Monte Carlo simulation of vote counts from Nigeria presidential elections

Sunday Tunmibi and Wole Olatokun

Abstract: This study was designed to examine the application of Monte Carlo simulation to analyse vote counts from the 2011 and 2015 presidential elections in Nigeria. The study adopted a simulation approach and the data were simulated using the R programming language. The actual and simulated data were analysed using frequency and percentage distribution as well as Spearman Rank Correlation tests. Findings revealed that the vote counts for People’s Democratic Party (PDP) and Congress for Progressive Change (CPC) in the 2011 presidential elections as well as vote counts for All Progressives Congress (APC) and PDP in the 2015 presidential elections do not conform to the distributional pattern of the respective simulated vote counts. A manipulation-free data generating process is expected to produce vote counts that are close to a random distribution. The study recommends the use of forensic analysis by the Independent National Electoral Commission (INEC) for future electoral reform in Nigeria.

Subjects: African Studies; Political Behavior and Participation; Information Technology

ABOUT THE AUTHORS

First author (corresponding author)
Dr. Sunday Tunmibi holds a B.Tech. in Electronics and Electrical Engineering obtained from Ladoke Akintola University of Technology, Ogbomosho, Nigeria and degrees of master and Ph.D. in Information Science obtained from University of Ibadan, Nigeria. His major fields of interest are election forensics, artificial intelligence, data mining and big data analysis. He has authored nine articles in international journals and three in proceedings of international conferences. Email: sundaytunmibi@gmail.com.

Prof. Wole Olatokun is professor of Information Science at the University of Ibadan, Nigeria. He is the former Director of Africa Regional Centre for Information Science, University of Ibadan, and currently Honorary Professor at the Information Studies Programme, School of Social Sciences, University of Kwazulu-Natal and Department of Information and Knowledge Management, University of Johannesburg, South Africa. His areas of research interest include Social Informatics, Knowledge Management, ICT for Development and Digital Scholarship. Email: woleabheyolatokun@yahoo.co.uk

PUBLIC INTEREST STATEMENT

Election forensic is a nascent field in Social Science. It uses statistical methods to verify whether election results are free from anomalies or not. The techniques adopted in election forensic are based on finding patterns in the electoral returns data and the possible deviation of election data from the expected distribution of this pattern. The study adopted secondary data gathered from the electoral body in Nigeria. We adopted agent-based approach (using Monte Carlo method) to detect anomalies in the Nigeria 2011 and 2015 presidential election results. The findings show that the elections conducted in Nigeria are filled with anomalies, as the collated votes for the two leading political parties in both elections deviate from the simulated votes. The study concluded with recommendations and suggestions for policy makers and government on the way forward and necessary steps that could help to avoid future election malpractices.
Keywords: Agent-based model; election forensics; Monte Carlo simulation; vote counts; Nigeria presidential election

1. Introduction

Election is a formal decision-making process by which a population chooses an individual to hold public office by voting. The history of Nigeria’s democratic experiments demonstrates that elections and electoral politics have generated so muchanimosity which has, in some cases, threatened the corporate existence of the country (Animashaun, 2010). Much of this animosity could be traced to unreliability of electoral result returns. Studies show that all elections conducted during the fourth republic in Nigeria (1999–2019) are filled with anomalies (Ogbeidi, 2010; Osinakachukwu & Jawan, 2011; European Union Election Observation Mission (European Union Election Observation Mission. 2015. Second preliminary statement on, 2015).

The 1999 general election was generally considered to be violence-free but not credible due to massive irregularities. There was a wide disparity between the number of voters observed at the polling stations and the final result that was reported from several states (Ogbeidi, 2010). The 2003 election, in the history of Nigeria politics, presented the first peaceful civilian transition. However, the elections were said to be more pervasively and openly rigged than the flawed 1999 polls (Onebamhoi, 2011). More so, this apparently phantom election recorded extraordinary high turnout figures generally in excess of 90 percent (Osinakachukwu & Jawan, 2011).

