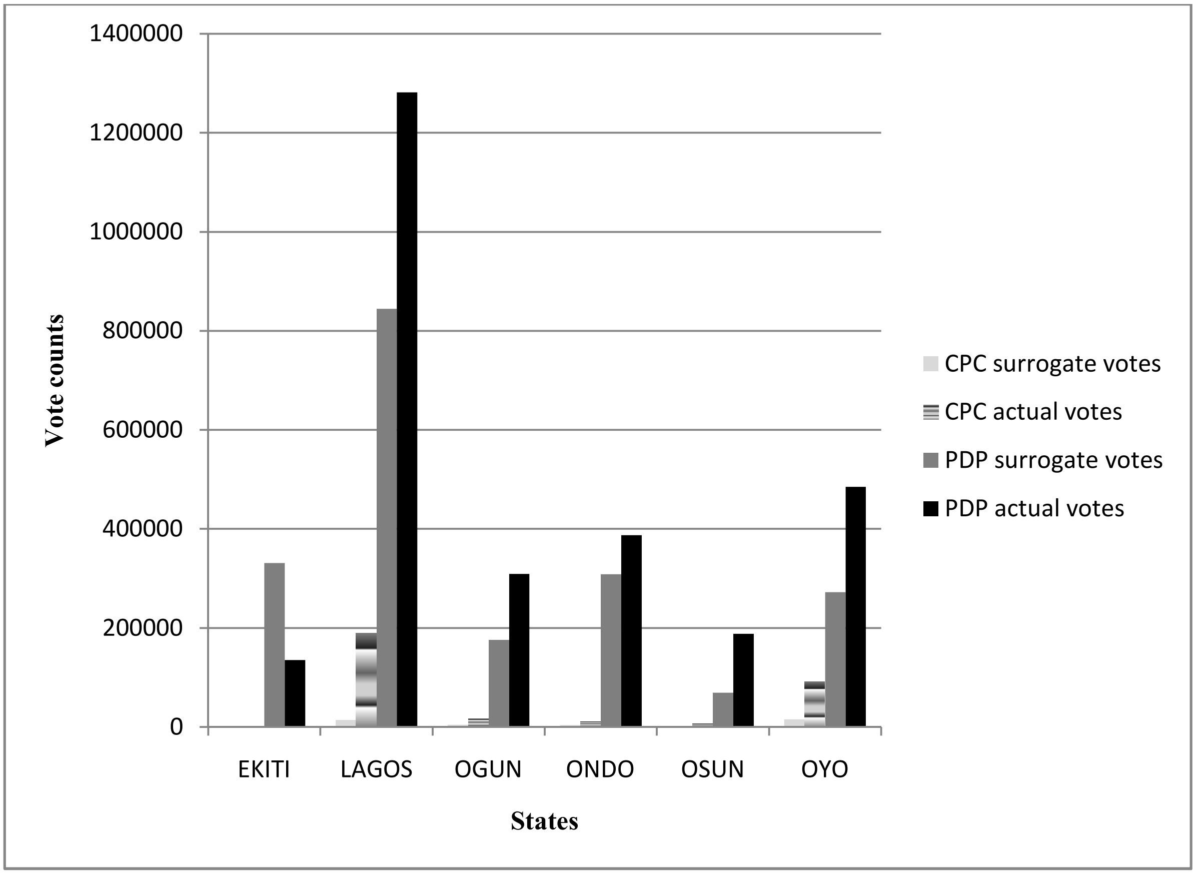

Figure 4.23: Monte Carlo surrogate and actual votes for South-west zone in 2011

Figure 4.24 shows the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the South-east geo-political zone. Figure 4.24 reveals that the actual vote counts and the surrogate votes are higher and very close for PDP in all the states in the South-east zone. Figure 4.24 also reveals that the actual vote counts and the surrogate votes are very close and very low for CPC in all the states in the South-east zone.

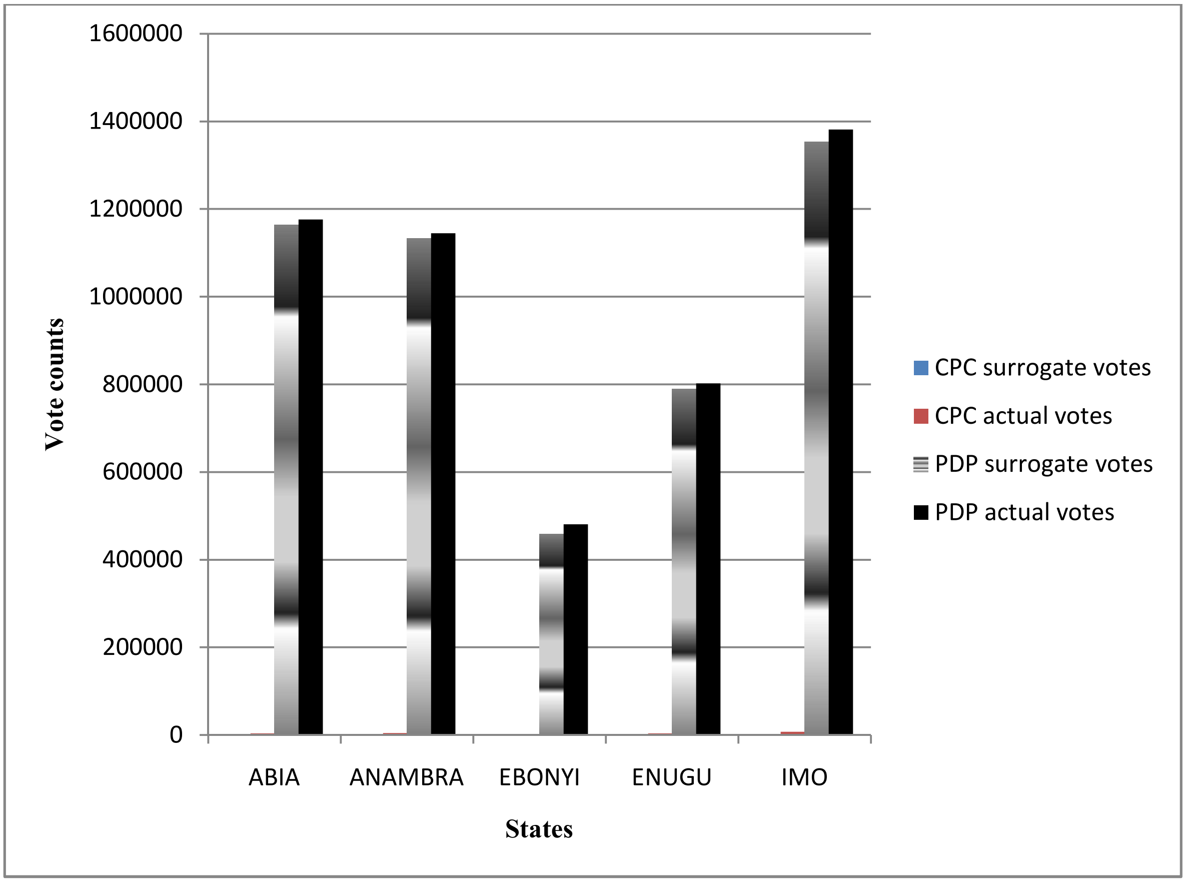

Figure 4.24: Monte Carlo surrogate and actual votes for South-east zone in 2011

Figure 4.25 presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the South-south geo-political zone. Figure 4.25 reveals that the actual vote counts and the surrogate votes are very low for CPC in all the states in the South-south zone. The figure also shows the actual vote counts and the surrogate votes are very high and very close for PDP in all the states in the South-south geo-political zone.

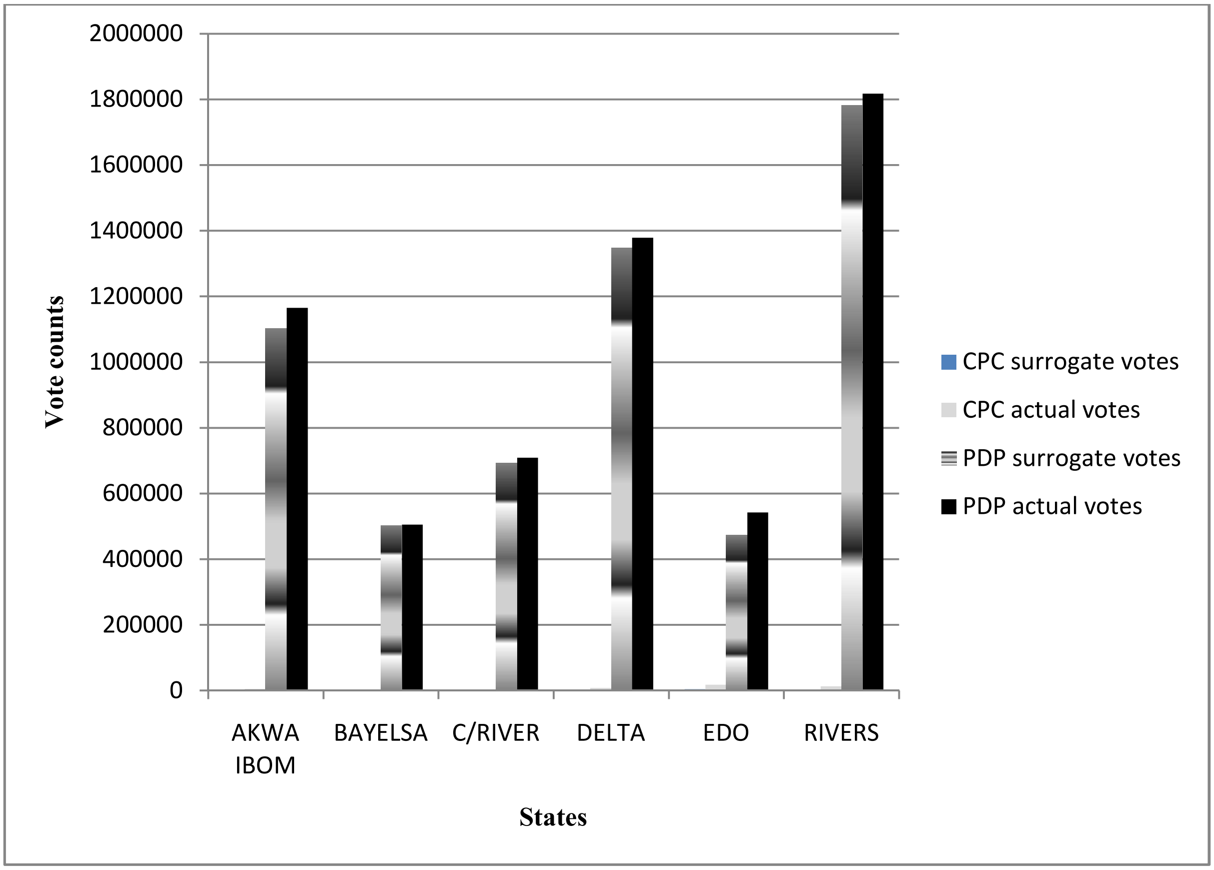

Figure 4.25: Monte Carlo surrogate and actual votes for South-south zone in 2011

Figure 4.26 shows the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the North-west geo-political zone. Unlike in the South-east and South-south zone, Figure 4.26 reveals that the actual vote counts and the surrogate votes differ for both the PDP and CPC. While the actual vote counts show that CPC recorded very high votes in all the states in the North-west zone, the simulated votes suggest that the results could have been different.

