ANALYSIS OF DICHOTOMOUS OUTCOMES:

A Case Study of Constituency Characteristics' Data in the 2006 General Elections in Uganda

By

OCHAMA A. AHMED
BSc. (With Educ.) Hons., Maths and Physics, IUIU
2002/HD15/1905U

A Research Thesis Submitted to the Directorate of Graduate Studies for the Award of the

Degree of Master of Statistics of Makerere University

December, 2012

DECLARATION

I, OCHAMA A. AHMED, declare that the information in this book is my original work and has never been produced in this nature of form and content.

SIGNED:……………………………….   DATE:…………………………

AHMED A. OCHAMA

APPROVAL

APPROVED BY:

…………………………… DATE:…………………………

DR BRUNO OCAYA

(SUPERVISOR)

DEDICATION

I dedicate this book, most especially, to my loving mother and family to whom there is no substitute to schooling; wife Nezmah, and children – Hashim, Hamza and Hanna – whose invaluable right to attention was severally interrupted in the course of writing this research.

ACKNOWLEDGEMENTS

ABSTRACT

Using election data on the 2006 Presidential poll following re-introduction of multi-party system and various surveys' data compiled around 2006, this study explored characteristics that could have guaranteed election victory to the incumbent. The study considered scores for Constituency ratings – derived or otherwise - on income and poverty levels, participation, incumbency, literacy levels, party dominance or allegiance of the Constituency to a party, effect of regional location and the nature of the Constituency. The study population comprised 214 Parliamentary Constituencies covering the whole of Uganda.

A systematic approach utilising Hosmer-Lemeshow logistic regression technique for analysis of dichotomous outcomes was used. On analyzing the variables, all other quantitative predictors, except participation of the Constituency voters, average monthly per Capita incomes of voters and dominance by a political party, were eliminated. These were found to be statistically reliable for the restricted model in correctly predicting 70.56% of the data.

The policy implications for multiparty systems suggest that political and coalition actors could optimise their decisions or vote chances by articulating alternative policy matters in their campaign strategies. Unobservable income and other forms of inducements or coercion on which data are seldom collected or lacking may influence polls' outcomes. Future studies could focus on the influence of these factors on election results. Technically, it remains incumbent upon the Electoral Commission and government to be held accountable for the entire electoral environment during the poll's cycle.

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

BJP   Bharatiya Janata Party

DP   Democratic Party

EA   Electoral Area

EC   Electoral Commission

FDC   Forum for Democratic Change

MP   Member of Parliament

NRMO   National Resistance Movement Organisation

UBOS   Uganda Bureau of Statistics

UPC   Uganda Peoples Congress

CHAPTER ONE

INTRODUCTION

1.1   Background to the Study

Statisticians in research devise methods for estimating the relationships between response and explanatory variables that today constitute an integral part of data analysis.

Recent developments in statistical analyses reveal increased interest in the application of regression techniques involving dichotomous outcomes; that is, the requirement to reliably classify observations into one of two (binary) outcomes for analysis and prediction. These require application of logistic regression or probit analyses techniques. The techniques are appropriate for representing only two attributes of interest. Precisely, logistic regression, proposed in the late 60's and early 70's (Canbrera, 1994); and becoming routinely available in statistical packages in the early 80's, specifies the probabilities of particular outcomes (e.g. "pass" and "fail") for each subject or case involved. It produces a regression equation that accurately predicts the probability that an individual or entity will fall in one category or another (Tate, 1992). A dichotomous dependent variable is therefore assumed for a binary logistic regression where upon the value being predicted in the regression is actually a probability; ranging from 0 to 1.

Within the context of elections, the partisan choice is the most important decision made by the citizenry[1]. The election response could be visualized in terms of two (dichotomous) possibilities; namely; winning or not winning majority votes in a constituency. Economic

---

[1] Andrew L., (2005), Economic Voting and Electoral Behaviour: How do Individual, Local and National Factors Affect the Partisan Choice?, Economics and Politics Volume 17 No. 2 0954-1985.

growth, income distribution, social politics, and even the decision to go to court are often affected by whether one party or another is victorious.

The 2006 general election in Uganda provides a good scenario for understudying constituency characteristics and their impact on polls' outcomes. Firstly, the 2006 elections were characterized by the rebirth of multiparty democracy in Uganda. This followed the 2005 referendum to determine the choice of political system Uganda was to adopt under universal adult suffrage. Previously, multiparty elections conducted in Uganda were held in 1980. That is, the one of 2006 coming after 25 years. Previous general elections in Uganda were, hitherto, held in 1996 and 2001 under the "No Party" Movement type of governance in which the guiding principle for elections at all levels was the non-partisan "individual merit." Secondly, it was the very first time to conduct multiple level elections concurrently.

Events leading to the multiparty contest of 2006 saw the number of districts in Uganda grow from 29 in 1996 to 59 in 2001, then to 69 in 2006. Similarly the number of presidential candidates grew from 3 in 1996 to 4 in 2001, then 5 in 2006. These changes inherently altered the political spectrum for the 811 candidates vying for Parliamentary Constituency positions, and at the lower levels, of elective politics. The stakes were higher in 2006 considering that the outcome of the elections in 2001 had subjected the victory of the incumbent Presidential Candidate to a Supreme Court process and that of many other elections contested in court. Notably the Supreme Court proceedings[2] enlisted affidavits from, among others, Statistics professionals[3] to adduce evidence against what was deemed contestable election results that gave victory to Candidate Yoweri Kaguta Museveni. The Electoral Commission had declared Candidate Yoweri Kaguta Museveni the victor having obtained 59.26% being the highest

---

[2] Kiiza Besigye vs Yoweri Kaguta Museveni (2006); Presidential Election Petition No.1.
[3] Odwe Report (2006).

percentage of total votes cast[4] compared to Candidates Besigye Kizza (FDC): 37.39%, Ssebaana Kizito (DP): 1.58%, Abed Bwanika (Independent): 0.95% and Miria Kalule Obote (UPC): 0.82%[5].

The Constituencies were characterized by striking aspects of poverty depicted by a 2002-2005 poverty mapping study.[6] For instance, in some sub-counties, particularly in the remote districts of eastern and northern Uganda, over 90% of the population was found to live below the poverty line. In the western and central parts, less than 20% of the population was found to be poor. In the urban areas, poverty levels were much lower. Also, the wealthier districts in the western and central regions were more advantaged in terms of lower pupil-teacher ratios relative to other regions. Districts in the north and east of Uganda had much higher pupil-teacher ratios than the rest of the country. A special case of low pupil-classroom ratio was found in the districts of Kotido, Moroto and Nakapiripirit attributed to low education attitude and cultural barriers. According to the Uganda National Household Survey 2002/2003, the adult literacy rate in Uganda was estimated at 69%. The same survey indicated that 17% of the population aged 15 years and above had never had any formal education and 44% had not completed primary education. The aggregate expenditure of local governments in 2004/2005 was Uganda Shillings 896.4 billion (US$610m). The main items of local government expenditure were: (1). Education 42%, (2). Statutory bodies 3%, (3). General administration 15%, (4). Finance 7%, (5). Health 15%, (6). Other expenditure 12% accounting for between 11.4% - 12.4% of central government transfers to local governments. Conditional grants were

---

[4] A Presidential Candidate needs at least 51% of the total valid votes cast to win an election.
[5] Electoral Commission (2006), Report on the 2005/2006 General Elections, Kampala.
[6] Spatial Trends of Poverty and Inequality in Uganda (2002-2005)

given to local governments to finance programmes agreed upon between the Central and Local Governments[7].

Key questions, though, remain unanswered whether the results of the 2006 elections could have been different or whether the outcome could have been verified using these characteristics. For strategic reasons, one other question begs; what politicians or parties could have focused on in order to win elections without having to resort to assertions or cumbersome legal processes to claim victory.

The layout of the study is as follows: Chapter One is the Introduction, Chapter Two gives the Literature Review, Chapter Three details with the Methodology of the study, Chapter Four provides the Interpretation of Results and Discussion of Findings. Chapter Five is the last part which gives the Summary, Conclusion and makes Recommendations from the study.

---

[7] The unconditional grant is the minimum grant to finance decentralised services, allocated according to a formula provided for in the constitution.