The 2007 elections were described as the worst in Nigeria’s history ranking among the worst conducted anywhere in the world in recent times (Onebamhoi, 2011). The elections were characterized by late arrival of electoral materials in the various polling units, inadequate polling materials, voter’s registration problems, no secrecy of the ballot, ballot paper problems, snatching of ballot boxes and destruction of ballot materials, violence, use of security agencies to intimidate the voters and rigged elections, no voting in some polling centres, use of government officials to commit electoral fraud, and omissions of some parties’ logo and candidate names on the ballot paper to disenfranchise their supporters (Kia, 2013). Although the 2011 presidential election was commended by observers as one of the most successful in Nigeria’s political history, cases of stuffing ballot boxes, under age voting and outright falsification of election results were also reported in some states (Yusuf & Zaheruddin, 2015). Likewise, the European Union Election Observation Mission. 2015. Second preliminary statement on (2015) reported that the 2015 general elections were marred by malpractices, despite being largely peaceful.

To address the issues of election anomalies, scholars have suggested the application of election forensics to analyse election data (Breunig & Goerres, 2011; Levin & Alvarez, 2013; Mebane, 2006). Election forensic techniques focus on finding patterns in the electoral returns data and the possible deviation of election data from the expected distribution of this pattern (Levin & Alvarez, 2013). Many of the common electoral forensic techniques are based on Benford’s Law while few others focus on Agent-based modelling.

Agent-based modelling is a modelling technology that is ideally suited to investigate outcomes that may emerge when large numbers of boundedly rational agents, using adaptive decision rules selected from a diverse portfolio of possibilities, interact with each other continuously in an evolving dynamic setting (MacGregor et al., 2006). Agent-based modelling has been deployed to analyse electoral outcomes (Cantu & 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 et al., 2012). Other methods are the application of scripting techniques based on object-oriented methods (Wilenisky & Rand, 2011) and the application of the Monte Carlo analysis (Cantu & Saiegh, 2011). 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 discrete-event simulation and Monte-Carlo methods. The Monte-Carlo approach has been specifically adopted in detecting fraudulent electoral data (Kobak et al., 2014).

Monte Carlo analysis is a general term that refers to research that employs random numbers, usually in the form of a computer simulation (Johnson, 2011). This article focuses on the application of Monte Carlo approach of the agent-based model to analyse election data in Nigeria. In this article, the authors examined if the actual vote counts of PDP and CPC in the 2011 presidential elections deviate from the simulated vote counts from the Monte Carlo simulation; and if the actual vote counts of APC and PDP in the 2015 presidential elections deviate from the simulated vote counts from the Monte Carlo simulation. In order to determine the correlation between the actual vote counts and their respective simulated vote counts from the Monte Carlo simulation, we hypothesised that:

\(H_{01}: \text{There is no significant correlation between the digital distribution of vote counts of the 2011 presidential election results in Nigeria and the digital distribution of their respective simulated votes from the Monte Carlo simulation.}\)

\(H_{02}: \text{There is no significant correlation between the digital distribution of vote counts of the 2015 presidential election results in Nigeria and the digital distribution of their respective simulated votes from the Monte Carlo simulation.}\)

2. Literature review

2.1. An overview on election forensics

Election forensics is a name coined to describe a nascent field of social science that intends to develop data analysis tools that can be used to detect discrepancy in election outcomes (Tunmibi & Olatokun, 2021; Levin & Alvarez, 2013; Mebane, 2010). The fundamental difficulty with the study of election fraud is its measurement—it may take many forms and those involved typically wish to hide these illicit activities. Scholars generally rely upon assessments by election observers to measure electoral fraud for quantitative cross-national studies and use media reports and petitions filed by aggrieved parties for single-country studies. But these measures are generated by participants with different interests, expectations, and standards across elections, which raises concerns about consistency and bias. Consequently, current measures used in cross-national studies may underestimate the extent of electoral fraud in new democracies, including those that appear to have fairly clean elections. This underlies the need for substantial caution about the robustness of empirical findings in this emerging large-n literature (Birch, 2007; Ichino & Schundeln, 2012; Kelley, 2012; Lehocq, 2003).