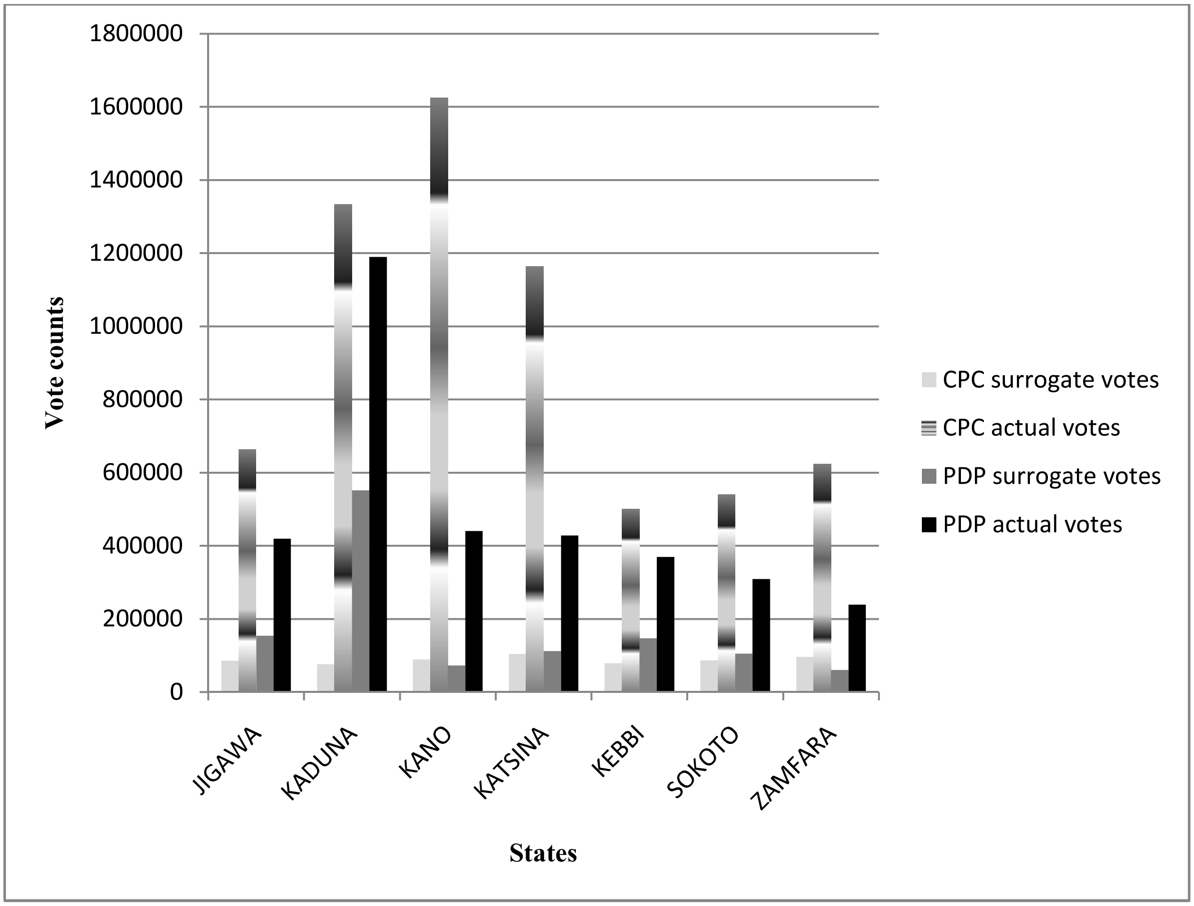

Figure 4.26: Monte Carlo surrogate and actual votes for North-west zone in 2011

Figure 4.27 presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the North-east geo-political zone. Figure 4.27 reveals that the actual vote counts for CPC was extremely high in Bauchi State and Borno State when compared with the simulated votes. With the exception of Gombe State, the actual vote counts and the surrogate votes agree on the leading party in all the states in the North-east zone.

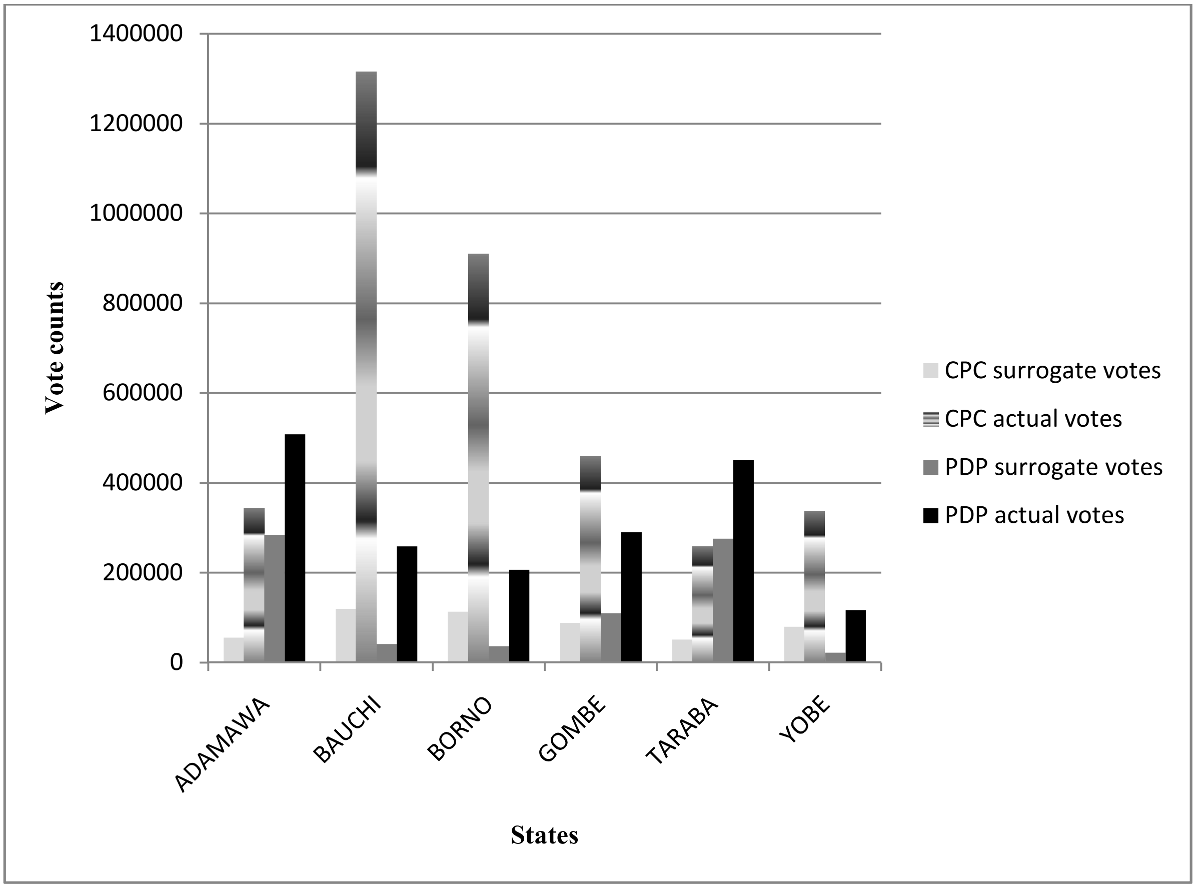

Figure 4.27: Monte Carlo surrogate and actual votes for North-east zone in 2011

Figure 4.28 presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDPand CPC in the North-central geo-political zone. With the exception of Niger State, the actual vote counts and the surrogate votes agree on the leading party in all the states in the North-central zone. While the actual vote counts reveal that CPC recorded very high votes in Niger State, the simulated votes suggest that the results could have been very close between the PDP and the CPC.

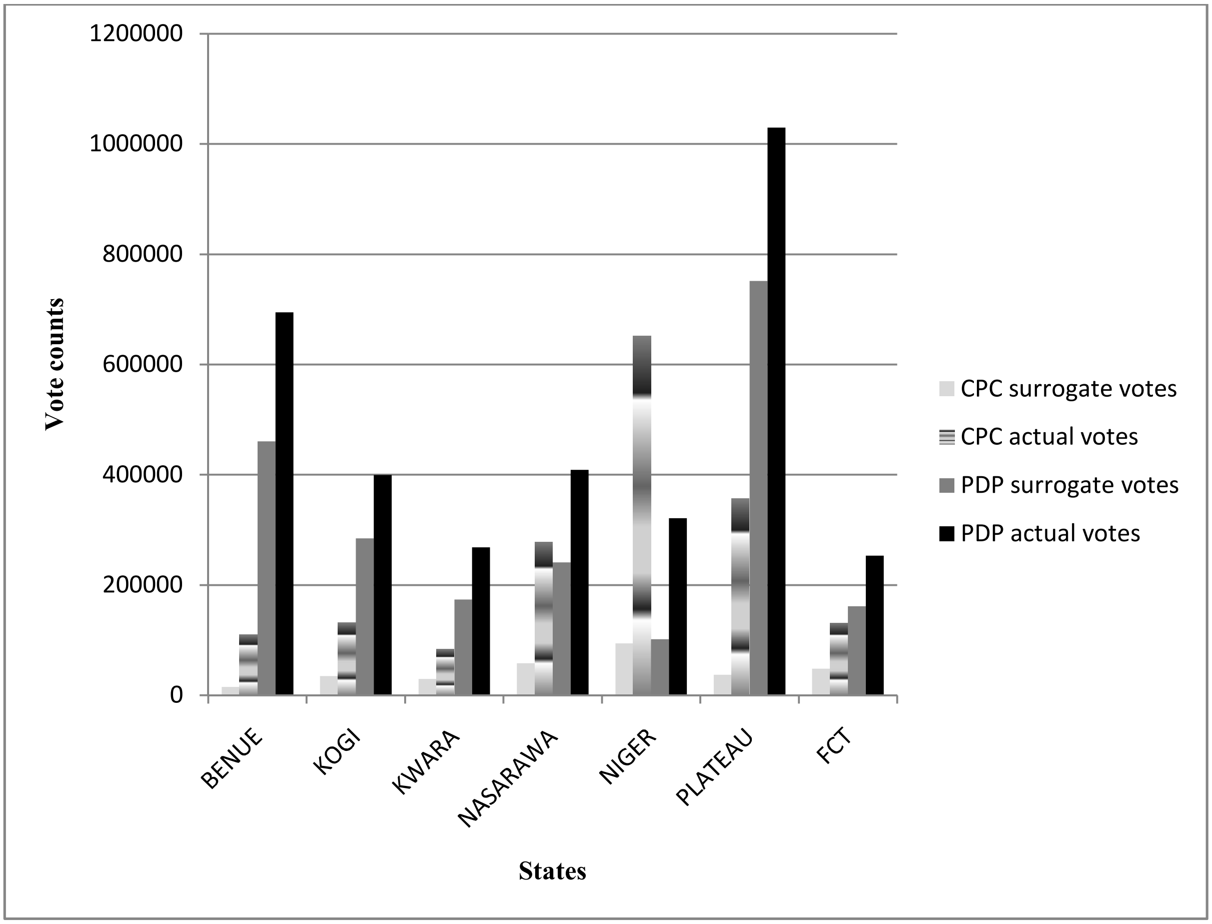

Figure 4.28: Monte Carlo surrogate and actual votes for North-central zone in 2011

Figure 4.29 shows the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in all the states for the presidential election in 2011. For clearer view, the distribution of the actual vote counts and the simulated votes have been analysed from Figure 4.23 to Figure 4.28. The mean (607978.0) and the standard deviation (429220.5) of the vote counts for PDPin all the states differ from the mean (458099.6) and the standard deviation (449434.8) of the simulated votes. Likewise, the mean (330131.2) and the standard deviation (438261.4) of the vote counts for CPCin all the states differ from the mean (40418.7) and the standard deviation (40979.6) of the simulated votes.

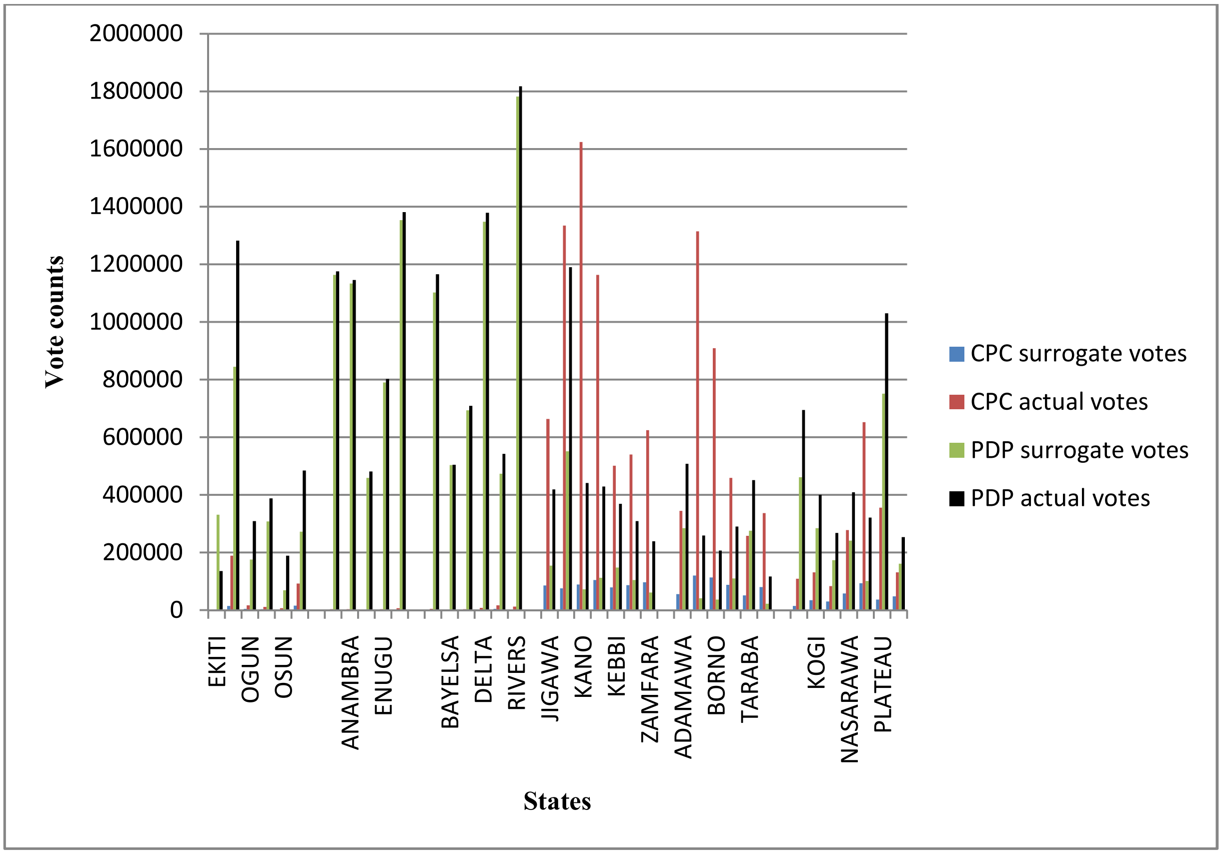

Figure 4.29: Monte Carlo surrogate and actual votes for all states in 2011

4.4 Test of Hypotheses

The tests of hypotheses are divided into three sub-sections. The first sub-section focuses on the Benford's frequency distribution of digits, the second sub-section focuses on the Monte Carlo surrogates, and the last sub-section focuses on test of variation in results across the geo-political zones in Nigeria. The tests of hypotheses were carried out on the results of the 2007, 2011 and 2015 presidential elections.