1.2   Statement of the Problem

Common cases of assessing election victories, or defeats, have depended on ballot results and accounts by various election practitioners and observers. Some of these accounts are not assertions that have been subjected to rigorous empirical verifications. Ideally, elections conclude with clear-cut and unchallenged victories by the candidates and ballot measure outcomes that represent the will of the electorate. Where the elections have been contested; short of vote recounts, grounds for claims have relied partly on Court procedures, or entirely, on qualitative observations to draw conclusions about the election outcomes. This is because of inadequacy, or luck of, data concerning considerations for characterising the elections.

In this thesis, the researcher examines constituency characteristics and how they could have influenced the outcome of the 2006 Presidential election in Uganda. The primary goal of this research is to analyse the factors at play during the election. The analysis of key constituency characteristics frames future and broader understanding of people-centred politics for politicians. Through identification of data handling techniques undertaken, associations between variables and scenarios is presented as an undertaking in statistical methodology which is often unavailable on record.

In the study, the researcher applies the methodology for analysis of dichotomous outcomes in order to predict the polls' outcome using constituency characteristics' data. An attempt is made to ascertain which variables are central to addressing the classification of unknown election outcomes by use of characteristics for correct cases, or otherwise, of the data.

This researcher found no attempt made to model polls' outcomes based on constituency characteristics as likely determinants for winning or losing an election in Uganda.

1.3   Purpose and Objectives of the Study

General Objective:

The main purpose of the study was to apply the methodology for analysis of dichotomous outcomes to predict polls' outcome using constituency characteristics' data in the general elections of 2006 in Uganda.

Specific Objectives:

The specific objectives dealt with the analyses of the influence of Constituency characteristics on the Presidential election outcome. These included:

1.3.1   To analyse the effects of Constituency characteristic variables on polls' outcome. *viz*:

   (a) Poverty levels.

   (b) Participation of the Constituency voters.

   (c) Incumbency of the Candidate for the seat of President.

   (d) Income levels of voters.

   (e) Literacy levels of voters.

   (f) Party Dominance.

   (g) Location of the Constituencies by region.

   (h) Nature of the Constituency.

1.3.2   To determine which variables in 1.3.1 were central in the prediction of winning the election.

1.3.3   To classify Constituencies, in terms of representativeness, correctness or otherwise, of the data according to characteristics for predicting a poll's outcome.

1.4   Research Questions

(a) Can winning an election be correctly predicted from the knowledge of the voting population, allocation of government resources or grants at the local levels, civic consciousness of the citizens, income levels of citizens, proportion of local government positions won by a party or its dominance and proportion the of citizenry deemed active in local politics?

(b) If winning an election can be correctly predicted, which variables are central in the prediction? Does the inclusion of a particular variable increase or decrease the probability of the specific outcome?

(c) How good is the model at classifying cases for which the outcome is unknown? In other words, how many constituencies are classified correctly and how many are not?

1.5   Significance of the Study

The primary objective of this research is met by contributing insights to the pre-election conditions. The evaluation of key constituency characteristics lays emphasis on the election environment for future and broader understanding of people centred politics for politicians. Through identification of data handling techniques considered, statisticians will enhance considerations for improving accuracy in statistical analyses.

1.6   Limitations of the Study

Limitations of the study arise from problems of endogeneity of factors which may actually have been unobservable. Hence to make a statement about causal effects in order to empirically evaluate the theoretical work requires addressing the endogeneity problem. Although the impact such errors may have on the accuracy of the election outcomes, it is not systematically investigated in this dissertation.

1.7 Assumptions of the Study

The study made the following assumptions:

(a) Elections were conducted under ideal democratic conditions.

(b) Characteristics are homogeneous across all constituencies but only vary in extent.

(c) Each factor (or each continuous explanatory variable) has an independent effect on the response variable. That is; the explanatory variables did not have joint effects.

CHAPTER TWO

LITERATURE REVIEW

A broad range of literature on elections and parties emphasise that politicians are motivated by the attractions of political office (Schlesinger, 1991; Harmel & Janda, 1994; Muller & Strom, 2000). Some studies have examined the role of income inequality and its impact on income-based voting, the link between competition for political office and incumbent performance at the lower government levels (Besley & Burgess, 2002) while others have considered the influence of literacy levels of citizens, disbursement of government grants and civic consciousness of voters. Effort to understand the determinants of polls outcomes take observed empirical regularities as their starting point. This chapter will highlight the works of various researchers and the linkages between factors and election outcomes.

2.1   INCOME LEVELS

The link between income and political preferences and behaviour has long played an important role in theorizing about economic policy making, redistribution and the welfare state. Classical models of redistribution such as Romer (1975), Roberts (1977) and Meltzer and Richard (1981) postulate that a citizen's political preferences are based on the differences between her own income and the average income in society. Models of social insurance such as Moene and Wallerstein (2001, 2003) and Iverson and Soskice (2001) assume that one's income affects levels of economic risk and tastes for insurance and consequently political preferences over the extent and form of the welfare state. These models have generated quite a large empirical literature testing implications such as the positive relationship between income inequality and

redistribution e.g. Milanovic (2000) and Kenworthy and Pontusson (2005) and negative relationship between inequality and economic growth (e.g. Alesina and Rodrick, 1974, Perotti, 1996). But surprisingly, little attention has been paid to the validity of the underlying assumptions on the influence of income on voting behavior.

Two findings seem worthy of note. The first is the income-based voting across parties and within the electoral systems is predictably responsive to the positioning of the parties on the redistributive policy dimension (Benoit and Laver, 2006). As predicted by the formal model, multinomial logit coefficients tend to be strongly correlated with the positions on Benoit-Lavers' redistributive policy positions.

The second notable finding relates to the role of income inequality. As discussed above, income inequality may affect income-based voting in two ways. The first is by polarizing the electorate so that more support goes to extreme parties. The second is a contextual effect where income inequality is related to the weight that voters place on income in their voting decisions. Both the explanatory and model-based analyses indicate that the second effect swamps the first and that voters in egalitarian societies place more weight on income in their voting decisions. This relationship contradicts the standard political economy argument that inequality would increase the salience of redistributive issues (Kitschelt, 2003:6).

2.2   LITERACY LEVEL

The connection between literacy and political engagement is predicated on the assumption that as individuals become more exposed to information about their environment, especially the public institutions and government, they will be more prepared to intervene to make such bodies more responsive to their needs. There is also the expectation that as individuals are engaged in political decisions about the myriad aspects of their lives, an intimate connection

emerges between literacy and democracy. At the individual level, several attitudes and behaviors are commonly recognized as political in nature: engagement in voting, seeking information about candidates or issues, participating in discussion of political issues, and affiliating with a political party or social movement (Torney-Purta and Richardson, 2004).

The key finding of Wolfinger and Rosenstone (1980) that education is the most important predictor of voting forms the basis for recent innovative work by Feddersen and Pesendorfer (1996, 1999), arguing that informal differences among voters can help explain the observed variation in political participation. The key understanding here is that high levels of civic literacy are necessary to sustain widespread political participation[8].

2.3   INCUMBENCY FACTOR

The incumbency advantage literature originated in the study of congressional elections in the United States. Starting in the late 1960s, scholars noticed an unambiguous empirical trend of lower congressional turnover and higher incumbent re-election. Even after enrolling for party membership, region, election, year, etc., researchers found statistical evidence for an incumbency advantage, "the extra percentage of the vote that appears to result simply from being an incumbent" (Fiorina, 1989).

Empirically in the United States, incumbents do far better than non-incumbents, and as a result, a number of explanations have been offered for this advantage. These explanations range from models of asymmetric information (Banks and Sundaram, 1993; Austen-Smith and Banks, 1989; Besley and Case, 1995; Rogoff, 1990) to the beneficial aspects of office entitlements such as franking privileges, the ability to shape redistricting plans, seniority privileges, and the potential for increased name recognition (Alford and Hibing; 1989). Nothing in these models,

---

[8] Henry Milner (2007), *Civic Literacy in Comparative Context – Why Canadians Must be Concerned*: Enjeux Publics, Vol.2, No.2, July 2001.

however, guarantees an advantage, and many of them could just as easily generate a disadvantage with minor changes in the underlying assumptions.

The Indian form of government is very similar to both the US and British systems. Like the US, India has a federal system with thirty-two semi-autonomous states and union territories joined under a single central government. A simple comparison of the reelection rates of all candidates suggests that incumbents are at a thirty-seven percent advantage compared to non-incumbents. As explained previously, however, this sample includes both very strong incumbents that derive their power from sources other than their office and conversely, very poor candidates who lose by very large margins and have little chance of building a vote base. On average, incumbents are more likely to be female, have substantially more political experience, and are more likely to belong to one of the major parties. In addition to the difference in incumbency status, the difference in the reelection rates thus likely reflects other differences between the candidates[9].