The operational problem in uncovering fraudulent elections is identifying the characteristics that make them distinct from valid ones. Still, it is usually impossible to be absolutely certain about the legitimacy of an election. A particularly appealing option to overcome this difficulty is to tease out possible evidences of fraud from the available data using mathematical algorithms (Cantu & Saiegh, 2011). Monte Carlo simulation provides a probable application of mathematical algorithm that can be used to simulate and analyse election data.

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 simulation...
environment such as represented by the artificial world (Heppenstall et al., 2012). Agents can be representations of animate entities such as humans that can roam freely around an environment or inanimate objects that have fixed locations but can change state. Each of the inanimate and animate agents outlined above can possess rules that will affect their behaviour and relationships with other agents and/or their surrounding environment. Rules are typically derived from published literature, expert knowledge, data analysis or numerical work and are the foundation of an agent’s behaviour. One rule-set can be applied to all agents or each agent (or categories of agents) can have its own unique rule set. These rules are typically based around “if-else” statements with agents carrying out an action once a specified condition has been satisfied. Agents can interact with each other and among themselves and with the environment. This relationship may be specified in a variety of ways, from simply reactive (that is, agents only perform actions when triggered to do so by some external stimulus) to goal-directed (that is, seeking a particular goal). The behaviour of agents can also be scheduled to take place synchronously or asynchronously.

Although there is no universal agreement on the precise definition of the term “agent,” definitions tend to agree on more points than they disagree. The fundamental feature of an agent is the capability of the component to make independent decisions. This requires agents to be active rather than purely passive (Macal & North, 2006). While introducing the need for agent-based model, Macal and North emphasised that the systems to be analysed and modelled are becoming more complex in terms of their interdependencies. More so, traditional modelling tools are no longer as applicable as they once were.

Agent-based model has its direct historical roots in complex adaptive systems (CAS) and the underlying notion that systems are built from the ground-up, in contrast to the top-down systems view taken by Systems Dynamics. CAS concerns itself with the question of how complex behaviours arise in nature among myopic, autonomous agents. In addition, agent-based model tends to be descriptive, with the intent of modelling the actual or plausible behaviour of individuals, rather than normative such as traditional operations research, which seeks to optimise and identify optimal behaviours (Macal & North, 2006).

The first social agent-based simulation was developed by Thomas Schelling (Schelling, 1978). Schelling applied cellular automata to study housing segregation patterns, in which agents represent people and agent interactions represent a socially relevant process. The Schelling model showed that it is possible to have patterns that are not necessarily implied or consistent with the objectives of the individual agents. Some years later, Epstein and Axtell extended the notion of modelling people to growing entire artificial societies through agent simulation called Sugarscape model. Sugarscape agents emerged with a variety of characteristics and behaviours, highly suggestive of a realistic, although rudimentary and abstract, society. Emergent processes were observed including death, disease, trade, wealth, sex and reproduction, culture, conflict and war, and externalities such as pollution (Epstein & Axtell, 1996; Macal & North, 2006).

2.2. Application of agent-based model and Monte Carlo technique

Laver and Sergenti (2012), on their own, started with the twin premises that understanding multiparty competition is a core concern for everyone interested in representative democracy, and that we must understand multiparty competition 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 model decision-making by party leaders, in what was clearly an analytically intractable setting, in terms of the informal rules of thumb that might be used by real human beings, rather than the 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 with social turnout built upon developments in different fields within social science. They described and analysed an agent-based model (ABM) of repeated elections in which voters and parties behaved simultaneously. They placed voters in a social context and let them interact with one another when choosing 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 on feedback from the electorate. Their model yielded significant turnout, divergent platforms, and numerous results consistent with the rational calculus of voting model and the empirical literature on social turnout.