4.4.1 Benford's frequency distribution of digits

This sub-section presents the tests of hypotheses between the digital frequencies of vote counts of the 2007, 2011 and 2015 presidential elections and their respective Benford's frequency distribution. The sub-section also presents the tests of hypotheses between the digital frequencies of voters' turnout of the 2011 and 2015 presidential elections and their respective Benford's frequency distribution. Spearman Correlation Coefficient was adopted to test the hypotheses.

Hypothesis One: There is no significant correlation between the digital distribution of vote counts of election results in Nigeria and the digital distribution of Benford's Law.

Table 4.22 shows the Spearman Correlation Coefficient testsfor the first digits' distribution of the vote counts for APC and PDP in the 2015 presidential election results; PDP, CPC, ACN and ANPP in the 2011 presidential election results; and total vote counts of all the 24 political parties in the 2007 presidential election results.

Table 4.22: Spearman Correlation Coefficients for first digits' distribution of votes

		Correlation Coefficient	p-value	N
2015	APC/Benford	0.962	0.000	9
	PDP/Benford	0.740	0.023	9
2011	PDP/Benford	0.928	0.000	9
	CPC/Benford	0.752	0.019	9
	ACN/Benford	0.962	0.000	9
	ANPP/Benford	0.726	0.027	9
2007	Total votes	0.681	0.043	9

Table 4.22 shows that the null hypothesis is rejected for the first digits' distribution of vote counts for APC ($r = 0.962$, p-value < 0.05) and PDP ($r = 0.740$, p-value < 0.05)in the 2015 presidential elections; the first digits' distribution of the vote counts for PDP ($r = 0.928$, p-value < 0.05), CPC ($r = 0.752$, p-value < 0.05), ACN ($r = 0.962$, p-value < 0.05) and ANPP ($r = 0.726$, p-value < 0.05) in the 2011 presidential elections; and the first digits' distribution of the total vote counts ($r = 0.681$, p-value < 0.05) in the 2007 presidential elections. This implies that the distribution of first digits of vote counts for APC and PDP in the 2015 presidential elections; the distribution of first digits of vote counts for PDP, CPC, ACN and ANPP in the 2011 presidential elections; and the distribution of first digits of total vote counts for all the 24 political parties in the 2007 presidential elections correlate with (conform to) the expected distribution of first digits proposed by Benford's Law.

The Spearman Correlation Coefficient testsfor the second digits' distribution of the vote counts for APC and PDP in the 2015 presidential election results; PDP, CPC, ACN and ANPP in the 2011 presidential election results; and total vote counts of all the 24 political parties in the 2007 presidential election results are presented in Table 4.23.

Table 4.23: Spearman Correlation Coefficients for second digits' distribution of votes

		Correlation Coefficient	p-value	N
2015	APC/Benford	-0.614	0.059	10
	PDP/Benford	0.234	0.515	10
2011	PDP/Benford	0.511	0.131	10
	CPC/Benford	0.199	0.581	10
	ACN/Benford	0.506	0.135	10
	ANPP/Benford	0.165	0.648	10
2007	Total votes	0.364	0.301	10

Table 4.23 reveals that the null hypothesis is accepted for the second digits' distribution of vote counts for APC (p-value > 0.05) and PDP (p-value > 0.05)in the 2015 presidential elections; the second digits' distribution of the vote counts for PDP (p-value > 0.05), CPC (p-value > 0.05), ACN (p-value > 0.05) and ANPP (p-value > 0.05) in the 2011 presidential elections; and the second digits' distribution of the total vote counts (p-value > 0.05) in the 2007 presidential elections. This indicates that the distribution of second digits of vote counts for APC and PDP in the 2015 presidential elections; the distribution of second digits of vote counts for PDP, CPC, ACN and ANPP in the 2011 presidential elections; and the distribution of second digits of total vote counts for all the 24 political parties in the 2007 presidential elections do not correlate with (do not conform to) the expected distribution of second digits proposed by Benford's Law.

Hypothesis Two: There is no significant correlation between the digital distribution of voters' turnout of election results in Nigeria and the digital distribution of Benford's Law.

The Spearman Correlation Coefficient tests for the first digits' distribution of voters' turnout in the 2015 and 2011 presidential election results are presented in Table 4.24. The table shows that the null hypothesis is accepted for the first digits' distribution of voters' turnout in both the 2015 (p-value > 0.05) andthe 2011 (p-value > 0.05) presidential elections.

Table 4.24: Spearman Correlation Coefficients for first digits' distribution of turnout

		Correlation Coefficient	p-value	N
2015	Turnout/Benford	0.374	0.321	9
2011	Turnout/Benford	0.067	0.864	9

Table 4.25 presents the Spearman Correlation Coefficient tests for the second digits' distribution of voters' turnout in the 2015 and 2011 presidential election results. The table also shows that the null hypothesis is accepted for the second digits' distribution of voters' turnout in both the 2015 (p-value > 0.05) andthe 2011 (p-value > 0.05) presidential elections.

Table 4.25: Spearman Correlation Coefficients for second digits' distribution of turnout

		Correlation Coefficient	p-value	N
2015	Turnout/Benford	-0.093	0.799	10
2011	Turnout/Benford	-0.171	0.637	10

The results imply that the distribution of first and second digits of voters' turnout in the 2015 and 2011 presidential election results do not correlate with (do not conform to) the expected distribution of first and second digits proposed by Benford's Law.

4.4.2 Monte Carlo Surrogates

This sub-section presents the tests of hypotheses between the surrogate votes in the 2015 and 2011 presidential elections and their respective actual vote counts. Spearman Correlation Coefficient was adopted in testing the hypotheses. As explained in chapter three, the tests of hypotheses three and four only reveal if the simulated data mimic the actual data.

Hypothesis Three: There is no significant relationship between the digital distribution of vote counts of election results in Nigeria and digital distribution of Monte Carlo simulation.
Table 4.26 shows the relationship between the actual vote counts,of APC and PDP in the 2015 presidential election and PDP and CPC in the 2011 presidential election, and the respective simulated votes. Table 4.26 shows that there is a significant relationship between the simulated votes and the actual votes of APC ($r = 0.890$, p-value < 0.05) and PDP ($r = 0.849$, p-value < 0.05)in the 2015 presidential elections as well as between the simulated votes and the actual votes ofPDP ($r = 0.881$, p-value < 0.05) and CPC ($r = 0.953$, p-value < 0.05) in the 2011 presidential elections.

Table 4.26: Spearman Correlation Coefficients for Monte Carlo simulated votes

		Correlation Coefficient	p-value	N
2015	APC/Surrogate's votes	0.890	0.000	37
	PDP/ Surrogate's votes	0.849	0.000	37
2011	PDP/ Surrogate's votes	0.881	0.000	37
	CPC/ Surrogate's votes	0.953	0.000	37

Therefore, the null hypothesis is rejected. This implies that there is a relationship between the simulated votes and their respective actual vote counts. It also implies that the simulated votes represent the actual vote counts for each of the political parties.

4.4.3 Test of variation in results across the geo-political zones

This sub-section presents the Kruskal-Wallis test for the variation in the vote counts for 2015 and 2011 presidential elections across the six geo-political zones in Nigeria. The zones and their respective states are South-west (Ekiti, Lagos, Ogun, Ondo, Osun and Oyo), South-east (Abia, Anambra, Ebonyi, Enugu and Imo), South-south (Akwa Ibom, Bayelsa, Cross River, Delta, Edo and Rivers), North-west (Jigawa, Kaduna, Kano, Katsina, Kebbi, Sokoto and Zamfara), North-east (Adamawa, Bauchi, Borno, Gombe, Taraba and Yobe) and North-central (Benue, Kogi, Kwara, Nasarawa, Niger and Plateau). For analysis in this chapter, FCT was added to the North-central zone. The result of the Kruskal-Wallis test (given as a Chi-square value) is described as H in Table 4.27.

Hypothesis Four: There is no significant difference in the voters' turnout across the geo-political zones in Nigeria.

Table 4.27: Kruskal-Wallis test for voters' turnout across geo-political zones

Dependent Variables	Degree of freedom	H	p-value
2015 presidential elections			
Voters' turnout	5	15.611	0.008
2011 presidential elections			
Voters' turnout	5	21.744	0.001

The Kruskal-Wallis testin Table 4.27 shows a significant difference inthe voters' turnout in 2015 presidential elections across the six geo-political zones ($H(5) = 15.611$, $p < 0.05$) and the voters' turnout in 2011 presidential elections across the six geo-political zones ($H(5) = 21.744$, $p < 0.05$). Therefore, the null hypothesis is rejected.

Unlike one way ANOVA, multiple comparism test such as post-hoc is not available for Kruskal-Wallis on SPSS. In order to know the nature of the differences in the voters' turnout across the six geo-political zones, 15 possible pairs of geo-political zones were selected and analysed using the Kruskal-Wallis test. The results reveal thatthe voters' turnout in the 2015 presidential election are not significantly different ($p > 0.05$) between the following zones: South-west and South-east; South-west and South-south; South-west and North-east; South-west and North-central; South-east and South-south; South-east and North-east; South-east and North-central; South-south and North-west; South-south and North-east; South-south and North-central; and North-east and North-central. In addition, the voters' turnout in the 2011 presidential election are not significantly different ($p > 0.05$) between the following zones: South-east and South-south; South-east and North-west; South-east and North-east; South-south and North-west; South-south and North-east; North-west and North-east; North-west and North-central; and North-east and North-central.

CHAPTER FIVE

DISCUSSION OF FINDINGS

5.0 Introduction

This chapter presents the discussion of findings based on the digital characteristics of the electoral results. This study adopted Benford's Law and Monte Carlo simulation and the discussion of major findings are presented in line with the two approaches. The first section presents the distribution characteristics of election results, followed by the discussion on the relationship between election results and Benford's distribution. The next section focuses on the relationship between election results and Monte Carlo simulation, while the final section presents the discusion on vote counts and voters' turnout across the geo-political zones in Nigeria.

5.1 Distribution Characteristics of Election Results

With the exception of the 2007 presidential elections, the distribution of first digits of vote counts for the parties in the 2015 and 2011 presidential elections, as shown in section 4.2 of this study, are left skewed as suggested by Benford (1938). This implies that lower digits, especially digit one, have more number of occurrence than the higher digits. However, from the different figures in section 4.2, there are differences between the distribution of the first digits of vote counts for each party and the expected Benford distribution. Also, the distribution of first digits of the voters' turnout for both the 2015 and 2011 presidential elections do not follow a left-skewed distribution and are completely different from the pattern predicted by Benford. The figures in section 4.2 also reveal that the distribution of second digits of vote counts for 2015, 2011 and 2007 presidential elections; and voters' turnout in the 2015 and 2011 presidential elections shows patterns that are different from the expected Benford's distribution for second digits.

According to Benford's Law,numbers should consistently fall into a pattern with low digits occurring more frequently in the first position than larger digits.Therefore, sequences of digits are usually expected to be uniformly or randomly distributed. However, the digital distribution of the first and second digits of the analysed election results in this study deviates from the Benford's Law. If the observed discrepancies are as a result of human manipulation of the election results, then, according to Benford's Law, some digits have been supplied in excess of what should have been in error-free results. Subsequent studies on election malpractice have also shown that people over-supply modes or under-supply digits while manipulating results (Mebane, 2006; Roukema, 2014).

Likewise, findingsfrom this study show that the last digits of vote counts for each party and voters' turnout in the 2015 and 2011 presidential elections are not uniformly distributed. Also, the distribution of the last digits of the total vote counts for each candidate in the 2007 presidential election is not uniform (Figure 4.9).In their study, Beber and Scacco (2008) as well as Levin, Cohn, Ordeshook and Alvarez (2009) argue that the last digits of electoral results should be uniformly distributed if there are no manipulations.They emphasised the fact that human tends to be biased when generating figures. Findings from this study support this assumption as digits 1 and 3 occurred more as last digits in most of the election results. The findings, therefore, reveal that the presidential election results from 2007 to 2015 deviate from uniformly distributed last digits.