While the American politics literature focuses primarily on the individual running for re-election, studies of Indian politics use the word "incumbent" to describe both parties in control of power at different levels of government and individual candidates. "Anti-incumbency factor" has become the favorite buzzword in contemporary Indian politics; both political commentators and politicians using it to explain the defeat of incumbent parties and incumbent legislators. It became especially popular in the 1990s when India entered a phase of multiparty competition. Prior to that, the Congress party had dominated Indian politics. Till 1970, this dominance was complete at both national and regional levels. Between 1970 and 1990, the

---

[9] Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections***Leigh L. Linden**† January 8, 2004 (https://webspace.utexas.edu/linden/www/Incumbency%20Disad.pdf; Monday 22/08/2011)

Congress grip started slipping as the party began facing greater opposition in state assembly elections, and in the Lok Sabha. The final transition to multiparty competition happened in the early 1990s as more regional parties entered the fray in both national and state politics. At the national level, this has led to an era of coalition governments that either fall before completing their term in office or are voted out of office in the next election. Re-election rates are similarly low at the state level. Out of the 71 state assembly elections that took place between 1989 and 1999, only 29 governments secured re-election. With the exception of few states like West Bengal and Bihar, incumbent governments are regularly defeated after one term in office.

While these outcomes provide evidence for an anti-incumbency trend at the macro-level, it is equally, if not more, important to understand the nature of electoral politics at the constituency level to get a complete picture. This is because it is not the case that any party that successfully gets re-elected has escaped the "curse" of anti-incumbency at the constituency level. The party may have lost a large proportion of the seats it held in the previous election but gained new seats previously held by other parties. Seat retention rates have been low historically, long before the onset of multiparty competition in 1990. Congress' retention rate decreased from 74% in 1991 to 36% in 1999. The BJP won an identical number of seats in both the 1998 and 1999 Lok Sabha elections and led the seat tally on both occasions, but retained only 63% of the seats. Looking at party performance at the constituency level in national elections since 1970, Kumar (2003) finds that national parties are more likely to retain a seat than non-national parties. Moreover, greater voter turnout and larger gaps between elections decrease the probability of the incumbent party securing re-election.

The final type of incumbency effect is that of candidate incumbency. Linden (2004) uses non-parametric regression to measure the effect of candidate incumbency in national elections since 1977. His results show that prior to 1991, incumbency had a positive effect on a candidate's

chances for re-election in national elections but since then, incumbents are 14% less likely to win than non-incumbents. He argues that this reversal of fortunes may be linked to the Congress Party no longer being the single dominant party. While Congress dominated national politics, incumbent experience and connections within the party were valuable to voters. However, multiparty competition in the 1990s made it harder for voters to predict which party will gain control of parliament, causing them to care less about experience in the previous regime. One potential problem with this explanation is that while the onset of multiparty competition and increased uncertainly about electoral outcomes could plausibly explain a decline in advantage due to incumbency; it is less clear why voters would start voting against incumbents. Under uncertainly about the national winner, an incumbent should have an even chance of winning or losing.

2.4   TURNOUT AND THE PARTISAN CHOICE

The relationship between turnout and partisan outcomes in five U.S. Presidential elections observed was found to be characterized by both regularities and change over time. Jell and Martinez (2005) assert that, as a rule, Democrats do benefit from higher turnout—and probably should have lost both the 1960 and the 1976 elections had turnout been much lower. Similarly, the Republicans benefit from low turnout, and conceivably could have lost the 2000 election legally as well as empirically with some slight and well placed increases in turnout. They observe that on the other hand, there is a limit to the explanatory or determinative power of turnout.

Their central concern was to determine whether or not increased levels of voter turnout in specific elections would have aided the Democratic party and whether decreased turnout would

have aided the Republican party in U.S. Presidential elections over the past 40 years. In the process of addressing this question, they assessed the following:

• *whether higher levels of turnout are associated with a greater tilt toward the Democrats into the electorate (consistent with the conventional wisdom);*

• *whether higher levels of turnout are associated with greater defection rates (consistent with DeNardo's hypothesis); and,*

• *whether the strength of the relationship between turnout and partisan outcomes has declined concomitant with the decline in the class-party cleavage in the United States.*

Their basic approach, which is similar to Lacy and Burden (1999) involved positing that U.S. ;citizens generally have three unordered choices in each election; namely, vote Democratic, vote Republican, or abstain. They first estimated vote choice (including abstention) as an unordered multinomial logit function of standard variables associated with both candidate preference and the likelihood of voting. From that estimation, they derived probabilities for each respondent's selection of each of the three choices (abstain, vote Democratic, or vote Republican). Higher turnout was simulated by starting with the actual voters in a given election and progressively adding actual abstainers who had the lowest estimated probability of abstaining. They simulated lower turnout by progressively subtracting actual voters who had the highest probability of abstaining from the pool of voters. The simulated results allowed them to make informed estimates about the most likely consequences in any given election as if turnout had been different.

Citrin, Shickler, and Sides (2003) also used a simulation methodology to address the same question with respect to U.S. Senate elections in 1994, 1996, and 1998. Using exit poll data, they calculated the additive effects of age, income, education, gender, ethnicity, union membership, and veteran status on voter choice among observed voters (exit poll respondents). They estimated the impact of increased turnout on electoral outcomes by assuming that parameters of candidate choice among voters would have been the same among nonvoters if the latter had actually voted. Despite this basic similarity, there were important differences in the methodologies. Citrin, Shickler, and Sides' comparisons between voters in exit poll surveys and to census distributions constrain them to a Columbia school (sociological) model of voter choice (2003, 79), while the estimation in this study allowed for the potential effects of a variety of attitudinal and other non-demographic variables on both the decision to vote and candidate preference. Second, Martinez and Jeff's methodology yielded estimated probabilities of abstention and candidate preference, enabling them to examine the likely impact of turnout at smaller discrete increments of turnout change. Third, since they could also estimate the probability of abstention among actual voters, their methodology allowed estimating the likely effects of lower turnout as well as higher turnout, an aspect of the relationship which had been neglected in literature. Thus, their simulation methodology offered a different approach to understand the entire turnout-partisan effects question than that proposed by Citrin and his coauthors. To the extent that the results of the two approaches converged, they became more confident that their conclusions were not dependent on the assumptions inherent in each. They examined five presidential elections (1960, 1964, 1976, 1984, and 2000) which represented the variation in conditions that might have affected the relationship between turnout and partisan outcomes. While their primary purpose was to understand partisan behavior in two-party Presidential races going forward in time, they did not focus purely on the last one or two

elections since it was easily possible in their study to generalize "across space" as advised in King, Keohane, and Verba (1994, 219)[10].

There were three important driving considerations for the case-selection methodology. First, the period covered the decline of class polarization in the U.S. party system noted by many authors[11]. Second, the cases included elections which appeared to be (and were) close, as well as a couple of landslides. The anticipated closeness of an election might very well change the strategic behavior of elites and potential voters, which could affect both the likelihood of voting and in some circumstances, voter preferences. While it is unlikely that slight variations in the underlying conditions would have altered the outcomes of landslide elections, they include two such cases both to serve as contrasts to close elections and to examine whether variations in turnout might have affected the winner's margin of victory. Third, changes in election laws over this time period generally eased registration requirements in an attempt to lower the costs of voting and increase turnout. Two of their selected cases precede the 1965 federal Voting Rights Act, which eliminated some significant state restrictions on voter registration.

In sum, their selected cases represented a diverse set of electoral conditions and outcomes. The 1960 election was a landmark contest between a photogenic young challenger and the ultimate political insider, eventually decided by a razor-thin margin. The Democratic partisan majority was at its post-war peak in 1964 and contributed to a landslide victory for President Johnson and congressional Democrats. Twelve years later, the subsequent Democratic majority had eroded considerably and Jimmy Carter eked out a close victory over President Ford. By 1984, a

---

[10] The unit of space here is time, and we include previous cases that address the same hypotheses and retain the same general characteristics.