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 et al. (2014) hypothesised that the frequency of reported round percentages should be increased if election results are manipulated or forged. They analysed raw data from seven federal elections held in the Russian Federation during the period from 2000 to 2012. They used Monte Carlo simulations to confirm high statistical significance of man-made fraud in all elections since 2004. They discovered that the number of polling stations reporting turnout and/or leader’s result expressed by an integer percentage, as opposed to a fractional value, was much higher than expected by pure chance.

In a similar study, Ananyev and Poyker (2018) examined the role electoral fraud plays in non-democracies. They focused on data obtained from the 2011 parliamentary elections in Russia and a regionally representative public opinion survey. Using indicators based on empirical studies and Monte Carlo simulation, they discovered that more manipulation occurred in the areas where the regime was more popular. In a separate study carried out in developing democracies, Ferrari et al. (2018) used Monte Carlo methods, nonparametric Bayes and path sampling methods to investigate fraud in the 2013 presidential elections in Kenya and 2014 presidential elections in Brazil. The electoral data for the Brazil election was aggregated to town level while that of Kenya was at ward levels. Their model shows signs of fraud in the Kenya election with no sign of fraud in the Brazil election.

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 fraud using recorded vote counts. Also, Rivest and Shen (2012) carried out a research on post-election auditing based on Bayes audit. The Bayes audit has the structure of two nested loops: an outer, “real-world” loop that is auditing the ballots one by one, and an inner “simulation” loop that is using Monte Carlo simulation to estimate the probability that the Bayesian model would generate an election upset (compared to the originally reported election outcome). Their study gave an experimental evidence of the effectiveness and efficiency of the Monte Carlo technique.

3. Research method
Monte Carlo simulation, a technique in agent-based modelling, was adopted for the research design. The research locations include the 119, 973 polling units across the country. According to Alli (2015), 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 Federal Capital Territory, Abuja has 562 polling units. The population of this study, collated across the 36 states and the FCT, Abuja, includes the 73,528,040 and 67,422,005 registered voters for the 2011 and 2015 elections, respectively. The 2011 and 2015 presidential elections were purposively selected and vote counts of leading political parties in the elections, All Progressives Congress (APC) and People’s Democratic Party (PDP) in 2015 election as well as People’s Democratic Party (PDP) and Congress for Progressive Change (CPC) in 2011 election, were simulated using R programming. Further analyses were carried out using frequency and percentage distribution and Spearman Rank Correlation tests.
PERT distribution approach was adopted for the Monte Carlo simulation owing to the fact that the Independent National Electoral Commission (INEC), the electoral body in Nigeria, did not provide detailed breakdown of election results at wards or polling units. The vote counts collated from all the polling units in each of the 36 states and the FCT, Abuja were aggregated to state level. The 2011 and 2015 presidential election results obtained online, as released by INEC, are therefore summaries of the election results collated at the state levels. These results are in the public domain (see Appendix A and B).

4. Data analysis
This section focuses on the simulation of the vote counts in the 2011 and 2015 presidential election results using Monte Carlo approach. 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 states (including the FCT) by simulating with a pert distribution. The pert distribution uses the minimum, expected and maximum values of the observed election data. The analysis was carried out on the two leading political parties of the presidential election results.

4.1. Monte Carlo Simulation of the 2015 presidential election’s result
This sub-section presents the Monte Carlo Simulation of the vote counts for APC and PDP (the two leading/major political parties) in each of the geopolitical zones and all the states in the 2015 presidential election. Figure 1 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 1 also reveals that the actual votes for APC and PDP in Lagos State are far beyond the simulated votes for both parties.

Key: APC = All Progressives Congress; PDP = People’s Democratic Party

Figure 2 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 2 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.