5.2 Relationship between Election Results and Benford's Distribution

To make comparison between the expected Benford's distribution and the observed distribution of digits in the election results, analysts advocate the use of the mean values of distribution of digits as predicted by Benford, with focus on the second digits (Mebane, 2006;Deckert, Myagkov and Ordeshook, 2011; Mebane, 2013). Findings from this study show that the mean values might not be sufficient to establish malpractice in the relationship between the observed mean and the expected mean. While disparity in mean values give rise to a higher suspicion of malpracticein the total vote counts of 2007 presidential elections, voters' turnout and vote counts of PDP and ACN in the 2011 presidential election as well as the vote counts of APC in the 2015 presidential election, approximately close mean values in the vote counts for CPC and ANPP in the 2011 presidential elections as well as voters' turn out and vote counts for PDP in the 2015 presidential elections give rise to a very low suspicion.

Findings from the Spearman Correlation coefficients test for first digits, which show that first digits of vote counts for 2007, 2011 and 2015 presidential elections conform to Benford's

Law, might not be sufficient to establish anomalies in the vote counts. A better approach in establishing relationship between the observed distribution of digits and Benford's digits distribution is to adopt the nonparametric regression analysis for second digits (Mebane, 2013; Berdufi, 2014). Findings from the results of the hypotheses, using Spearman Correlation coefficients test for second digits,show that the vote counts for 2007, 2011 and 2015 presidential elections, and the voters' turnout in the 2011 and the 2015 presidential elections do not conform to Benford's Law. This implies that the approximately close values between the observed and expected mean in the vote counts of CPC and ANPP in the 2011 presidential elections as well as voters' turn out and vote counts for PDP in the 2015 presidential elections were not statistically significant enough to suggest that the results were error-free. Findings from the Spearman Correlation coefficients test therefore suggest the presence of anomalies in the collated presidential election results of 2007, 2011 and 2015.

5.3 Relationship between Election Results and Monte Carlo Simulation

First, normality tests were carried out using Kolmogorov-Smirnov and Shapiro-Wilk test of normality to give more insights into the distribution of the actual vote counts. The results of the normality tests reveal that the actual vote counts failed the normality test. This implies that the results of the normality tests suggest that the actual vote counts are not normally distributed. Therefore, a PERT distribution was used to simulate the actual election data. The PERT distribution was also considered suitable for the election data because there are fewer details in the state collated results released by INEC.

Findings from the pattern of the PERT distribution for the simulated data and the pattern of distribution of election data show that some of the vote counts (Figure 4.38 and Figure 4.46) differ from the expected simulated votes. This variation in the pattern of distribution suggests that the actual vote counts do not represent a random distribution. However, since the observed deviations are not based on statistically significant proof, there is no statistical evidence to back up the suggestion that the deviation in pattern represents a non-random process in the vote counts of the 2011 and 2015 presidential election results.

The test on hypothesis three shows that there is a significant relationship between the digital distribution of vote counts of election results in Nigeria and digital distribution of Monte Carlo simulation.Monte Carlo analysis employs random numbers in simulating data in a way that mimics the actual data as close as possible.The assumption is based on the expectation that election without manipulation should have vote counts that are random (Oleg, 2011; Johnson, 2011; Kobak, Shpilkin and Pshenichinkov, 2014). The results of the Spearman Correlation Coefficients tests on the relationship between the vote counts and Monte Carlo simulated votes show that the surrogate votes represent (or mimic) the actual vote counts from both the 2015 and 2011 presidential elections.

Nevertheless, Monte Carlo simulation is a weaker election forensic tool because the parameters for simulation of data could be subjective. On the contrary, Benford's Law is a natural and universal law. The parameters for modelling with Benford's Law are universal and objective. In comparison, Benford's Law is a better election forensic tool for analysing and detecting anomalies in electoral data.

5.4 Voters' Turnout across the Geo-political Zones in Nigeria

The results of the tests on hypothesis four in this study reveal that voters' turnout across the geo-political zones vary for both the 2011 and 2015 presidential elections. The findings from this study show that voters' turnout in the 2015 presidential elections was higher in the Northern part of the country than in the Southern part (with the exception of the South-south geo-political zone). The South-south geo-political zone has the highest voters' turnout in the 2015 presidential election. This could be attributed to the fact that the winner of the election is a native of the zone. The findings also indicate that a lot of electorates from the South-west and South-east geo-political zones did not turn out to vote.

The findings on the 2011 presidential elections reveal that the South-south and South-east geo-political zones recorded the highest voters' turnout. Although voters' turnout was higher in the Northern part of the country, the recorded turnout from the Southern part of the country was close to the recorded turnout from the Northern part of the country. Similar to the 2015 presidential elections, the South-west geo-political zone has the lowest voters' turnout in the 2011 presidential elections. Perhaps, this could be attributed to the attitudes or behaviours of the elites in the zone.

Considering the fact that the findings show higher turnout from some geo-political zones, it could be suggested that analyst, testing for election malpractice, should focus on locations with massive voters' turnout and votes. While it is possible that the massive voters' turnout and votes were a result of gerrymandering and voters' preferences, it is also possible that they were a result of malpractice. This assumption was also supported by Myagkov, Ordeshook and Shakin (2009); Levin, Cohn, Ordeshook and Alvarez (2009); and United States Agency for International Development (2017) in their report which revealed that data from locations with massive support for a particular candidate could signal malpractices.

CHAPTER SIX

MODEL FOR APPLICATION OF DIGITS-BASED TEST

6.0 Introduction

This chapter presents a model for the application of digits-based test and Monte Carlo simulation based on the findings of this study. The model, which can serve as a framework for future studies on election forensics, focused on holistic approach of combining Benford's Law, Monte Carlo simulation, Last-digits tests and election reports to detect anomalies in election data. The first section presents the application of election forensic tests; followed by the setbacks of digits-based tests and Monte Carlo analysis; and the chapter concludes with the proposed model for election forensic tests which is represented in a diagrammatic form.

6.1 The Application of Election Forensic Tests

This study has so far focused on the application of Benford's Law and Monte Carlo simulation in examining the digital distribution pattern of presidential election results in Nigeria. The results of analysis carried out in this study reveal that the digital distribution of the vote counts for the political parties and the voters' turnout do not follow the distributional pattern predicted by Benford's Law. Likewise, the deviation in the pattern of distribution of the Monte Carlo simulated votes and the actual vote counts could suggest anomalies in the actual vote counts.

The findings from the digit-based tests (Benford's distribution tests) in this study suggest that the election data are not error-free. The election forensics methods adopted for this study do not constitute, on their own, definitive proof that election malpractice did or did not occur in the adopted presidential elections in Nigeria. Nevertheless, there is a greater confidence in the probabilistic and statistical assessment of the election data if the different election forensics approaches are put in an appropriate framework.

6.2 Setbacks of Digits-based test and Monte Carlo analysis

The digits-based tests and Monte Carlo simulation carried out in this study focused on vote counts and voters' turnout. However, there are setbacks associated with the application of digits-based test using only vote counts and voters' turnout as variables. The sub-sections below present these major setbacks, based on the findings from this study.

6.2.1 Gerrymandering

Gerrymandering refers to thefact that, often in drawing legislative districts, imbalances are created so that one party hasa systematic advantage over the others. Gerrymandering does not necessarily suggest intentional manipulation as the imbalances may be created to reflect transient opinions rather than longstanding partisan divisions (Mebane, 2010b). The election data adopted for this study indicate that some candidates or party have greater support from some particular regions of the country, which often means that such candidates or party would have a higher share of votes in the states that constitute such region.

A major challenge with this can be seen in the application of first-digits test. For illustration, if the total vote cast in a state is approximately one million, it is expected that the first digit of the vote counts for the leading candidate, favored by gerrymandering, would be between 5 and 9. In some extreme cases, the first digits would be either 8 or 9 as seen in some of the presidential election results in this study. Probabilistically, this suggests that there is no equal chance of occurrence for digits 1 to 9 in such states. Hence, there is no strong logical or theoretical basis for adopting the first digits test in detecting discrepancyin Nigeria election data.

6.2.2 Voters' Preferences

The election data adopted for this study also indicate some level of biased voters' preferences for some candidates and parties in some states or geo-political zones. Voters' preferences, if independent and random, do not constitute a challenge to digits-based test. However, voters' preferences constitute a challenge if it is biased towards a candidate or party in such a way that all or almost all of the votes are in favor of such candidate or party. This implies that vote counts from such electorates do not reflect an independent and random process that is required in the application of digits-based election forensics methods.

Many scholars, as shown in this study, have argued that vote count is a process that strongly resembles the Benford's Law distribution in the seconddigits it produces. Their arguments could be valid on the fact that random votes from different voters are gathered from different locations. Hence, the election process exhibits some statistical mixtures that could fit into the proposed Benford's Law for the second digits. Nevertheless, there is no definite proof that voters' preferences could not influence the result of the application of second digits test in detecting election malpractice.

An alternative approach is to disaggregate the election results to the ward level before applying the last-digits test for each of the states. At the ward level, the analyst could easily check for extreme cases of voters' preferences such as situations where the same number of votes was recorded across polling units or all or almost all votes are for one party. Such cases could be isolated before applying the last-digits test. The submission of scholars is that no last digit will be repeated more frequently than any other in a series ofrandom draws. Hence, the last digit of the total number of votes cast at a polling unit is expected to correlate with the last digit of the vote count at the next polling unit. However, unlike in the more advanced democratic countries, application of last-digits test at ward levels remains a major challenge in Nigeria because INEC does not present detailed breakdown of election results.

6.3 Proposed Model for Election Forensic Tests

Gerrymandering and voters' preferences are two major factors that influence voters' turnout and vote counts. The challenges of voters' preferences, especially when the preferences are based on ethnicity, tribe or language, stem from the problem of gerrymandering. Gerrymandering, in simple terms, could be defined as grouping of voters into constituencies. These preferences do not necessarily reflect a fortuitous distribution of electorates' votes, as they often shift most of the votes to the favour of a particular candidate. Figure 6.1 presents the main variables that should be involved in the application of Benford's distribution and Monte Carlo methods.

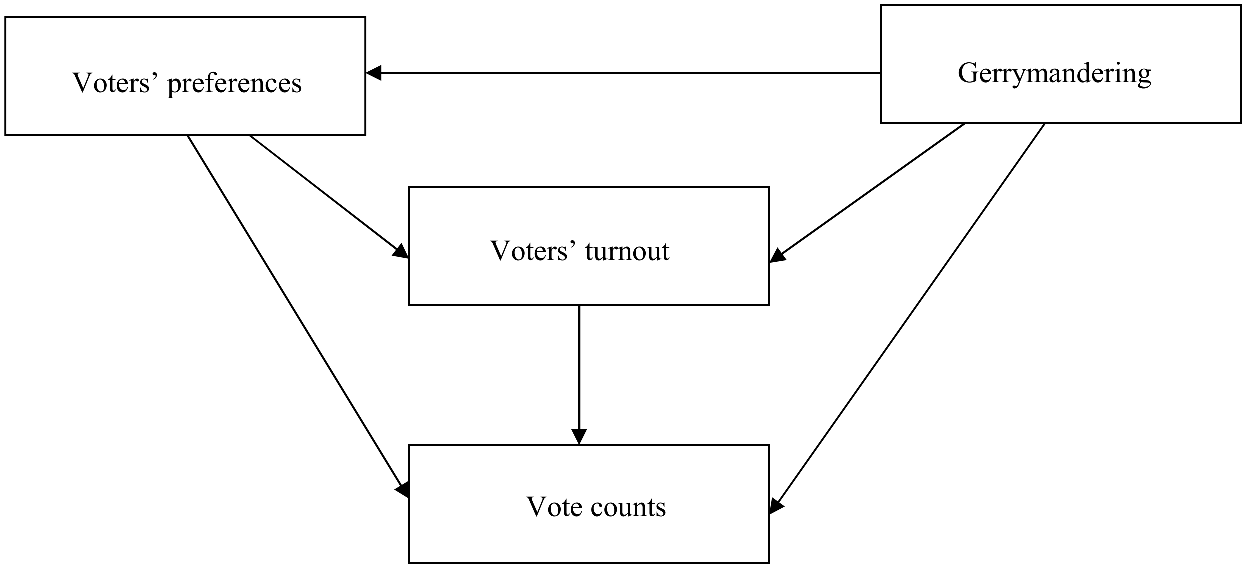

Figure 6.1: Variables influencing application of Benford's Law and Monte Carlo analysis

The figure shows that gerrymandering influences voter’ preferences, voters’ turnout and vote counts. The figure also shows that voters’ preferences influence voters’ turnout and vote counts; while voters’ turnout influences vote counts.

A diagrammatic representation of a model for the application of digits-based and Monte Carlo tests is presented in Figure 6.2. The model shows the different phases that should be observed in applying election forensic tests.