[11] Abramson, Aldrich, and Rohde 2002, 113–15; Dalton 1996; Nieuwbeerta and de Graaf 1999.

surge in Republican partisanship precipitated President Reagan's landslide re-election. Thus, in order to gain some leverage in understanding the effects of underlying partisan distributions and short-term forces on the relationship between turnout levels and electoral outcomes, they selected a Democratic landslide year (1964), a Republican landslide year (1984), two close Democratic victories (1960 and 1976), and an election in which the Republican won a razor-close electoral college vote while losing an equally close popular vote (2000).

2.5   CIVIC PARTICIPATION

Researchers in both political science and psychology have identified personal beliefs as an important factor influencing behavior or state of civic consciousness. The idea of efficacy is also a prominent part of theory and research in education (Stenz & Lambert, 1977). Notable links have been made between efficacy beliefs and academic or social achievement (Eccles, Wigfield, & Schiefele, 1998).

The concept of political efficacy and its relationship to the stability of democracies or the participation of citizens is one belief that has received considerable attention from political researchers. Political efficacy is the "feeling that individual political action does have an impact on the political process…" (Campbell, Gurin, & Miller, 1954, p.g. 187). There are at least two political theories that support the importance of instilling a positive sense of political efficacy in citizens (Finkel, 1985). The first theory attributes the stability of democratic regimes to at least a modest sense of efficacy among a majority of citizens. Citizens who believe that the government is responsive to their needs are more likely to submit to government authority or to grant it legitimacy, thereby contributing to a stable democracy. A second theory, participatory democracy, emphasizes the outcome effects for individuals of a

strong sense of political efficacy. It argues that individuals with positive political efficacy are more likely to participate in the political process, thereby ensuring an engaged citizenry which contributes to more than a threshold level of legitimacy.

Initial measurement of political efficacy dates to a study by Campbell et al. (1954), where an index was developed for use in predicting voting. Since that time a number of studies have assessed the reliability and validity of the measurement, which has led to further refinement of the concept of political efficacy into internal and external political efficacy (Acock, Clarke, & Stewart, 1985; Asher, 1974; Balch, 1974; Craig et al., 1990; Hayes & Bean, 1993). Internal political efficacy has been defined as "beliefs about one's own competence to understand and participate effectively in politics," whereas external political efficacy is "beliefs about the responsiveness of governmental authorities and institutions to citizens' demands" (Craig et al. 1990). The existence of a balanced pluralism among civil society interests so that none can establish absolute dominance[12].

Political discussions may include encouragement to engage in other political or civic activities. Some research shows that people who are asked to participate in civic activities are more frequent participants than those who are never asked (Keeter, Zukin, Andolina, & Jenkins, 2002b).

Bandura (1997) has applied his concept of perceived self-efficacy to other domains such as politics. His definition of self-efficacy as it applies to politics is the "belief that one can produce effects through political action" (p. 483).

---

[12] A strong civil society is identified by a very ideal model in Mouzelis: (8) (Mouzelis, 1995: 225-226)

Almond and Verba (1963) describe how political socialization is an informal learning process in which individuals gain knowledge and attitudes about political figures, processes, and systems. Their work describes the importance of adult political socialization, which is contrasted with the relative weakness of childhood socialisation. The study links the political attitudes of citizens in the U.S., Britain, Germany, Italy, and Mexico and how those states function. The authors look at three main dimensions of political culture, those being knowledge of the political system, feelings towards the political system, and attitudes towards the self as a political actor. They describe some characteristics of a balanced civic culture including: orientation toward the political and government system; pride in aspects of one's nation; expect fair treatment from government authorities; the capacity and ability to talk freely and frequently about politics and to be emotionally involved in elections; a tolerance towards opposition parties; membership in a voluntary association; the valuing of active participation in local government activities, parties, and in civic associations; a self-confidence in one's competence to participate in politics and finally civic cooperation and trust.

2.6   PARTY DOMINANCE

Definitions of party dominance can usefully be distinguished by means of four criteria: the threshold for dominance; the inclusion or exclusion of opposition features; in presidential systems of government: the presence or absence of divided government; and the time-span taken into account[13]. The simplest definition of party dominance relies only on vote or seat share; it is not interested in the make-up of the opposition, does not take into account presidential systems of government and is limited to a single legislature (1976:44). The most complex definition includes all four dimensions.

---

[13] *European Journal of Political Research* **43**: 173–197, 2004

The most promising definitions of party dominance and party system typologies have been applied to a data set of 18 African countries with a total of 59 elections. These were all sub-Saharan African countries that have held three or more consecutive multi-party elections up until the end of 2002. It was demonstrated that at least half of these countries have or had a dominant party and party system by most definitions. The second part approaches the same issue from a different angle, looking at the number of parties and how to count them.

The conclusion is that Sartori's counting rules, party system typology and his definition of a dominant party are still indispensable to arrive at an accurate classification of party systems and their dynamics in general, and of dominant party systems in particular.

The advantages of Sartori's framework of analysis are fivefold. First, Sartori's counting rules are not strictly based on relative size, but on the number of relevant actors in electoral competition and government formation. Second, Sartori's analysis employs a conception of dominance absent in continuous measures of party number. Third, the distinction between dominant and dominant authoritarian party systems encourages an identification of the nature of dominance. Fourth, dominant party systems are embedded in a typology of party systems. In fact, there are two typologies: one for structured party systems and one specifically designed for Africa's fluid polities and unstructured party systems. Finally, and decisively, Sartori's counting rules, his definition of a dominant (authoritarian) party system, and typology of party systems provide a unified and coherent framework of analysis that is sensitive to context and time. By contrast, the quantitative counting rules dominating the literature have no intrinsic relationship to party systems, making any attempt to go from one to the other extremely hazardous.

Though one-party dominance is a central feature of contemporary political systems in sub-Saharan Africa little effort has been spent on its systematic explanation as a dependent variable. Three potential main causes can be identified: Dominant parties benefit from the rules of the game such as electoral systems and systems of government, they command a huge support base for social, economic, political and historical reasons or they just play foul and secure dominance by undemocratic means. Testing respective hypotheses for 38 dominant and non dominant party systems in sub-Saharan Africa reveals that institutional variables offer almost no general explanatory value, hardly better do social, economic and other, "rational" determinants of voting behaviour. An authoritarian regime environment and a coercive historical origin of dominant parties explain at least two thirds of all cases.

A combination of variables offers the best explanation with only eight deviating cases.

In political systems that explicitly prohibit parties from contesting seats in local elections, national and regional parties often unofficially back candidates in local elections, even if this support is not denoted on the actual ballot. This is certainly the case in India, where party officials at the state level proffer statistics about the number of local governments controlled by the party and village political leaders will admit to being aligned with a certain party, even though local elections operate on a non-partisan basis by law (Packel 2007). Ghana also witnesses this practice, where parties unofficially sponsor individual candidates (Ayee, 2004). Thus, permitting parties to compete in local elections allows the inescapable link between local government and higher levels of government to exist openly, allowing for greater transparency. Where a strong national party system exists, parties have significant incentives to mobilize at the local level in order to improve their electoral chances at higher levels (Blair 2000).

2.7   THE INFLUENCE DUE TO POVERTY

The Ugandan economy has registered considerable success in recent years in terms of growth rates in comparison to other countries in the East African sub-region. The impressive economic performance has however not been matched with positive social indicators. The high level of poverty indices and massive under employment indicate that a very large part of the population is yet to benefit from the country's economic recovery. The under-employment in Uganda, a poverty problem in reality, is even more serious than it appears to be. The number of under-employed or unemployed people is estimated at about 3.8 million[14]. The growth in the labour force is at least 300,000 per year, out of which only 100,000 people are absorbed, mostly in the agricultural sector. As regards the distribution of household income, 49% of the total households earn between US$0-50 per month while 32% earn between US$50-100 per month. These ratios are reflective of mainly the rural situation, since only 19% of the households in urban areas earn between US$ 0-50. This indicates high income inequality in Uganda and the critical issue of under-employment, particularly in rural areas.

To understand Constituency characteristics related to poverty incidences, this study examines survey work conducted in Kenya.