Key: APC = All Progressives Congress; PDP = People’s Democratic Party
Figure 2. Monte Carlo surrogate and actual votes for South East zone in 2015.

Figure 3. Monte Carlo surrogate and actual votes for South South zone in 2015.

Figure 3 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 3 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 PDP were alarming and highly exceed the simulated votes in Akwa Ibom State, Delta State and Rivers State.

Key: APC = All Progressives Congress; PDP = People’s Democratic Party

Figure 4 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. The simulated votes were higher for APC in all the states in the North West geo-political zone. The figure also shows an extremely higher actual vote counts for APC in Kano State when compared to the simulated votes.

Key: APC = All Progressives Congress; PDP = People’s Democratic Party

Figure 5 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 5 shows that the actual vote counts and the simulated votes are close for PDP in Borno State, Gombe State and
Figure 4. Monte Carlo surrogate and actual votes for North West zone in 2015.

Figure 5. Monte Carlo surrogate and actual votes for North East zone in 2015.

Figure 6. Monte Carlo surrogate and actual votes for North Central zone in 2015.
Yobe State. On the contrary, the actual vote counts were extremely high for APC in Bauchi State when compared with the simulated votes.

**Key:** APC = All Progressives Congress; PDP = People's Democratic Party

**Figure 6** 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.

**Key:** APC = All Progressives Congress; PDP = People's Democratic Party

**Figure 7** shows the Monte Carlo simulated votes (surrogate) and the actual votes for APC and PDP in all the states for the presidential election in 2015. For clarity, the distribution of the actual vote counts and the simulated votes were presented from **Figure 1** to **Figure 6**. The mean (416,889.8) and the standard deviation (413,798.6) of the vote counts for APC in all the states differ from the mean (139,155.8) and the standard deviation (87,957.74) of the simulated votes. Likewise, the mean (347,382.8) and the standard deviation (315,993.0) of the vote counts for PDP in all the states differ from the mean (225,596.4) and the standard deviation (152,096.4) of the simulated votes.

**Key:** APC = All Progressives Congress; PDP = People's Democratic Party

**4.2. Monte Carlo simulation of the 2011 presidential election’s result**

This sub-section presents the Monte Carlo Simulation of the vote counts for PDP and CPC (the two major political parties) in each of the geopolitical zones and all the states in the 2011 presidential election. **Figure 8** presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in the South West geo-political zone. **Figure 8** shows that the actual vote counts and the surrogate votes are higher for PDP in all the states in the South West zone. **Figure 8** 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, Ogun, Ondo and Osun states.

**Key:** CPC = Congress for Progressive Change; PDP = People’s Democratic Party

![Figure 7. Monte Carlo surrogate and actual votes for all states in 2015.](https://doi.org/10.1080/23311886.2021.1914397)
Figure 8. Monte Carlo surrogate and actual votes for South West zone in 2011.

Figure 9. Monte Carlo surrogate and actual votes for South East zone in 2011.

Figure 9 shows the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in the South East geo-political zone. Figure 9 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 9 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.

Key: CPC = Congress for Progressive Change; PDP = People's Democratic Party

Figure 10 presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in the South South geo-political zone. It also reveals that the actual vote counts and the surrogate votes are very low for CPC in all the states in the South East zone. In addition, the figure 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.

Key: CPC = Congress for Progressive Change; PDP = People's Democratic Party

Figure 11 shows the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in the North West geo-political zone. Unlike in the South East and South South zones, Figure 11
Figure 10. Monte Carlo surrogate and actual votes for South South zone in 2011.

![Monte Carlo surrogate and actual votes for South South zone in 2011](image)

Figure 11. Monte Carlo surrogate and actual votes for North West zone in 2011.

![Monte Carlo surrogate and actual votes for North West zone in 2011](image)

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.