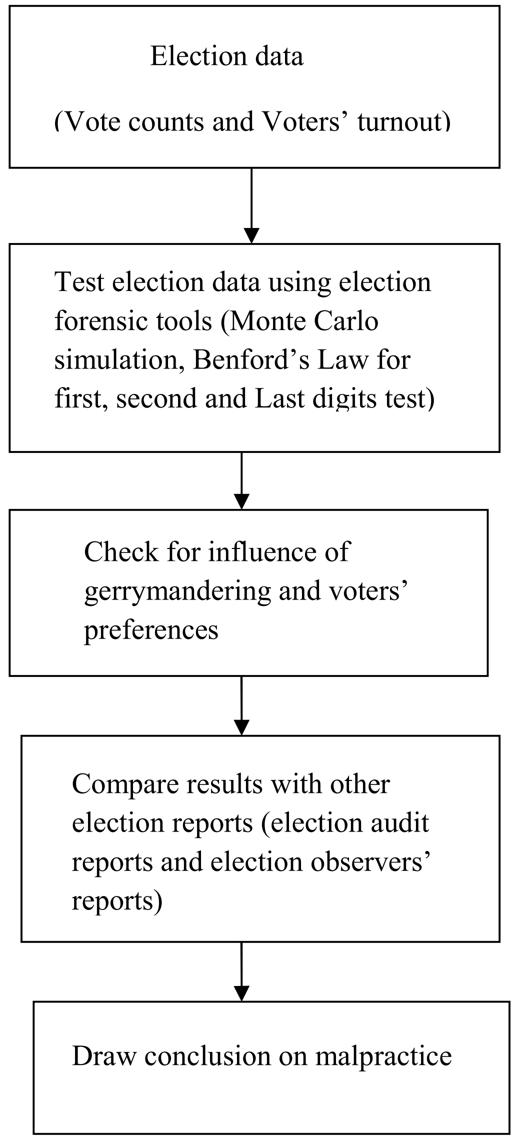

Figure 6.2: Model for the application of election forensic tests

Based on the findings from this study, the model in Figure 6.2 suggests that Monte Carlo simulation, and Benford's Law for first, second and Last digits tests should be combined to test for anomalies in election data (vote counts and voters' turnout). This is necessary because, in isolation, none of the methods is sufficient enough to give a clear view on the suspicion of election malpractice. If all the methods lead to similar indication of anomalies in the election data, there is a greater confidence in the suspicion that malpractice has occurred. However, as earlier stated, the result from the combined election forensic tools might not be sufficient proof to draw conclusion on election malpractice.

Therefore, it is important for the analyst to check for the influence of gerrymandering and voters' preferences on the results of the analysis. This will help to ascertain if the deviation in pattern detected during the analysis could be attributed to malpractices, rather than geographical or cultural influences. If the results of the analysis indicate that the election data might have been affected by gerrymandering and voters' preferences as discussed in section 6.2, the analyst should compare the result of the tests with other election reports before drawing conclusion on election malpractices.

Irrespective of the results of the election forensic analysis, it is important to compare all results with other election reports, such as election audit reports and election observer's reports. Information from on-the-ground monitors of elections could give more insight into malpractices. More so, the combined knowledge gained from the reports of the election forensic analysis and the other election reports from on-the-ground assessment would help the analysts to draw an objective and better conclusion.

CHAPTER SEVEN

SUMMARY, CONCLUSION AND RECOMMENDATIONS

7.0 Introduction

This chapter summarises the findings from this study. It also proposes areas of further studies. The first section presents the summary of findings, the second section presents the conclusion, the third section presents the contribution to knowledge, the fourth section focuses on the recommendations and the last section presents suggestions for further studies.

7.1 Summary of Findings

This research work examined the application of Benford's Law and Monte Carlo Simulation in modelling the presidential elections conducted in Nigeria between 2007 and 2015. The specific objectives of the study are: to examine the digital distribution characteristics of presidential electoral results in Nigeria; to determine if vote counts deviate from uniformly distributed Benford's digits; to determine if voters' turnout deviate from uniformly distributed Benford's digits; to examine if the actual vote counts represent the generated surrogate vote counts from the Monte Carlo simulation; and to present a model for the application of election forensic tests.

For the hypotheses, Spearman Correlation Coefficient was adopted to test for the relationship between the observed distribution of first and second digits in the results (vote counts and voters' turnout) and Benford's distribution for first and second digits. In addition, the frequency distributions of digits as well as the mean of the distributions in the vote counts and voters' turnout were also compared with frequencies and expected mean for first and second digits as predicted by Benford. According to Frank Benford, deviation of the distributions from the expected Benford's values suggests errors in the results. Monte Carlo simulation of the election results was carried out to produce surrogates vote counts for the elections. Spearman Correlation Coefficient was also adopted to test for the relationship between the vote counts from the elections and the Monte Carlo's surrogate votes. Kruskal-Wallis test was adopted to test for the difference in voters' turnout across the six geo-political zones in Nigeria.

The findings from this study show that the distributions of the first and second digits of vote counts for the political parties in the 2015, 2011 and 2007 presidential elections differ from Benford's distribution for first and second digits. Likewise, the distributions of the first and second digits of voters' turnout in the 2015 and 2011 presidential election results differ from the expected first and second Benford's distribution. The mean of the first and second digits of vote counts and voters' turnout in the elections also differ from the mean predicted by Benford for the first and second digits. The only exception was the mean of the second digit of voters' turnoutin the 2015 presidential elections which gave an approximatedvalue of the expected mean.However, it can be argued that the excessive occurrence of digits 2 and 9 in the distribution of the second digits led to a closer match between the mean of the distribution and the expected mean predicted by Benford. The distribution of the last digits of vote counts and voters' turn out in the election results are also not uniform. The last digits of vote counts and turnout in election results are expected to be uniformly distributed if there are no anomalies.

The results of the hypotheses tests that were carried out to examine the relationship between the digital distribution of vote counts in the 2015, 2011 and 2007 presidential election results and Benford's distribution show that there is no significant relationship between the second digits' distribution and Benford's distribution. Also, the tests of hypotheses reveal that there is no significant relationship between the digital distribution of the voters' turnout in the 2015 and 2011 presidential elections and Benford's distribution. The tests of hypotheses between the vote counts for the political parties in the 2015 and 2011 presidential elections and their respective Monte Carlo surrogates show that there is a significant relationship between the actual vote counts and the surrogates. This implies that the simulated data represent the actual data. Also, the Kruskal-Wallis tests show that the distribution of voters' turnout for the elections varies across the geo-political zones. The result of the Kruskal-Wallis tests does not deviate from what is expected in a general election.

In summary, the findings from this research workshow coherence between Benford's Law and Monte Carlo Simulation on the digital distribution of vote counts in the presidential election results in Nigeria. The digital distribution of vote counts and voters' turnout in the election results differ from the Benford's distribution. Likewise, the distributional pattern of vote counts differ from the observed pattern of the PERT distribution adopted for the Monte Carlo simulated votes. Therefore, the findings, based on the application of the two methods, suggest anomalies in the presidential election results examined in this study.

7.2 Conclusion

The importance of free and credible elections in the growth of democracy in Nigeria necessitates the need for election forensic analysis. The election forensic methods adopted for this study focus on Benford's Law, Monte Carlo simulation and last digits test. These methods were applied to find patterns in the distribution of election data that deviate from the expected distribution. The presence of such deviation in the distributional pattern of election data suggests that the data has been manipulated.

The results of analysis carried out in this study reveal that the digital distribution of the vote counts for 2015, 2011 and 2007 presidential elections, as well as the voters' turnout for the 2015 and 2011 presidential elections do not conform to the distributional pattern predicted by Benford's Law and the uniformly distributed last digits. The results of analysis also show that the digital distribution of vote counts and voters' turnout in the 2015 and 2011 presidential elections do not conform to Monte Carlo simulated data. An error-free data generating process is expected to produce counts, especially for the first two digits, where digit 0 through 9 should follow a particular pattern, with lower digits occurring more frequently than the higher digits. Also, the digital distribution of the vote counts for each of the political parties is expected to produce a pattern which shows that the counts are independently and randomly generated.

The findings from this study, therefore, suggest that the electoral process did not produce independent and randomly generated votes in the presidential elections. The findings also suggest that the 2015, 2011 and 2007 presidential election results are not error-free. However, only the Benford's distribution test and last digits test were carried out on the 2007 presidential election data as a result of the fewer details in the result published by INEC. Therefore, the findings, which reveal that the digital distribution of the vote counts for the 24 presidential candidates of the 2007 elections do not conform to expected distributional pattern, might not be sufficient to suggest that the election was not error-free.

The study concludes with the presentation of a model which could be adopted by scholars in the election forensics field. The model suggests that factors such as gerrymandering and voters' preferences influence vote counts and voters' turnout. A proper understanding of the relationship between these factors could guide an election forensic analyst in making better conclusion. The model also suggests that election forensic test should be adopted as a compliment to other election reports on election malpractices.

7.3 Contribution to Knowledge

This study investigated the applicability of Benford's Law and Monte Carlo simulation in modelling presidential election results released by INEC between 2007 and 2015. In adopting two different models, this study has contributed to the existing literature by revealing the significance of digits and digital patterns in the vote counts and voters' turnout of electoral data. As Nigeria continues to struggle with the problem of electoral integrity, this study can provide additional forensic tools and insight for INEC, election observers and researchers.

A major significant contribution of this study is the application of forensic techniques to election results in Nigeria beyond what is known in the literature. More so, unlike the reports from domestic and international observers which could be selective and subjective, the analysis carried out in this study produced empirical results that could be more objective. This is so because the forensic analysis was carried out on the whole data, gathered at the state level, across the nation, as reported by INEC. In addition, the information and conclusions from this study are based on forensic evidence and statistical metrics, not mere allegation of election malpractices.

The application of first and second digit's distribution of Benford's Law, and the simulation with Monte Carlo technique, has helped to broaden the knowledge base and give more insights into election data pattern in Nigeria. This research work has also given insight into the applicability of last digit-based test on electoral malpractices. The result of the forensic analysis in this study, which shows that the adopted presidential election results in Nigeria deviate from the expected randomness in error-free elections, has therefore contributed to knowledge in the forensic analysis of malpractices in the presidential elections in Nigeria.

This study has contributed in developing analytical insight into election malpractices in Nigeria through the use of formal forensic models and empirical cases. The proposed model has also provided a foundation for future research tailored towards developing more accurate measures of electoral malpractices. However, since the adopted methods in this study basically produce statistical anomaly in suspicion of malpractices, the results of the election forensics analysis should be adopted as a complement to the information gathered from election audits, aswell as reports from election observers.

7.4 Recommendations

Based on the findings, the following recommendations are made:

1. INEC should improve on the quality of election data by providing detailed election results with consistent format across the different election year. The public should be able to obtain resultscollated at the lower levels (polling units and wards) from INEC. This will help researchers in making better and robust analysis on election data for each election year. It will also help to increase the researchers, observers and electorates' trust and confidence in the election process.
2. This study has shown that election forensic analysis could be used to flag suspicious election data. As shown in the proposed model for election forensic tests, the results of the election forensic analyses could augment the reports from on the ground observationby election observers. It is recommended that domestic and international observers should adopt election forensic analysis as a complement to the reports from on the ground assessment.
3. This study has also shown that an error-free data generating process is expected to produce vote counts with a particular pattern which shows that the counts are close to a natural process. It isrecommended that the Independent National Electoral Commission should adopt the proposed model for insight into suspicious electoral data pattern and adjust the electoral process in line with findings.

7.5 Suggestions for Further Studies

1. One of the major limitations of this study is the unavailability of detailed election result from INEC. With the exception of the results for the 2011 and 2015 presidential elections, the 2007 presidential election results released by INEC lacked the necessary details that are needed for a robust application of digits-based tests. The missing details include breakdown of vote counts for each presidential candidates at the state and the total number of registered voters at each state. There is no doubt that INEC, so far, has made efforts to improve from one election year to the other. However, the wide margin

between the registered voters and the accredited voters, as presented in the 2015 presidential election results, should be a major concern for researchers. Scholars can examine the significance, with respect to malpractices, of this wide margin on the result of the presidential elections.

2. In addition, this research work focused on data collated at the state level rather than the wards or polling units. While this might not be a limitation for the application of Benford's Law and Monte Carlo simulation which require data sets to have four or more digits for a good fit, it is a limitation for the application of the last digits test proposed by Beber and Scacco (2008). Future studies can therefore consider focusing on electoral malpractices within a state or local government area, instead of within the entire country. Although Benford's distribution and Monte Carlo simulation can help to suggest malpractices in election results collated at the state level, the method proposed by Beber and Scacco should help to give more insight into malpractices in election results collated at the wards or polling units. It is assumed that manipulators of results at the polling units and the ward focused on the last digits, favor some numerals over others, underestimate the likelihood of digit repetition in sequences of random integers, demonstrate a preference for pairs of adjacent digits and avoid pairs of distant numerals.
3. Furthermore, this research work concentrated on patterns in digits' distribution of vote counts and voters' turnout to detect suspicious data. The vote of each electorate is expected to be an independent random value, with negligible effect on the final collation. However, it can be argued that voters' preference as a result of gerrymandering and strategic voting pattern of a group could distort the pattern of distribution. This could lead to an inaccurate conclusion that errorhas occurred in an error-free election. In Nigeria, observation shows that some voters cast their ballots for a particular candidate because their community or ethnic group has a preference for such candidate. This implies that the voter's choice is not determined by the voter's own preferences but by the preferences of a community or group. Scholars could look into providing a theoretical basis for the influence of voters' preferences on digits' distribution of election results.
4. This study basically considered Benford's Law and Monte Carlo simulation as election forensic techniques. However, findings from other scholars tend to suggest that there are other forensics approaches, apart from the agent-based and Zipfian models, that could be

applied to studying election malpractices. Further studies can look into these additional forensic techniques in order to have more insights on electoral malpractices in Nigeria.