According to the report – "*Decentralization, Accountability and the MPs Elections: The Case of the Constituency Development Fund in Kenya,*[15]" over the last ten years the Kenyan Government intensified the use of decentralized programs in its strategy to tackle poverty and reverse regional disparities. A prime example of this policy is the Constituency Development

---

[14] Uganda National Integrated Survey 1992-93;The Household Budget Survey 1989-90.
[15] R. G. Romero (2009) University of Oxford

Fund (CDF) launched in 2003. The CDF allocates resources to all of the 210 constituencies, taking into account constituency poverty levels. The CDF is designed to consider local needs and preferences, by stipulating that the MP of each Constituency should decide along with members of the local community how to use the funds to tackle poverty. In practice however, there have been concerns that the participation of residents in the decision making has been limited, and that the role of MPs as legislators, implementers and auditors of the fund reduces incentives for transparency and good management of the fund. Despite the weak institutional framework supporting the CDF, MPs are still accountable to residents in general elections, when residents have the opportunity to assess the performance of their MPs and reward or punish MPs with their vote. Suffice to note that in Kenya's last general election of December 2007, only 40% of the incumbent MPs were re-elected despite the fact that the great majority (80%) of them were contending for re-election. Unlike the presidential election, international electoral observers and Kenyan political parties did not dispute the MPs election outcome, so the election results for MPs.

The outcome of the MP elections raises serious questions about the factors that influenced voting behavior, and whether the way MPs managed the Constituency Development Fund affected their re-election chances in a country where voting has traditionally been motivated by ethnic ties rather than the performance of politicians. Some insights into these questions can be gained from the two nationally representative voter surveys that the Centre for the Study of African Economies (CSAE) conducted just before and after the 2007 elections, especially when combined with publicly available information of how MPs used the CDF over the period 2003-2007.

A potential reason for the difference in the average number of CDF projects run is the allocations that constituencies received. The MPs that lost in the 2007 elections got on average the highest CDF allocations, USD 758,294, as their constituencies had on average a larger number of poor people than the constituencies of the MPs that won or that did not contest in the 2007 elections. In contrast the Constituencies of the MPs who got re-elected received on average less CDF allocations (USD 739,336) than those MPs that lost in the elections and than those who did not contest in elections. The MPs who did not contest in the elections run, on average, fewer projects than those MPs who were contesting in the elections despite the fact that they received larger allocations than those who won the elections and a similar amount to those who lost the elections.

At the time of the pre-election survey, most respondents, 96% according to the survey, planned to participate in the elections. The survey showed that the leading candidates - Kibaki and Odinga - were in a virtual statistical tie (with 40.2% of intended vote for Kibaki and 46.7 for Odinga). It is true therefore that any one of the Presidential candidates could have won the election depending on the voter turnout. This pre-election survey was also reported to have been corroborated by a major exit poll conducted by the University of California, San Diego.

2.8   REGIONAL EFFECTS

Voting behaviour is determined by one's location[16]. One's community or region might have a strong tradition of support for a party and he becomes socialized into that tradition through family and peers. One reason is that parties and their union affiliates might have first established themselves among sugar, rice or banana workers and so their strength is based on

---

[16] http://gtuwi.tripod.com/voting3.html; 29th July 2011 16:55:01 GMT

the social and demographic profile of persons in those areas. The People's Progressive Party in Guyana has its strength among the rural Guyanese in the sugar and rice growing areas. This, in turn, overlaps with race because historically East Indians have worked longest in these areas. The People's National Congress gets its support from the more urban, Afro-Guyanese.

Location overlaps with other social characteristics. The rural population tends to be less skilled and less educated in contrast to the urban voters. So, there are differences between PPP and PNC voters based on location, race, and levels of education.

In Trinidad too, the tradition has been for the People's National Movement (founded by Eric Williams) to have more support from non-agricultural or urban regions. The United National Congress (formerly the United Labour Front) of Basdeo Panday gets more support from agricultural workers which his party was very instrumental in organising in the early days. Again, there is an overlap with other characteristics. Agricultural workers are mainly rural East Indians of lower education compared to supporters of the PNM.

The association between race and community voting is shown by a study of voting behaviour in the 1976 general elections in Trinidad and Tobago. The dominance of the PNM over the years could be explained by the fact that of the country's 36 constituencies, 19 had Afro-Trinidadians making up 50% or more of the voters, 13 had Indo-Trinidadians in the majority and four had mixed-race populations. In 1992, for instance the United Labour Congress (UNC) won all the rural East Indian constituencies.

Where people vote along racial lines as they tend to do in Trinidad, the party with a racial majority in more than half of the constituencies has a clear advantage. Because these constituencies are geographically concentrated in rural and urban regions there is also a

geographical split in the vote and this reinforces the sense of two societies coexisting tensely under one government. This gives rise to calls or sentiments favouring secession into two nations during periods when racial tensions are high.

In Jamaica, community and regional patterns of voting also exist. It is not based on race but on community loyalty to a party. For many years, the Jamaica Labour Party had been stronger in rural areas, especially in the traditional sugar and banana parishes (St. Thomas, Clarendon, St. Catherine) where BITU organisation had provided support among the agro-proletariat. The PNP had been stronger in Kingston and St. Andrew, capital cities, main towns and generally, the more urban areas. The lesser significance of sugar and banana workers now meant that there is no safe JLP parish. In the last two elections, the PNP won the vote in all parishes.

Except for 1980, the PNP has always won the majority of votes and seats in Kingston and St. Andrew. It was this that drove Mr. Seaga to carve out a garrison constituency in West Kingston to ensure continued re-election. This started the garrison phenomenon. But with its natural urban working class support, housing and social policies it was easy for the PNP to respond and outdo the JLP with its garrison constituencies.

Garrison constituencies and garrison communities within constituencies have become the prototype of community-based voting. There are between 11 and 13 garrison constituencies. Party loyalty is strictly enforced by area dons. But constituents born into these communities are highly socialised into the pattern of party loyalty. The natural tendency to support a party along with the zeal of these supporters often lead to over-voting and other abuses to the electoral system. A violent form of tribalism between garrison communities is also common.

2.9   RURAL, URBAN or PERI-URBAN CONSTITUENCIES

Although the vast majority of studies analyzing voting behavior and party competition in Africa have concentrated on ethnic or regional divisions, the rural-urban dimension has not been completely ignored. In a recent study of economic voting in Zambia's 1996 elections, Posner and Simon (2002) include rural-urban location as a control variable, and find that voters in urban areas were actually more likely to support the incumbent than were rural voters. Although contrary to the trend highlighted above, this finding is explicable given the Zambian context, because the ruling MMD party was born out of the urban protest movements of the late 1980s, and because its leader President Chiluba was a former trade unionist with a strong following amongst the urban industrial sector. This trend has since reversed in Zambia, with the proportion of urban residents supporting the incumbent in 2005 (still the same MMD, but at this time led by the late President Levy Mwanawasa), 13% lower than the equivalent proportion of rural respondents.

Kimenyi and Romero (2008) note that urban residents in Kenya were less likely to express incumbent support prior to the 2007 elections, although this relationship appears not to be significant, and they do not discuss it in any detail. In a study of economic voting in Ghana using survey data from 1999, Youde (2005) demonstrates a significant negative relationship between urban location and incumbent support. Again however, the rural-urban variable is simply included as a control in this study, and no attempt is made to explain the relationship. This limited literature therefore provides mixed evidence for the effects of urban-rural location on voting behavior, and offers little by way of explanation.

Somewhat more usefully, the hostility of Ghanaian urbanites towards the incumbent party that Youde highlights what had already been noted with regards to the 1992 and 1996 elections by Bawumia (1998) and Nugent (1999), both of whom explain it largely as a reaction to the effects of the economic Structural Adjustment Program (SAP) that had been pursued by President Jerry Rawlings since 1983. While the SAP certainly created winners and losers in both urban and rural areas, both Bawumia and Nugent note that the majority of the negative impact was borne by urban residents, particularly the job losses resulting from public sector retrenchment, but also the more effective system of urban taxation that resulted from the reforms. In contrast, the most visible benefits resulting from the SAP were heavily skewed towards rural areas, where increased investment in basic public services such as pipe-borne water, electricity, and roads had a much more noticeable impact. Furthermore, the removal of price distortions meant that agricultural producers in the countryside, especially cocoa farmers, received vastly improved prices for their crops. In addition, Nugent (1999) argues that the opposition parties were unable to compete effectively in rural areas, due to their limited ability to communicate with rural voters.