Key: CPC = Congress for Progressive Change; PDP = People’s Democratic Party

**Figure 12** presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and CPC in the North East geo-political zone. **Figure 12** reveals that the actual vote counts for CPC were extremely high in Bauchi and Borno states 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.

Key: CPC = Congress for Progressive Change; PDP = People’s Democratic Party

**Figure 13** presents the Monte Carlo simulated votes (surrogate) and the actual votes for PDP and 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
Figure 12. Monte Carlo surrogate and actual votes for North East zone in 2011.

Figure 13. Monte Carlo surrogate and actual votes for North Central zone in 2011.

Figure 14. Monte Carlo surrogate and actual votes for all states in 2011.
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.

Key: CPC = Congress for Progressive Change; PDP = People’s Democratic Party

Figure 14 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 a clearer view, the distribution of the actual vote counts and the simulated votes have been presented in Figures 8-13. The mean (607,978.0) and the standard deviation (429,220.5) of the vote counts for PDP in all the states differ from the mean (458,099.6) and the standard deviation (449,434.8) of the simulated votes. Likewise, the mean (330,131.2) and the standard deviation (438,261.4) of the vote counts for CPC in all the states differ from the mean (40,418.7) and the standard deviation (40,979.6) of the simulated votes.

4.3. Test of hypotheses
In this section, results of the tests of hypotheses are presented.

H₀₁: There is no significant correlation between the digital distribution of vote counts of the 2011 presidential election results in Nigeria and the digital distribution of their respective simulated votes from the Monte Carlo simulation.

Table 1 shows the correlation between the actual vote counts and the respective simulated votes in the 2011 presidential election. We infer from Table 1 that there is a significant correlation between the simulated votes and the actual votes of PDP (p-value < 0.05) and CPC (p-value < 0.05) in the 2011 presidential elections. Therefore, the null hypothesis is rejected.

Key: PDP = People’s Democratic Party; CPC = Congress for Progressive Change

H₀₂: There is no significant correlation between the digital distribution of vote counts of the 2015 presidential election results in Nigeria and the digital distribution of their respective simulated votes from the Monte Carlo simulation

The correlation between the actual vote counts and the respective simulated votes in the 2015 presidential election is presented in Table 2. We infer from Table 2 that there is a significant correlation between the simulated votes and the actual votes of APC (p-value < 0.05) and PDP (p-value < 0.05) in the 2015 presidential elections. Therefore, the null hypothesis is rejected.

Key: APC = All Progressives Congress; PDP = People’s Democratic Party

The results of the hypotheses imply that there is a correlation between the simulated votes and their respective actual vote counts. This also implies that the simulated votes represent (conform to) the actual vote counts for each of the political parties.

Table 1. Spearman correlation coefficients for votes in 2011 presidential election

|                | Correlation Coefficient | p-value | N  |
|----------------|-------------------------|---------|----|
| 2011 PDP/Surrogate’s votes                  | 0.881     | 0.000  | 37 |
| 2011 CPC/Surrogate’s votes                  | 0.953     | 0.000  | 37 |
5. Discussion
The tests of hypotheses show 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 fraud should have vote counts that are approximately random (Johnson, 2011; Kobak et al., 2014; Oleg, 2011). 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.

Pert distribution was used to simulate the actual election data because there are fewer details in the state collated results released by the Independent National Electoral Commission. Similar to Cantu and Saiegh (2011), who were able to separate clean elections from the fraudulent elections by creating a set of simulated elections using Monte Carlo method to simulate vote counts and Bayes classifier as learning algorithm, 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 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.

The findings corroborate reports from the European Union Election Observation Mission. 2015. Second preliminary statement on (2015) on the 2015 presidential elections and the submission of Yusuf and Zaheruddin (2015) on the 2011 presidential elections. Both studies revealed that the 2011 and 2015 presidential elections were marred with electoral malpractices. However, the Monte Carlo analysis alone might not provide sufficient evidence to back up the suggestion that the deviation in pattern represents a fraudulent process in the vote counts of the 2011 and 2015 presidential election results. Hence, it is important to compare the results of the Monte Carlo analysis with other election reports, such as election audit reports and election observer’s reports.