REFERENCES

Addis, C. L. 2009. Iran's 2009 presidential elections. Congressional Research Service Report. Retrieved October 20, 2014, from https://www.fas.org/sgp/crs/mideast/R40653.pdf.

Adigwe, I. 2016. Lotka's Law and productivity patterns of authors in biomedical science in Nigeria on HIV/AIDS. *The Electronic Library* 34.5: 789-807.

Afe, A. E. 2015. Governorship election litigation and political stability in Ondo State, Nigeria, 1983-2013. *International Journal of Arts and Sciences* 8.2: 139-153.

Agbaje, A. and Adejumobi, S. 2006. Do votes count? The travails of electoral politics in Nigeria. *Africa Development* 31.3:25-44.

Alli, Y. 2015. INEC: Nigerians'll vote in 119,973 polling units. *The Nation*. Retrieved August 29, 2015, from http://thenationonlineng.net/inec-nigeriansll-vote-119973-polling-units-2/.

Alvarez, R.M., Hall, T.E.and Hyde, S.D.2008. *Election fraud: detecting and deterring electoral manipulation*. Washington, DC: Brookings Institution Press. 149-161.

Andriotis, P., Oikonomou, G. and Tryfonas, T. 2013. Jpeg steganography detection with Benford's Law. *Digital Investigation* 9: 246-257.

Animashaun, K. 2010. Regime character, electoral crisis and prospects of electoral reform in Nigeria. *Journal of Nigeria Studies* 1.1: 1-33.

Awopeju, A. 2011. Election rigging and the problems of electoral act in Nigeria. *Afro Asian Journal of Social Sciences* 2: 1-17.

Beber, B. and Scacco, A. 2008. *What the numbers say: A digit-based test for election fraud using new data from Nigeria.* Paper prepared at the Annual Meetingof the American Political Science Association, Boston, MA.

Benenson, I., Martens, K. and Birfir, S. 2008. ParkAgent: an agent-based model of parking in the city. *Computers, Environment and Urban Systems* 32: 431-439.

Benford, F. 1938. The law of anomalous numbers. *Proceedings of the AmericanPhilosophical Society* 78.4: 551-572.

Berdufi, D. 2014. Statistical detection of vote count fraud (2009 Albanian parliamentary election and Benford's Law). *Mediterranean Journal of Social Sciences* 5.2: 755-771.

Birch, S. 2007. Electoral systems and election misconduct.*Comparative Political Studies* 40.12: 1533–1556.

Blydenburg, J. C. (1976). An application of game theory to political campaign decisionmaking. *American journal of political science* 20.1: 51-65.

Bradford, S. C. 1934. Sources of information on specific subjects. *Engineering*137: 85-86.

Breunig, C. and Goerres, A. 2011. Searching for electoral irregularities in an established democracy: applying Benford's Law tests to Bundestag elections in unified Germany. *Electoral studies* 30.3: 534-545.

Cantu, F. and Saiegh, S. M. 2011. Fraudulent democracy? An analysisof Argentina's infamous decadeusing supervised machine learning.*Political Analysis* 19.4: 409–433.

Casti, J. L. 1997. *Would-be-worlds: how simulation is changing the frontiers of science*.New York: Wiley.

Chen, Y. and Leimkuhler, F. F. 1986. A relationship between Lotka's law, Bradford's law, and Zipf's law. *Journal of the American Society for Information Science* 37.5: 307-314.

Crooks, A. T., Castle, C. J. E., and Batty, M. 2008. Key challenges in agent-based modelling forgeo-spatial simulation. *Computers, Environment and Urban Systems 32*.6: 417−430.

Crost, B., Felter, J. H., Mansour, H. and Rees, D. I. 2013. Election fraud and post-election conflict: evidence from the Philippines. *Discussion Paper* No. 7469.

Deckert, J., Myagkov, M. and Ordeshook, P. C. 2011. Benford's Law and the detection of election fraud.*Political Analysis*19: 245-268.

Durtschi, C., Hillison, W. and Pacini, C. 2004. The effective use of Benford's Law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting* 17-34.

Edoh, H. 2004. Corruption: political parties and the electoral process in Nigeria. In M. Jibo and A.T. Simbine. Eds.*Contemporary Issues in Nigerian Politics*,Ibadan: JODAD Publication.

Epstein, J. M. and Axtell, R. 1996. *Growing artificial societies: social science from the bottom up*. Cambridge, MA: MIT Press.

Enabulele, A. O. and Ewere, A. O. 2010. When the shepherd becomes the wolf: what is wrong with the Independent National Electoral Commission of Nigeria? *Journal of Politics and Law* 3.2: 184-198.

European Union Election Observation Mission.April 13, 2015. Second preliminary statement on 2015 general elections of the federal republic of Nigeria. Retrieved March 23, 2015, from http://eueom.eu/files/pressreleases/english/130415-nigeria-ps2_en.pdf.

Fewster, R. M. 2009. Teachers corner: A simple explanation of Benford's Law. *The American Statistician* 63.1: 26-32

Filippov, M. and Ordeshook, P. C. 1996. Fraud or fiction: who stole what in Russia's December 1993 elections. *Social Science Working Paper* 963, California Institute of Technology.

Fowler, J. H. and Smirnov, O. 2005. Dynamic parties and social turnout: an agent-based model. *American Journal of Sociology* 110.4: 1070-1094.

Frequencies. 2002. Zipf ranking, Benford's Law, and Ratio analysis. *The Journal of Size Law Applications* 1.2: 1-24.

Green, R.2014. Rethinking Transparency in U.S. Elections. *Faculty Publications.* Paper 1724.Retrieved September 16, 2017, fromhttp://scholarship.law.wm.edu/facpubs/1724.

Hazell, R. 2010. *Fixed term parliament*. London: The Constitution Unit, University College London.

Heppenstall, A. J., Crooks, A. T., See, L. M. and Batty, M. 2012. *Agent-based models of geographical systems*. Dordrecht: Springer. 85-105.

Howard, M. M. and Roessler, P. G. 2006. Liberalizing electoral outcomes in competitive authoritarian regimes. *American Journal of Political Science* 50.2: 365-381.

Ichino, N. and Schundeln, M. 2012. Deterring or displacing electoral irregularities? spillover effects of observers in a randomized field experiment in Ghana. *The journal of politics* 74.1: 292-307.

Jimenez, R. 2011. Forensic analysis of the Venezuelan recall referendum. *Statistical Science* 26.4: 564-583.

Johnson, P. E. 2011. *Monte Carlo analysis in academic research*. Department of Political Science and Center for Research Methods and Data Analysis, University of Kansas. Retrieved March 23, 2015, from http://pj.freefaculty.org/ps909/Johnson-mc-20110428.pdf.

Judge, G. and Schechter, L. 2009. Detecting problems in survey data using Benford's Law. *Journal of Human Resources*, 44.1.

Kellerman, L. 2014. *Evaluating the effectiveness of Benford's Law as an investigative tool for forensic accountants*. Dissertation submitted for the degree of Magister Commerciiin Forensic Accountancy at the Potchefstroom Campus of the North-West University, South Africa.

Kelley, J. 2012. *Monitoring democracy: when international election observation works, and why it often fails*. Princeton,NJ: Princeton University Press.

Kia, B. 2013. Electoral corruption and democratic sustainability in Nigeria. *IOSR Journal of Humanities And Social Science* 17.5: 42-48.

Klimek, P., Yegorov, Y., Hanel, R. and Thurner, S. 2012. Statistical detection of systematic election irregularities. *PNAS* 109.41: 16469-16473

Kobak, D., Shpilkin, S. and Pshenichinkov, M. 2014. *Integer percentages as electoral falsification fingerprints*. Preprint arXiv: 1410.6059.

Ladan-Baki, I. S. 2016. Electoral violence and 2015 general elections in Nigeria. *Global journal of human-social science* 16.1: 23-29.

Laver,M. and Sergenti,E. 2012.*Party competition: an agent-based model*. Princeton University Press: Princeton, NJ.

Leemann, L. and Bochsler, D. 2014. A systematic approach to study electoral fraud. *Electoral Studies* 35:33-47

Lehoucq, F. 2003. Electoral fraud: cause, types, and consequences.*Annual Review of Political Science* 6: 233–256.

Leimkuhler, F. F. 1967. The Bradford distribution. *Journal of Documentation* 23: 197-207.

Levin, I., Cohn, G. A., Ordeshook, P. C. and Alvarez, R. M. 2009. *Detecting voter fraud in an electronic voting context: an analysis of the unlimited reelection vote in Venezuela*. Online Proceedings of the Electronic Voting Technology Workshop.

Levin, I. and Alvarez, R. M. 2013.*Introduction to the virtual issue: election fraud and electoral integrity.*Virtual Issue of Political Analysis. Retrieved April 12, 2015, from http://www.oxfordjournals.org/our_journals/polana/virtual_issue_efraud.pdf.

Lotka, A. J. 1926. The frequency of distribution of scientific productivity. *Journal of the Washington Academy of Science* 16.12: 317-323.

Macal, C. M. and North, M. J. 2006. *Tutorial on agent-based modelling and simulation part 2: how to model with agents*. Argonne National Laboratory, Argonne.

MacGregor, J. N., Edward, P. C, and Thomas, C. O. 2006. Acomparison of heuristic and human performance on open versions of the traveling salesperson problem.*Journal of Problem Solving* 1.1: 33-43.

Maendeleo Policy Forum 2016. Elections in the Sub-Saharan Africa: general trends, challenges and opportunites. Retrieved September 15, 2017, from http://www.africa.undp.org/.../undp-rba-maendeleo-elections-subsahara-africa-2016.pdf.

Magaloni, B. 2010. The game of electoral fraud and the ousting of authoritarian rule. *American Journal of Political Science* 54.3: 751-765.

Maria, A. 1997. Introduction to modelling and simulation. *Proceedings of the 1997 Winter Simulation Conference*. U.S.A. 7-13.

Mebane, W. R. 2004. The wrong man is president! overvotes in the 2000 presidential election in Florida. *Perspectives on Politics* 2.3: 525-535.

Mebane, W. R. 2006. *Election forensics: vote counts and Benford's Law*.Political Methodology Society. University of California at Davis.

Mebane, W. R. 2008. *Election forensics: the second-digit Benford's Law test and recent American elections*. In R.M.Alvarez, T.E.Hall and S. D. Hyde. 2008.*Election fraud: detecting and deterring electoral manipulation*. Brookings Institution Press, Washington, D.C.

Mebane, W. R. 2010a. A layman's guide to statistical election forensics.*Election Guide Digest.* Washington DC: IFES.

Mebane, W. R.2010b. *Election fraud or strategic voting?* Paper presented at the 2010 annual meeting of the Midwest Political Science Association, Chicago, IL.

Mebane, W. R. 2013. *Election Forensics: The Meanings of Precinct Vote Counts' Second Digits*. Paper presented at the 2013 Summer Meeting of the Political MethodologySociety, University of Virginia, July 18-20, 2013.

Myagkov, M., Ordeshook, P. C., and Shaikin, D. 2009. *The forensics of election fraud: with applications to Russia and Ukraine.* New York: Cambridge University Press.

Nagaraju, V., VIjender, C., Bikshapathi, V. and Lingam, P. R. 2012. Application of game theory to assess conflicting election strategies – an empirical study. *International eJournal of Mathematics and Engineering* 196: 1875-1880.

National Democratic Institute. April, 2008. Final NDI report on Nigeria's 2007 elections. Retrieved July 28, 2017, from http://www.ndi.org/files/2313_ng_report_election07_043008.pdf.

Nigerian Muse.April 19, 2011. Details of 2011 presidential election results in Nigeria. Retrieved July 28, 2017, from http://www.nigerianmuse.com/20110419040622zg/sections/general-articles/details-of-2011-presidential-election-results-in-nigeria/.

Nigrini, M.J. (1996). Taxpayer compliance application of Bendford's law. *Journal of the American Taxation Association*. 18.1: 72-92.

Nigrini, M. J. 2000. Digital analysis using Benford's Law. Test and statistics for auditors. Vancouver, BC: Global audit publications.

North, M. J., and Macal, C. M. 2007. *Managing business complexity: discovering strategic solutionswith agent-based modelling and simulation*. New York: Oxford University Press.

Ogbeidi, M. M. 2010. A culture of failed elections: revisiting democratic elections in Nigeria, 1959-2003. *HAOL* 21: 43-56.

Ojo, E. O. 2012. Leadership crisis and political instability in Nigeria, 1964-1966: the personalities, the parties and the policies. *Global Advanced Research Journal of History, Political Science and International Relation* 1.1: 006-017.

Okeyim, M. O., Adams, J. A. and Ojie, P. 2012. Managing election violence in Nigeria. *International Journal of Arts and Sciences* 5.2: 43-48.