Taken together, these arguments begin to offer a reasonable framework for understanding why African urbanites might be less disposed than rural residents to support the incumbent government. This framework is added to by Conroy-Krutz (forthcoming), who claims that in addition to the negative urban impact of SAPs and the relative efficiency of urban as opposed to rural campaigning by opposition parties, incumbent parties across Africa engage in widespread voter bribery in rural areas. Incumbency gives governing parties a significant comparative advantage in the distribution of minor consumption goods during election campaigns. The electoral impact of such distribution, which Conroy-Krutz terms "petty

patronage", is assumed to be greater in rural areas, where voters can be bought more cheaply and more reliably. Therefore higher levels of voter bribery in rural areas make rural voters more likely than urbanites to support the incumbent. In addition, it may also be the case that the socio-demographic profile of rural residents makes it easier for incumbents to mobilize them. Rural voters tend to be older, less educated, and more female. These factors, along with the possibility that people in rural areas are more likely to vote on the basis of loyalty to identity groups and dominant governing parties, or to be controlled by influential traditional rulers, may reduce the demands of rural residents, and render them less autonomous in their voting decisions than urbanites. Moreover, these factors should be expected to further compound the effect of voter bribery noted by Conroy-Krutz.

Van de Walle (2003: 313) argues that party competition in Africa does not take place primarily along a rural-urban cleavage. Rather, the overriding tendency is for party support to be based on ethnic or regional constituencies, with groups linked to parties through clientele ties. This state of the world is consistent with the argument that a predominantly rural population induces incumbents to implement policies that favor rural interests. It is important to be clear that the argument being made here concerns interests, not identities. It is not the intention of this paper to claim that voters identify as either rural or urban, nor is it to deny that they could. Rather, the argument is that rural voters are likely to share a variety of common interests, for example in primary education, rural feeder roads, rural electrification, or agricultural prices, which may differentiate them from urban voters. As a result, politicians can court rural voters not by appealing to a voter's rural identity, but by implementing (or promising) policies that serve these particular rural interests.

CHAPTER THREE

METHODOLOGY

The main purpose of this research was to apply the methodology for analysis of dichotomous outcomes in order to predict polls' outcomes. In doing so, the research utilised data on constituency characteristics during the general elections of 2006 in Uganda to address the following research questions:

(a) Can winning an election be correctly predicted from the knowledge of the voting population, allocation of government resources or grants at the local levels, civic consciousness of the citizens, income distribution of citizens, proportion of local government positions won by a party or its dominance and proportion of the citizenry deemed active in local politics?

(b) If winning an election can be correctly predicted, which variables are central in the prediction? Does the inclusion of a particular variable increase or decrease the probability of the specific outcome?

(c) How good is the model at classifying cases for which the outcome is unknown? In other words, how many constituencies are classified correctly and how many are not?

This chapter, therefore, provides descriptions of the study area, design, population, sample size, sampling techniques, data collection tools, data presentation and analysis, ethical considerations, pre-test, limitations and dissemination of results.

3.1 The Framework for the Logistic Regression Method

The Bernoulli distribution having parameter π_i for the dichotomous dependent variable, Y_i, $\Pr\{Y_i = y_i\} = \pi^{y_i}(1 - \pi_i)^{1-y_i}$, is the assumed distribution for the conditional mean of the dichotomous outcome;

$$\text{Where } y_i = \begin{cases} 1 \ for \ a \ win \\ 0, otherwise \end{cases}$$

The expected value and variance for Y_i are:

$$E(Y_i) = \mu_i = \pi_i, and$$

$$Var(Y_i) = \sigma^2(1 - \pi_i).$$

This assumption implies that the same probability is maintained across the range of predictor values.

From a practical point of view, it is important to note that if the predictors are discrete factors and the outcomes are independent, the Bernoulli distribution is used for the individual zero-one, $(0,1)$[17], data.

The model for logistic regression analysis, LRA, assumes that the outcome variable, Y, is categorical (e.g. dichotomous) although LRA does not model this variable directly. Rather, the LRA is based on probabilities associated with the values of Y. For simplicity, and because it is the case most commonly encountered in practice, we assume that Y is dichotomous, taking on values of 1 (i.e. the positive outcome or success) and 0 (i.e. the negative outcome or failure). In theory, the hypothetical, population proportion of cases for which $Y = 1$ is defined as $p = P(Y = 1)$. Then, the theoretical proportion of cases for which $Y = 0$ is $1 - p = P(Y = 0)$.

---

[17] The binomial distribution, $Y_i \sim B(n_i, \pi_i)$, for grouped data consisting of counts of successes in each group yields exactly the same likelihood function and therefore the same estimates and standard errors.

This restriction yields a number of advantages. For one, coding the variable in this fashion implies that the expected value of y is simply the probability that $y = 1$:

$$E(y_i|x_i, \beta)| = 1. \Pr(y_i = 1|x_i, \beta) + 0. \Pr(y_i = 0|x_i, \beta)$$

$$= \Pr(y_i = 1|x_i, \beta).$$

Thus, a second interpretation of the binary specification is obtained as a conditional mean specification. It follows that the binary model can be written as a regression model:

$$y_i = \left(1 - F(-x_i'\beta)\right) + \varepsilon_i,$$

Where ε_i is a residual representing the deviation of the binary y_i from its conditional mean.

Then

$$E(\varepsilon_i|x_i, \beta) = 0$$

$$Var(\varepsilon_i|x_i, \beta) = F(-x_i'\beta)(1 - F(x_i'\beta))$$

The conditional mean is, thence, useful in the interpretation and discussion of binary model residuals.

In the absence of other information, p is estimated by the sample proportion of cases for which $Y = 1$. However, in the regression context, it is assumed that there is a set of predictor variables, $X_1, X_2, \ldots, X_p$, that are related to Y and, therefore, provide additional information for predicting Y.

For theoretical, mathematical reasons, LRA is based on a linear model for natural logarithm of the odds i.e. the log-odds in favour of $Y = 1$. Note that in the LRA model, p is a conditional probability of the form $P(Y = 1|X_1, X_2, \ldots, X_p)$. That is, it is assumed that "success " is more

or less likely depending on combinations of values of the predictor variables. The log-odds, as defined above, also known as the *logit* transformation of p, and the accompanying analytical approach are collectively known as *logit analysis*.

The simple logistic model has the form;

$$\log(Y) = natural \log(odds) = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta X \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . (1)$$

The equation to predict the probability of occurrence of the outcome of interest is as follows:

$$\pi = probability(Y = outcome\ of\ interest, X|x, a\ specific\ value\ of X)$$

$$= \left(\frac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . (2)$$

Where π is the probability of outcome of interest or "event," α (the $y-$intercept and β the value of the coefficient which determines the direction of the relationship between X and the logit of Y) is the regression coefficient, $e = 2.71828$ is the base of the system of natural logarithms. X can be categorical or continuous but Y is always categorical.

Extending the logic of the simple logistic regression to multiple predictors one can construct a complex logistic regression for Y as follows:

$$\log(Y) = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . (3)$$

Therefore,

$$\pi = probability(Y = outcome\ of\ interest, X|X_1 = x_1, X_2 = x_2 \dots, X_n = x_n) =$$

$$\left(\frac{e^{\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n}}{1+e^{\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n}}\right) \dots\dots\dots\dots\dots\dots\dots\dots . (4)$$