This research work has contributed to literature in showing that Monte Carlo simulation, using PERT distribution approach, can be used to analyze election data with a view to detect irregularities in the data. Based on the findings of this study and the reviewed reports in literature, the 2011 and 2015 presidential election results were marred with irregular data.

6. Conclusions and recommendation
This paper presents the application of Monte Carlo simulation to analyse vote counts from the 2011 and 2015 presidential elections in Nigeria. The results reveal that the digital distribution of the vote counts for PDP and CPC in the 2011 presidential elections as well as vote counts for APC and PDP in the 2015 presidential elections do not conform to the respective distributional pattern of the simulated vote counts. 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 generated. The findings, therefore, suggest that the electoral process did not produce independent and randomly generated votes in the 2011 and 2015 presidential elections.

Table 2. Spearman correlation coefficients for votes in 2015 presidential election

| Year | Party Combination | Correlation Coefficient | p-value | N |
|------|-------------------|--------------------------|---------|---|
| 2015 | APC/Surrogate's votes | 0.890 | 0.000 | 37 |
|      | PDP/Surrogate's votes | 0.849 | 0.000 | 37 |
It is recommended that the Independent National Electoral Commission should intensify effort in making elections in Nigeria closer to the hitch-free elections obtained in more stable democracies. There should be a reform of the electoral act to produce a framework that addresses issues of vote buying and other election fraud, litigations, prosecution of electoral offenders within the shortest possible period, and structural reform of the electoral bodies. The electoral reform should also allow for use of forensic analysis of electoral data by plaintiff or defendant at the election petitions and tribunals. In addition, relevant authorities should ensure that the security forces deployed to maintain peace and order during elections do not interfere in the election process or participate in any form of election fraud.

Acknowledgements

We thank the team at the Laboratory for Interdisciplinary Statistical Analysis, Department of Statistics, University of Ibadan for advising and assisting with the simulation.

Funding

The authors received no direct funding for this research.

Author details

Sunday Tunmibi1

E-mail: sundaytunmibi@gmail.com

ORCID ID: http://orcid.org/0000-0001-8262-9459

Wole Olatokun1

ORCID ID: http://orcid.org/0000-0003-2780-2161

1 Department of Data and Information Science, Faculty of Multidisciplinary Studies, University of Ibadan, Ibadan, Nigeria.

Declaration of interest statement

We (Sunday Tunmibi and Wole Olatokun) declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Citation information

Cite this article as: Monte Carlo simulation of vote counts from Nigeria presidential elections, Sunday Tunmibi & Wole Olatokun, Cogent Social Sciences (2021), 7: 1914397.

References

Akin, O. 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/.

Ananyev, M., & Pokyer, M. 2018. Do dictators signal strength with electoral fraud? Retrieved February 18, 2021, from http://dx.doi.org/10.2139/ssrn.2712064

Animashaun, K. (2010). Regime character, electoral crisis and prospects of electoral reform in Nigeria. Journal of Nigeria Studies, 1(1), 1–33.

Birch, S. (2007). Electoral systems and election misconduct. Comparative Political Studies, 40(12), 1533–1556. https://doi.org/10.1177/0010414006292886

Breunig, C., & 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. https://doi.org/10.1016/j.electstud.2011.03.005

Cantu, F., & Saiegh, S. M. (2011). Fraudulent democracy? An analysis of Argentina’s infamous decade using supervised machine learning. Political Analysis, 19(4), 409–433. https://doi.org/10.1093/pan/mpr033

Epstein, J. M., & Axtell, R. (1996). Growing artificial societies: Social science from the bottom up. MIT Press.

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.