Oleg, K. 2011. Theory and practice of falsified elections. *Munich Personal RePEc Archive*. Paper number 35543.

Omodia, S. M. 2012. Election, governance and the challenge of national integration in the Nigerian fourth republic. *British Journal of Arts and Social Sciences* 5.2: 307-314.

Onebamhoi, O. N. 2011. Curbing electoral violence in Nigeria: the imperative of political education. *International Multidisciplinary Journal, Ethiopia* 5.5: 99-110.

Onoja, A. 2009. Nature, character and outcomes of post election challenges in Nigeria. *Journal of the Historical Society of Nigeria* 18: 79-104.

Oromareghake, P. B. 2013. Electoral institutions/processes and democratic transition in Nigeria under the fourth republic. *International Review of Social Sciences and Humanities* 6.1: 19-34.

Osimen, G. U. and Ologunowa, C. S. 2013. 2011 post-electoral violence in Nigeria: lesson for 2015 general election. *American International Journal of Research in Humanities, Arts and Social Sciences* 4.1: 45-55.

Osinakachukwu, N. P. and Jawan, J. A. 2011. The electoral process and democratic consolidation in Nigeria. *Journal of Politics and Law* 4.2: 128-138.

Pericchi, L. and Torres, D. 2011. Quick anomaly detection by the Newcomb-Benford law, with applications to electoral process data from the USA, Puerto Rico and Venezuela. *Statistical Science* 26.4: 502-516.

Petucci, 2005. Benford's Law: can it be used to detect irregularities in first party automobile insurance claims? *Journal of economic crime management* 3.1: 1-35

Phelps, S. 2012. Applying dependency injection to agent-based modelling: the JABM framework. *CCFEA Working Paper*, WP056-12.

Qiu, L. and Phang, R. 2020. Agent-based modelling in political decision making. Oxford Encyclopedia of Political Decision Making. Oxford University Press.

Quratul-Ann, M. 2013. Virtual Pakistan 2013: an agent-based modelling analysis. *College Undergraduate Research Electronic Journal*, University of Pennsylvania ScholarlyCommons.

Raimi, R. 1976. The first digit problem.*American Mathematical Monthly* 83.7: 521–538.

Rivest, R. L. and Shen, E. 2012. A Bayesian method for auditing elections.*Proceedings of the 2012 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE '12)*. USENIX.

Roukema, B. F. 2014. A first-digit anomaly in the 2009 Iranian presidential election. *Journal of Applied Statistics* 41.1: 164-199.

Schelling, T. C. 1978. *Micromotives and macrobehaviour.* New York: Norton.

Sheriff,G. I., Abdullahi, N. L. and Kabir, M. 2015. The 2015 general elections: A review of major determinants of paradigm shift in voting behaviour and political participation in Nigeria. *International Journal of Research in Humanities and Social Studies* 2.9: 8-16.

Simidchieva, B. I., Engle, S. J., Clifford, M., Jones, A. C., Peisert, S., Bishop, M., Clarke, L. A., and Osterweil, L. J. 2010. Modelling and analyzing faults to improve election process robustness. In Proceedings of the 2010 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE '10).

Udu, L. E. 2015. INEC and the 2015 general elections in Nigeria: matters arising. *Research on Humanities and Social Sciences* 5.12: 96-108.

United States Agency for International Development 2017. A guide to election forensics: Research and innovation grants working papers series. Retrieved September 17, 2018, from https://pdf.usaid.gov/pdf_docs/PA00MXR7.pdf.

Wattenberg, F. and Szabo, C. 2013. *Agent-based and community, modelling, Netlogo and the 2012 presidential election*. Paper presented at the Twenty-fifth Annual International Conference on Tehnology in Collegiate Mathematics, Boston, Massachusetts, March 21-24, 2013.

Wilensky, U.and Rand, W. 2011.*An introduction to agent-based modelling: modelling natural, social and engineered complex systems with NetLogo*. MIT Press, Cambridge, MA.

Yagboyaju, D. A. 2011. Nigeria's fourth republic and the challenge of a faltering democratization. *African Studies Quarterly* 12.3: 93-106.

Yusuf, I. and Zaheruddin, O. 2015. Challenges of electoral processes in Nigeria's quest for democratic governance in the fourth republic. *Research on humanities and social sciences* 5.22: 1-10.

Zipf, G. K. 1934. *Human behaviour and the principle of least effort*. Addison-Wesley, Reading: Massachusetts.

APPENDICES

Appendix 1: An official letter of request for data collection

UNIVERSITY OF IBADAN, NIGERIA
AFRICA REGIONAL CENTRE FOR INFORMATION SCIENCE

25 Years Of Academic And Research Excellence

DIRECTOR:
Wole Michael OLATOKUN, Ph.D
B.A, MInfSc, Ph.D (Ibadan.)

OFFICE OF THE DIRECTOR
Tel: 08059767412, 08165848675
E-mail: woleabbeyolatokun@yahoo.co.uk
Secretariat: 09037397987

16 December, 2015

INDEPENDENT NATIONAL ELECTORAL COMMISSION
RECEIVED
- 5 JAN 2015
OFFICE OF THE HON. CHAIRMAN

The Chairman
Independent National Electoral Commission,
FCT, Abuja.

Dear Sir,

LETTER OF INTRODUCTION

This is to introduce to you **TUNMIBI, Sunday Olapade** with **Matric No 147161** who is a PhD student of the Africa Regional Centre for Information Science (ARCIS). ARCIS is a Centre in the University of Ibadan established in 1990 to train high level manpower, undertake research, and provide information system development and management services to organisation in Nigeria and other West Africa countries.

Students in the (PhD) programme are required, as part of their training, to undertake academic projects, term papers and assignments that entail understudying the information systems/services of public and private sector organisations in Nigeria and other countries.

It is in this connection that the above-named student needs to obtain or collect data and information about or from your institution for his project titled **"APPLICABILITY OF AGENT-BASED AND ZIPFIAN MODELING IN FORENSIC ANALYSIS OF ELECTION RESULTS IN NIGERIA"**

I request you kindly give him the necessary assistance possible.

Thank you.

Dr. W. M. Olatokun
Ag. Director

DIRECTOR
AFRICA REGIONAL CENTRE FOR INFORMATION SCIENCE
(A.R.C.I.S)
UNIVERSITY OF IBADAN, NIGERIA

25th Anniversary 1990 - 2015

No 6 Benue Road, P. O. Box 22133. University of Ibadan, Ibadan, Nigeria
Email: dir_arcis@ui.edu.ng Website: www.arcis.ui.edu.ng

VISION: To become a global frontline centre of excellence for human resource development, research and service delivery in information science and technology

Appendix 2:Correlation matrix for the 2015 and 2011 presidential elections

2015 presidential election

Registered and accredited voters

Descriptive Statistics

	Mean	Std. Deviation	N
Registered voters in 2015	1822216	1026219.013	37
Accredited voters in 2015	858013.24	464121.161	37

Correlations[a]

		Registered voters in 2015	Accredited voters in 2015
Registered voters in 2015	Pearson Correlation	1	.859**
	Sig. (2-tailed)		.000
Accredited voters in 2015	Pearson Correlation	.859**	1
	Sig. (2-tailed)	.000	

**. Correlation is significant at the 0.01 level (2-tailed).

a. Listwise N=37

Votes cast for APC, PDP and other political parties

Descriptive Statistics

	Mean	Std. Deviation	N
APC's votes in 2015	416889.76	413798.615	37
PDP's votes in 2015	347382.76	315993.010	37
Votes of other parties in 2015	8364.35	8233.647	37

Correlations[a]

		APC's votes in 2015	PDP's votes in 2015	Votes of other parties in 2015
APC's votes in 2015	Pearson Correlation	1	-.350*	.034
	Sig. (2-tailed)		.033	.840
PDP's votes in 2015	Pearson Correlation	-.350*	1	.093
	Sig. (2-tailed)	.033		.586
Votes of other parties in 2015	Pearson Correlation	.034	.093	1
	Sig. (2-tailed)	.840	.586	

*. Correlation is significant at the 0.05 level (2-tailed).

a. Listwise N=37

Valid, rejected and total votes cast

Descriptive Statistics

	Mean	Std. Deviation	N
Valid votes cast in 2015	772636.86	424540.440	37
Rejected votes in 2015	22824.84	11468.557	37
Total votes cast in 2015	795461.70	431062.040	37

Correlations[a]

		Valid votes cast in 2015	Rejected votes in 2015	Total votes cast in 2015
Valid votes cast in 2015	Pearson Correlation	1	.560**	1.000**
	Sig. (2-tailed)		.000	.000
Rejected votes in 2015	Pearson Correlation	.560**	1	.578**
	Sig. (2-tailed)	.000		.000
Total votes cast in 2015	Pearson Correlation	1.000**	.578**	1
	Sig. (2-tailed)	.000	.000	

**. Correlation is significant at the 0.01 level (2-tailed).

a. Listwise N=37

2011 presidential election

Votes cast for PDP, CPC, ACN, ANPP and other political parties

Descriptive Statistics

	Mean	Std. Deviation	N
PDP's votes in 2011	607978.03	429220.489	37
CPC's votes in 2011	330131.16	438261.432	37
ACN's votes in 2011	56191.92	97623.515	37
ANPP's votes in 2011	24784.11	88154.503	37
Votes for other parties in 2011	13645.03	10644.552	37

Correlations[a]

		PDP's votes in 2011	CPC's votes in 2011	ACN's votes in 2011	ANPP's votes in 2011	Votes for other parties in 2011
PDP's votes in 2011	Pearson Correlation	1	-.236	.019	-.138	-.229
	Sig. (2-tailed)		.160	.911	.416	.173
CPC's votes in 2011	Pearson Correlation	-.236	1	-.220	.526**	.601**
	Sig. (2-tailed)	.160		.191	.001	.000
ACN's votes in 2011	Pearson Correlation	.019	-.220	1	-.051	.336*
	Sig. (2-tailed)	.911	.191		.765	.042
ANPP's votes in 2011	Pearson Correlation	-.138	.526**	-.051	1	.433**
	Sig. (2-tailed)	.416	.001	.765		.007
Votes for other parties in 2011	Pearson Correlation	-.229	.601**	.336*	.433**	1
	Sig. (2-tailed)	.173	.000	.042	.007	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. Listwise N=37

Appendix 3: R programming codes and tables for the Monte Carlo simulation

2011 presidential elections

```
mydata = read.csv('d2011.csv', header=T)
head(mydata)
library(mc2d)
attach(mydata)
names(mydata)
#... PERT distribution
#... CPC vote
macmum<-min(mydata[,5])
mimum<-min(mydata[,6])
exp <- median(mydata[,3])

n = 37
set.seed(30)
samples <-rpert(n, min=mimum, mode=exp, max=macmum, shape=4)

mean(samples)
CPC_vote <- replicate(1000, mean(rpert(n, min=mimum, mode=exp, max=macmum, shape=4)))
summary(CPC_vote)

#..... Loop
n = 37
for(i in n){
CPC_vote <- replicate(1000, rpert(n, min=mimum, mode=exp, max=macmum, shape=4))
}

s = seq(1,37,1)
for(i in s){
CPC_state[i]= median(CPC_vote[i,])
}

mydata = read.csv('d2011.csv', header=T)
head(mydata)
library(mc2d)
attach(mydata)
names(mydata)
#... PERT distribution
#... PDP vote
macmum<-min(mydata[,5])
mimum<-min(mydata[,6])
exp <- median(mydata[,3])

n = 37
```

```
set.seed(30)
samples <-rpert(n, min=mimum, mode=exp, max=macmum, shape=4)

mean(samples)
PDP_vote <- replicate(1000, mean(rpert(n, min=mimum, mode=exp, max=macmum, shape=4)))
summary(PDP_vote)

#..... Loop
n = 37
for(i in n){
PDP_vote <- replicate(1000, rpert(n, min=mimum, mode=exp, max=macmum, shape=4))
}

s = seq(1,37,1)
for(i in s){
PDP_state[i]= median(PDP_vote[i,])
}
```

2015 presidential elections

```
mydata = read.csv('d2015.csv', header=T)
head(mydata)
library(mc2d)
attach(mydata)
names(mydata)
#... PERT distribution
#... APC vote
macmum<-min(mydata[,5])
mimum<-min(mydata[,6])
exp <- median(mydata[,3])

n = 37
set.seed(30)
samples <-rpert(n, min=mimum, mode=exp, max=macmum, shape=4)

mean(samples)
APC_vote <- replicate(1000, mean(rpert(n, min=mimum, mode=exp, max=macmum, shape=4)))
summary(CPC_vote)

#..... Loop
n = 37
for(i in n){
APC_vote <- replicate(1000, rpert(n, min=mimum, mode=exp, max=macmum, shape=4))
}
```

```
s = seq(1,37,1)
for(i in s){
APC_state[i]= median(APC_vote[i,])
}

mydata = read.csv('d2015.csv', header=T)
head(mydata)
library(mc2d)
attach(mydata)
names(mydata)
#... PERT distribution
#... PDP vote
macmum<-min(mydata[,5])
mimum<-min(mydata[,6])
exp <- median(mydata[,3])

n = 37
set.seed(30)
samples <-rpert(n, min=mimum, mode=exp, max=macmum, shape=4)

mean(samples)
PDP_vote <- replicate(1000, mean(rpert(n, min=mimum, mode=exp, max=macmum, shape=4)))
summary(PDP_vote)

#..... Loop
n = 37
for(i in n){
PDP_vote <- replicate(1000, rpert(n, min=mimum, mode=exp, max=macmum, shape=4))
}

s = seq(1,37,1)
for(i in s){
PDP_state[i]= median(PDP_vote[i,])
}
```