For *logit* analysis techniques, the procedure hereunder was adopted.
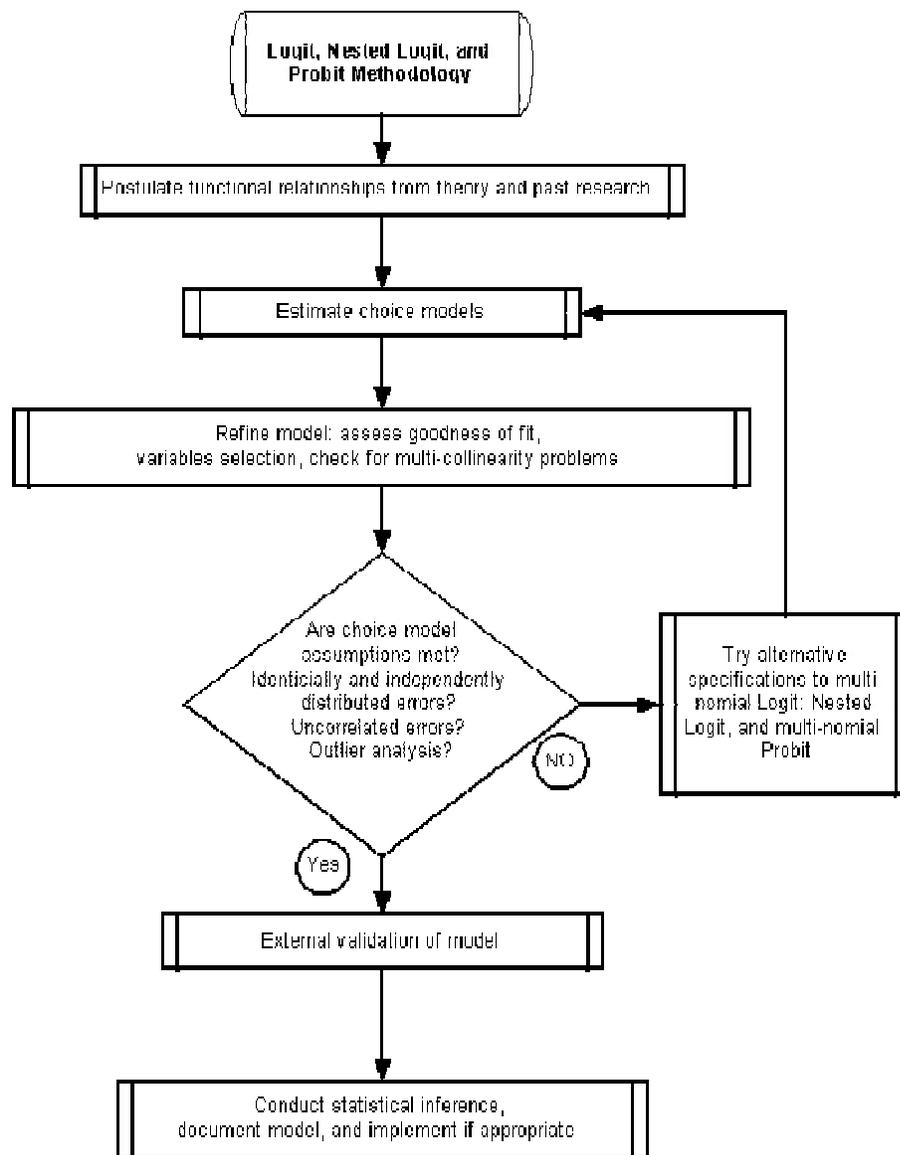
Figure 1:   **Logit, Nested Logit and Probit Methodology[18]**

[18] Adopted from http://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/v2chapter5.html; June, 2011

3.2   Data Collection

The research utilises data from the 2006 general elections results with hindsight on parliamentary elections results for incumbent considerations from the 2001 general elections. Data on income levels was obtained from the Uganda National Household Survey 2005/6 (UNHS III)[19]. Details of the data sources have been compiled in the *Appendix*.

Table 1 presents summary statistics.

3.3   Data Handling Techniques

The researcher aligned the data from the point of view that the incumbent party, National Resistance Movement Organisation (NRMO) which was transformed from a political movement, NRM, produced the eventual victor at the Presidential level in the 2006 election. In effect, other political parties were considered in this study as a coalition - having cooperated as the Interparty Cooperation or "the opposition". Therefore, in the research, responses were coded "1" if votes favoured the incumbent (NRMO), denoting victory or "0" otherwise.

The researcher took note of the fact that the 2006 general elections had three substantive elective levels for which to cast votes, namely; Presidential, Parliamentary and Local Government Council (Sub-county) Chairperson positions. Therefore accounts of preferences of the constituents have also been considered at all the three levels. The data collected have no citations for non-response cases although proxy statistics have been heavily utilised.

3.7   Description of Methods Used

The research adopted the approach of Hosmer and Lemeshow (1989) and Agresi and Finlay (1986) that support the implementation of stepwise logistic regression technique as an effective data analysis methodology. More specifically, the backward logistic regression was employed in which all independent variables were initially included in the model. At subsequent steps in

---

[19] UNHS III Report: Spatial Trends of Poverty and Inequality in Uganda, 2002/05.

the procedure, those variables found to be less significant were eliminated from the model until the remaining variables were all deemed "important."

To estimate $\beta_i s$ coefficients for the independent variables in the logistic regression model, the maximum likelihood regression fitting procedure (Warren, 1978) was explored. In general, the maximum likelihood technique was used to maximize the log-likelihood function indicating how likely it was to obtain the observed values of Y given the values of the independent variables and parameters, i.e. $\beta_i s$ (Menard, 1995).

The likelihood ratio chi-square test was used to assess significance in logistic regression since the errors were assumed to follow a binomial distribution. The test assigned a $p - value$ to each variable for assessing significance[20].

The Choice of Statistical Analysis

Logistic regression has been used in this study because it is less constrained by statistical assumptions[21]. It is known to provide more powerful and consistent predictions when some statistical assumptions are violated (Kvamme, 1983; Press and Wilson, 1978) and readily accepts mixtures of normal, ordinal, interval and ratio scaled independent variables. A model-building strategy based on an interactive backward selection of variables to obtain the best model on grounds of parsimony and goodness-of-fit was adopted for which the probability, $p - value \ll 0.05$, was rejected at 5% significance level, i.e. $0.1 \leq p \geq 0.5$.

---

[20] The most important variable being the one with the least $p - value$ $i.e$ $p<=0.05.$
[21] Discriminant Function Analysis:
    (i)   Assumes multivariate normality data
    (ii) Assumes equal covariance matrices
    (iii) Does not readily accept mixtures of categorical and interval-scale independent variables.

3.7.1 Multiple Regression and Ordinary Least Squares (OLS) Estimation

Before implementing logistic regression, the researcher examined linear regression models concept in exploring the relationship between a dependent variable Y and one or multiple independent variables X_1 to X_n; if Y varies systematically with different values for X_1 to X_n. The model is usually described by the following formula (Spicer, 2004):

$$Y = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_n X_n + \varepsilon) + r$$

Where ε is a constant representing the random error (any effects other than X_1 to X_n), β_0, β_1 and β_n are regression coefficient parameters (the amount Y changes by each one-unit change in β_i while other independent variables are held constant. Each regression coefficient represents the slope of the regression line; the larger the β_i the more influence the independent variable X_i has on the dependent variable, Y. r is the residual, i.e. the discrepancies between predicted Y-values and the actual Y-values. A common way to calculate the regression coefficients is to use ordinary least squares (OLS) estimation.

Here values for the regression coefficients, β_i, are chosen that minimize r, i.e. minimizing the summed up squared differences between all predicted Y values and the corresponding actual Y values. Hypothetically, if this model was applied to data where the dependent variable Y were categorical, the composite variable in the brackets above would no longer predict scores but probabilities of a case being in the category. So each regression coefficient, β_i, would then be the change in probability of a case belonging to the category given a one-unit increase in the independent variable, β_i. Although this might be interpretable and might seem well suited, there are several reasons why OLS regression should not be used when dealing with dependent variables that are categorical. One important reason is that the generated probabilities may fall outside the 0 to 1 range (Spicer, 2004). Instead, binary logistic regression is the recommended method.

3.7.2 Data Requirements and Related Methods

The sample size required to perform reliable estimations with the binary logistic regression method is known to vary with the number of independent variables, i.e. the more the independent variables, the more data required. However, the recommended minimum of cases per independent variable varies; Spicer (2004) refers to a recommendation of 50 cases per independent variable, while Garson (2009) refers to recommendations of 10 cases per independent variable. This study considered eight independent variables for 214 records.

3.7.3 Reporting Results

Evaluations involving logistic regression included the overall model evaluation, statistical tests of individual predictors, goodness-of-fit statistics and validation of predicted probabilities.

3.7.4 Overall Evaluation of the Model

The overall evaluation was used to demonstrate whether the logistic model fits the data closer than the intercept-only model (i.e. the model where the independent variables do not influence the dependent variable). $p-values$ smaller than 0.05 indicated that the independent variables most likely influenced the dependent variables.

3.7.5 Statistical Tests of Individual Predictors

The statistical significance of individual regression coefficients (i.e. $\beta_i s$) was tested using the Wald Chi-square statistic. For the dataset used, the test result ($p > 0.05$), suggested that an alternative model without the intercept might be applied to the data or a variable under consideration did not actually influence the dependent variable.

3.7.6 Goodness-of-fit Test Statistics

Goodness-of-fit statistics were used to assess the fit of a logistic model against actual outcomes. One inferential test and two descriptive measures have been presented in Table 12. The inferential goodness-of-fit test - Hosmer-Lemeshow (H-L) test - yielded a $\chi^2(8)\,of\;0.6714$ and was deemed only slightly significant (i.e. $p > 0.05$), suggesting the model was fit to the data fairly well. In other words, the null hypothesis of a good model fit to data was tenable.