Ferrari, D., McAlistcr, K., & Mebane, W. R. (2018). Developments in positive empirical models of election frauds: Varying dimensions [Paper presentation]. At the 2018 annual meeting of the midwest political science association, Chicago, IL, April 5–8, 2018.

Fowlen, J. H., & Smirnov, O. (2005). Dynamic parties and social turnout: An agent-based model. American Journal of Sociology, 110(4), 1070–1094. https://doi.org/10.1086/426554

Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (2012). Agent-based models of geographical systems (pp. 85–105). Springer.

Ichino, N., & 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. https://doi.org/10.1017/S0022381611001368

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-mc20110428.pdf.

Kelley, J. (2012). Monitoring democracy: When international election observation works, and why it often fails. Princeton University Press.

Kio, B. (2013). Electoral corruption and democratic sustainability in Nigeria. IOSR Journal of Humanities and Social Science, 17(5), 42–48. http://www.iosrjournals.org/iosr-jhs/papers/Vol17-Issues/5/01754248.pdf

Kobak, D., Shpilkin, S., & Pshenichnikov, M. (2014). Integer percentages as electoral falsification fingerprints. Preprint arXiv, 1410.6059. https://doi.org/10.1214/16-AOS1904

Lover, M., & Sergenti, E. (2012). Party competition: An agent-based model. Princeton University Press.

Lehoucq, F. (2003). Electoral fraud: Cause, types, and consequences. Annual Review of Political Science, 6(1), 233–256. https://doi.org/10.1146/annurev.polisci.6.121901.085655

Levin, I., & 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.

Macal, C. M., & North, M. J. (2006). Tutorial on agent-based modelling and simulation part 2: How to model with agents. Argonne National Laboratory.
MacGregor, J. N., Edward, P. C., & Thomas, C. O. (2006). A comparison of heuristic and human performance on open versions of the traveling salesperson problem. Journal of Problem Solving, 1(1), 33-43. https://doi.org/10.7771/1932-6246.1005

Mebone, W. R. (2006). Election forensics: Vote counts and Benford’s Law. Political Methodology Society. University of California at Davis. Mebone, W. R. (2010). A layman’s guide to statistical election forensics. In Election guide digest: IFES. Nagaraju, V., Vlijnder, C., Bikshapathi, V., & 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.

Ogbeldi, M. M. (2010). A culture of failed elections: Revisiting democratic elections in Nigeria, 1959–2003. HAOL, 21, 43-56.

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

Onebamhochi, O. N. (2011). Curbing electoral violence in Nigeria: The imperative of political education. International Multidisciplinary Journal, Ethiopia, 5(5), 99–110. http://doi.org/10.4314/afrev.v5i5.9

Osinkachukwu, N. P., & Jawan, J. A. (2011). The electoral process and democratic consolidation in Nigeria. Journal of Politics and Law, 4(2), 128–138. https://doi.org/10.5539/jpl.v4n2p128

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

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

Rivest, R. L., & 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. Bellevue, Washington, USA.

Schelling, T. C. (1978). Micromotives and macrobehaviour. Norton.

Tunmibi, S. & Olatokun, W. (2021). Application of digits based test to analyse presidential election data in Nigeria. Commonwealth & Comparative Politics, 59(1), 1–24. https://doi.org/10.1080/14662043.2020.1834743

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

Yusuf, I., & 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.

APPENDIX A
The result of the 2011 presidential election in Nigeria is available at:

http://www.nigerianmuse.com/20110419040622zg/sections/general-articles/details-of-2011-presidential-election-results-in-nigeria/

APPENDIX B
The result of the 2015 presidential election in Nigeria is available at:

www.inecnigeria.org/wp-content/uploads/2015/04/summary-of-results.pdf
You are free to:

Share — copy and redistribute the material in any medium or format.
Adapt — remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
No additional restrictions
You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Cogent Social Sciences (ISSN: 2331-1886) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com