Simulation of the vote counts of the 2011 presidential election

	STATES	CPC_state (simulated)	PDP_state (simulated)	Actual CPC	Actual PDP
1	EKITI	1502	331365	2689	135009
2	LAGOS	14448	844569	189983	1281688
3	OGUN	4758	175810	17654	309177
4	ONDO	3599	308235	11890	387376
5	OSUN	2027	69235	6997	188409
6	OYO	15672	272123	92396	484758
7	ABIA	466	1163763	3743	1175984
8	ANAMBRA	534	1133225	4223	1145169
9	EBONYI	301	459283	1025	480592
10	ENUGU	675	790452	3753	802144
11	IMO	792	1353440	7591	1381357
12	AKWA IBOM	637	1102480	5348	1165629
13	BAYELSA	199	502936	691	504811
14	C/RIVER	813	692819	4002	709382
15	DELTA	934	1348028	8960	1378851
16	EDO	4167	473206	17795	542173
17	RIVERS	1050	1782121	13182	1817762
18	JIGAWA	85565	154083	663994	419252
19	KADUNA	76016	551185	1334244	1190179
20	KANO	89148	72641	1624543	440666
21	KATSINA	104295	111934	1163919	428392
22	KEBBI	79160	147503	501453	369198
23	SOKOTO	86816	104985	540769	309057
24	ZAMFARA	96618	60584	624515	238980
25	ADAMAWA	55385	284655	344526	508314
26	BAUCHI	119673	41471	1315209	258404
27	BORNO	113486	36412	909763	207075
28	GOMBE	88348	109480	459898	290347
29	TARABA	51165	275646	257986	451354
30	YOBE	79871	22052	337537	117128
31	BENUE	15191	460733	109680	694776
32	KOGI	34572	284546	132201	399816
33	KWARA	29776	173487	83603	268243
34	NASARAWA	58336	240852	278390	408997
35	NIGER	93814	101374	652574	321429
36	PLATEAU	37283	751619	356551	1029865
37	FCT	48398	161354	131576	253444

Simulation of the vote counts of the 2015 presidential election

	STATES	APC_state (simulated)	PDP_state (simulated)	APC Actual	PDP Actual
1	EKITI	109610	279420	120331	176466
2	LAGOS	150969	202351	792460	632327
3	OGUN	156778	183490	308290	207950
4	ONDO	146199	213211	299889	251368
5	OSUN	164765	168362	383603	249929
6	OYO	161850	159537	528620	303376
7	ABIA	9521	433506	13394	368303
8	ANAMBRA	7273	471481	17926	660762
9	EBONYI	14092	394695	19518	323653
10	ENUGU	6858	483090	14157	553003
11	IMO	51927	362648	133253	559185
12	AKWA IBOM	16179	475215	58411	953304
13	BAYELSA	3953	459925	5194	361209
14	CROSS RIVER	17266	417216	28368	414863
15	DELTA	10786	458413	48910	1211405
16	EDO	113077	254657	208469	286869
17	RIVERS	12483	459863	69238	1487075
18	JIGAWA	234779	64019	885988	142904
19	KADUNA	194090	134433	1127760	484085
20	KANO	249204	46618	1903999	215779
21	KATSINA	258460	31236	1345441	98937
22	KEBBI	226489	67484	567883	100972
23	SOKOTO	219098	81876	671926	152199
24	ZAMFARA	221945	87940	612202	144833
25	ADAMAWA	161655	185874	374701	251664
26	BAUCHI	255938	38416	931598	86085
27	BORNO	261526	24031	473543	25640
28	GOMBE	216682	96694	361245	96873
29	TARABA	124428	240350	261326	310800
30	YOBE	258844	24709	446265	25526
31	BENUE	151602	199528	373961	303737
32	KOGI	170476	167313	264851	149987
33	KWARA	185612	133556	302146	132602
34	NASARAWA	129137	255019	236838	273460
35	NIGER	221465	83864	657678	149222
36	PLATEAU	122561	264464	429140	549615
37	FCT	131188	242561	146399	157195

Appendix 4:Detailed result of the 2015 presidential election

INDEPENDENT NATIONAL ELECTORAL COMMISSION
2015 PRESIDENTIAL ELECTION
MARCH 28, 2015
SUMMARY OF RESULTS

S/N	NAME OF STATE	CODE	NO OF REGD. VOTERS	NO OF ACCRED. VOTERS	AA	ACPN	AD	ADC	APA	APC	CPP
					VOTES RECEIVED BY PARTIES						
1	ABIA	AB	1,349,134	442,538	315	2,194	448	569	2,766	13,394	1,046
2	ADAMAWA	AD	1,518,123	709,993	495	1,166	595	1,012	1,549	374,701	839
3	AKWA IBOM	AK	1,644,481	1,074,070	1,600	443	474	608	384	58,411	412
4	ANAMBRA	AN	1,963,427	774,430	547	1,259	475	534	2,303	17,926	1,279
5	BAUCHI	BA	2,053,484	1,094,055	131	232	173	189	964	931,598	991
6	BAYELSA	BY	605,637	384,789	45	38	69	118	70	5,194	44
7	BENUE	BN	1,893,596	754,634	315	1,464	254	539	945	373,961	567
8	BORNO	BO	1,799,669	544,759	145	243	393	201	878	473,543	310
9	CROSS RIVER	CR	1,144,288	500,577	279	514	705	749	582	28,368	381
10	DELTA	DT	2,044,372	1,350,914	1,475	916	735	888	478	48,910	813
11	EBONYI	EB	1,071,226	425,301	426	1,214	1,133	2,704	2,452	19,518	2,345
12	EDO	ED	1,650,552	599,166	159	1,284	450	512	709	208,469	325
13	EKITI	EK	723,255	323,739	94	538	864	474	482	120,331	330
14	ENUGU	EN	1,381,563	616,112	441	479	260	478	715	14,157	237
15	GOMBE	GM	1,110,105	515,828	104	192	169	247	773	361,245	407
16	IMO	IM	1,747,681	801,712	389	956	797	1,617	2,256	133,253	733
17	JIGAWA	JG	1,815,839	1,153,428	394	540	587	375	2,527	885,988	1,553
18	KADUNA	KD	3,361,793	1,746,031	218	474	273	546	1,613	1,127,760	824
19	KANO	KN	4,943,862	2,364,434	436	778	708	657	2,770	1,903,999	1,552
20	KATSINA	KT	2,842,741	1,578,646	183	402	283	498	1,671	1,345,441	976
21	KEBBI	KB	1,457,763	792,817	214	361	450	472	2,685	567,883	1,794
22	KOGI	KG	1,350,883	476,839	700	1,089	427	761	1,001	264,851	967
23	KWARA	KW	1,181,002	489,360	249	817	530	488	1,165	302,146	910
24	LAGOS	LA	5,827,846	1,678,754	1,785	3,038	4,453	2,072	2,177	792,460	1,125
25	NASARAWA	NS	1,222,054	562,059	40	96	74	105	310	236,838	131
26	NIGER	NG	1,995,679	933,607	307	441	403	614	2,006	657,678	1,264
27	OGUN	OG	1,709,409	564,975	584	3,072	1,927	1,964	1,930	308,290	978
28	ONDO	OD	1,501,549	618,040	386	2,400	1,397	1,227	1,139	299,889	1,012
29	OSUN	OS	1,378,113	683,169	373	1,781	1,667	937	1,306	383,603	1,028
30	OYO	OY	2,344,448	1,073,849	6,331	8,979	6,282	5,000	4,468	528,620	6,674
31	PLATEAU	PL	1,977,211	1,076,833	178	381	279	406	618	429,140	237
32	RIVERS	RV	2,324,300	1,643,409	1,096	525	1,104	1,091	513	69,238	577
33	SOKOTO	SO	1,663,127	988,899	249	595	714	702	3,462	671,926	1,894
34	TARABA	TR	1,374,307	638,578	902	811	586	323	1,306	261,326	1,033
35	YOBE	YB	1,077,942	520,127	101	164	213	112	632	446,265	329
36	ZAMFARA	ZF	1,484,941	675,049	125	238	290	294	1,310	612,202	655
37	FCT	FCT	886,573	344,056	138	342	240	283	674	146,399	347
	TOTAL		67,422,005	31,746,490	22,125	40,331	30,673	29,665	33,537	15,424,921	36,300
					AA	ACPN	AD	ADC	APA	APC	CPP

S/N	NAME OF STATE	HOPE	KOWA	NCP	PDP	PPN	UDP	UPP	NO OF VALID VOTES	NO OF REJECTED VOTES	TOTAL VOTES CAST
1	ABIA	125	173	745	368,303	424	213	330	391,045	10,004	401,049
2	ADAMAWA	267	752	1,212	251,664	1,163	289	334	636,018	25,192	661,210
3	AKWA IBOM	192	160	381	953,304	327	224	144	1,017,064	11,487	1,028,551
4	ANAMBRA	357	311	887	660,762	537	265	1,121	688,584	14,825	703,409
5	BAUCHI	46	128	207	86,085	128	79	87	1,020,338	19,437	1,039,775
6	BAYELSA	18	52	95	361,209	62	30	35	367,067	4,672	371,739
7	BENUE	115	105	683	303,737	439	66	74	683,264	19,867	703,131
8	BORNO	88	158	107	25,640	149	91	41	501,920	13,088	515,008
9	CROSS RIVER	237	317	890	414,863	864	289	1,487	450,514	15,392	465,906
10	DELTA	166	311	670	1,211,405	393	354	261	1,267,773	17,075	1,284,848
11	EBONYI	989	313	1,850	323,659	1,166	624	4,859	363,888	29,449	393,337
12	EDO	22	175	516	286,869	729	150	77	500,451	22,334	522,785
13	EKITI	94	108	377	176,466	388	60	145	300,691	8,754	309,445
14	ENUGU	110	203	761	553,003	407	1,623	290	573,173	12,459	585,632
15	GOMBE	46	97	227	96,873	157	25	37	460,599	12,845	473,444
16	IMO	157	156	784	559,185	414	264	1,917	702,964	28,957	731,921
17	JIGAWA	357	423	548	142,904	853	338	197	1,037,564	34,325	1,071,889
18	KADUNA	105	176	754	484,085	549	79	78	1,617,482	32,719	1,650,201
19	KANO	292	288	697	215,779	485	234	136	2,128,821	43,626	2,172,447
20	KATSINA	47	215	330	98,937	254	117	72	1,449,426	32,288	1,481,714
21	KEBBI	215	448	519	100,972	547	207	238	677,003	38,119	715,122
22	KOGI	144	190	395	149,987	476	180	156	421,328	17,959	439,287
23	KWARA	118	214	394	132,602	325	81	102	440,080	21,321	461,401
24	LAGOS	255	1,000	1,430	632,327	1,041	269	244	1,443,686	52,289	1,495,975
25	NASARAWA	4	48	222	273,460	164	23	33	511,547	10,094	521,641
26	NIGER	186	305	550	149,222	449	116	118	813,671	31,012	844,683
27	OGUN	380	432	815	207,950	4,339	567	597	533,172	26,443	559,615
28	ONDO	189	223	848	251,368	734	184	321	561,056	21,379	582,435
29	OSUN	132	255	767	249,929	599	124	159	642,615	20,758	663,373
30	OYO	838	1,312	1,895	303,376	2,842	1,069	3,665	881,352	47,254	928,606
31	PLATEAU	58	138	653	549,615	554	54	29	982,388	18,304	1,000,692
32	RIVERS	542	2,274	565	1,487,075	492	303	156	1,565,461	19,307	1,584,768
33	SOKOTO	283	475	685	152,199	605	269	180	834,259	42,110	876,369
34	TARABA	161	153	873	310,800	680	124	439	579,677	23,039	602,716
35	YOBE	67	104	120	25,526	101	30	32	473,796	17,971	491,767
36	ZAMFARA	14	122	404	144,833	374	83	58	761,022	19,157	780,179
37	FCT	83	165	473	157,195	269	95	96	306,805	9,210	316,015
	TOTAL	7,435	13,076	24,455	12,853,162	24,475	9,208	18,220	28,587,564	844,519	29,432,083
		HOPE	KOWA	NCP	PDP	PPN	UDP	UPP			

a	TOTAL NUMBER OF REGISTERED VOTERS	67,422,005
b	TOTAL NUMBER OF ACCREDITED VOTERS	31,746,490
c	TOTAL NUMBER OF VALID VOTES	28,587,564
d	TOTAL NUMBER OF REJECTED VOTES	844,519
e	TOTAL NUMBER OF VOTES CAST	29,432,083