3.7.7 Assessment of Predicted Probabilities

Logistic regression predicts the logit of an event outcome from a set of predictors. Because the logit is the natural log of odds or $\ln\left(\frac{\pi}{1-\pi}\right)$, it can be transformed back to the probability scale. The resultant predicted probabilities can then be revalidated with the actual outcome to determine if high probabilities are indeed associated with events and low probabilities with nonevents. The degree to which predicted probabilities agreed with actual outcomes was expressed as a measure of association in a Classification Table – Table 12 (Pg 56).

CHAPTER FOUR

INTERPRETATION OF RESULTS AND DISCUSSION OF FINDINGS

The analytical goal of the study was to determine the influence of the explanatory variables on the outcome of the election. These variables used were Poverty levels, Participation of the Constituency voters, Incumbency of the Candidate for the seat of President, Income levels of voters, Literacy levels of voters, Party Dominance, Location of the Constituencies by region and Nature of the Constituency.

In order to analyse the result, "Win" or "Loss", the statistical software (*Eviews 3.1*) tool was used. To generate binary logistic regression results for "Binary," the software functionality – "*Binary Choice-logit, Probit, Extreme Value*" was chosen. The estimated coefficients, their standard errors, the significance of the variables - checked using *Wald Chi*-square $p - value$ were obtained to which interpretations have been given (see List of Tables, Page *vii*).

Table 1:   **SUMMARY STATISTICS** – Dependent Variable Frequencies

Value	Count	Percent	Cumulative Count	Percent
0	63	29.44	63	29.44
1	151	70.56	214	100.00

Application of Binary Logistic Regression to the Data

Logistic regression for seven dummy variables was performed; with the results for the outcome/ dependent variable, WL, and regressors CC, IPL, LL, PCI, PD and PW. The descriptive statistics (mean and standard deviation) for the regressors have been summarized in the table below (The independent variables are specified in the *Appendix*).

Table 2:   Descriptive Statistics for key Variables

	Win or Loss	Civic Consciousness	%Individuals in Poverty	Literacy Level	Household Per Capita Income	Party Dominance	Ratio of Voters to Census
Mean	0.705607	0.696987	4.041542	58.24458	29.22140	0.405514	0.077523
Median	1.000000	0.694999	0.660000	58.13000	27.21000	0.395000	0.060000
Maximum	1.000000	0.906150	91.00000	93.14000	72.48000	1.000000	0.390000
Minimum	0.000000	0.479048	0.590000	11.00000	14.66000	0.000000	0.010000
Std. Dev.	0.456838	0.088129	17.17689	16.50073	12.97813	0.158394	0.050135
Skewness	-0.902244	0.000617	4.877223	-0.403187	1.225524	0.279100	3.822992
Kurtosis	1.814044	2.940778	24.78785	3.685224	5.296006	4.482604	20.91928
Jarque-Bera	41.57545	0.031287	5081.250	9.984620	100.5736	22.37816	3384.425
Probability	0.000000	0.984478	0.000000	0.006790	0.000000	0.000014	0.000000
Observations	214	214	214	214	214	214	214

Table 3: **Effect of Constituency Location ; Rural (RC), Urban UC), Peri-Urban (PUC)**

Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
RC	1.078810	0.280755	3.842527	0.0001
UC	-0.762140	0.457738	-1.665015	0.0959
PUC	1.109308	0.207114	5.356015	0.0000

Categorically, 22% of the Constituencies were located in urban areas, 32% in rural areas and 48% in isolated rural-urban areas (or Peri-Urban Constituencies). The chances of winning the majority votes at the Constituency for the Presidential Candidate were decreased by a multiplicative factor of 0.76 ($P = 0.0959$), rendering the Urban Constituency characteristic insignificant according to the criteria (i.e. $p < 0.05$). The chances for winning were increased by a multiplicative factor of 1.07 for a rural Constituency ($P = 0.0001$) and the chances were found to increase by a multiplicative factor of 1.01 for Peri-Urban Constituencies ($P = 0.0000$). On dropping the UC variable, the results remained unchanged in assessing the influence of the omitted Urban Constituency variable. It appears that the "rural" element in the nature of the Constituencies was a significant aspect in influencing the outcome of the election.

Nature of Constituency, WL_{Ni}:

$$log(WL)_{Ni} = \gamma_1 RC + \gamma_2 UC + \gamma_3 PUC + \emptyset \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (1)$$

$$log(WL)_{Ni} = 1.08(RC) - 0.76(UC) + 1.11(PUC) + \emptyset$$

Table 4: REGIONAL EFFECTS: Central (DC), Eastern (DE), Northern (DN) and Western (DW)

Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
DC	1.263692	0.314105	4.023146	0.0001
DE	1.139434	0.406202	2.805093	0.0050
DN	-0.485508	0.259437	-1.871388	0.0613
DW	2.926739	0.592612	4.938708	0.0000

In terms of location, 27% (31% of all voters) of the Constituencies were located in the Central region, 16% of the Constituencies (17% of all voters) in the Eastern region, 29% of the Constituencies (25% of all voters) in the Northern region and 28% of the Constituencies (28% of all voters) in the Western region; indicating a fair distribution of voting population of voters across the regions. Analysis results reveal positive chances of winning majority votes in the regions for the Presidential Candidate were significantly positive in the Central, Eastern and Western region Constituencies by respective multiplicative factors of 1.26 (P = 0.0001), 1.34 (P = 0.0050), 1.26 (P = 0.0001) and 2.93 (P = 0.0000). Contrary to the preceding, the chances of the Presidential Candidate winning in the Northern region Constituencies were reduced by a multiplicative factor of 0.49 (P = 0.0613). Clearly the Northern Constituencies presented insignificant chances of achieving victory according to criteria (i.e. P <0.05).

Logistic Regression for Regional Effects, WL_{Ri}:

$$\log(WL)_{Ri} = \alpha_1 DC + \alpha_2 DE + \alpha_3 DN + \alpha_4 DW + \lambda \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

$$\log(WL)_{Ri} = 1.26(DC) + 1.34(DE) - 0.49(DN) + 2.93(DW) + \lambda$$

Table 5: Incumbency Effect Dummy Variables:

> Local Government (DLG), Seats for Member of Parliament (DMP) won by Party Candidates and Constituencies won by Incumbent (DIP) in Preceding Election as Candidate for the Seat of President (DIP)

Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
DMP	0.812500	0.366727	2.215546	0.0267
DIP	2.678300	0.587370	4.559814	0.0000

Three incumbency perspectives were analysed, namely; the Local Government positions won in the election by the overall Presidential Candidate (DLG), Seats for Member of Parliament won by the Party Candidates allied to the party of the Presidential Candidate (DMP) and the Constituencies won by the incumbent Presidential seat Candidate of the winning party (DIP). Analysis results show positive chances of winning majority votes by the Presidential seat Candidate with a multiplicative factor of 0.93 ($P = 0.0211$) if the Candidates vying for Parliamentary seats at the Constituencies also won those seats. The chances were even higher that the incumbent's chances of winning elections in all the Constituencies were increased by a multiplicative factor of 2.64 ($P = 0.0000$). Thus the incumbency effect for winning elections at the Constituencies by the Presidential Candidate and for the shared party Candidates at the same level contributed significantly to a positive chance for winning the election. Conversely, results show that the chances of the Presidential Candidate winning in the Constituencies were negatively affected by the influence of local government positions won by a multiplicative factor of 0.46 ($P = 0.4617$). This result represents an insignificant influence according to criteria (i.e. $P < 0.05$).

Incumbency Factor, WL_{Ii}:

$$\log(WL)_{Ii} = \lambda_1 DIP + \lambda_2 DMP + \lambda_3 DLG + \gamma \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

$$\log(WL)_{Ii} = 2.64(DIP) + 0.93DMP - 0.37(DLG) + \gamma$$

Combined Effect of Dummy Variables:

(WL_D, the result for *Regional, Nature of Constituency* and *Incumbency* of variables)

$$\log(WL)_D = log(WL)_{Ri} + log(WL)_{Ni} + \log(WL)_{Ii}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Results of Binary Logistic Regression for the Non-dummy Variables

The following results were obtained[22]:

Table 6:   LRA output for Predictors

Convergence achieved after 6 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-12.96830	2.559873	-5.065995	0.0000
CC	12.66590	3.125560	4.052361	0.0001
IPL	-0.012124	0.011718	-1.034638	0.3008
LL	-0.087794	0.029228	-3.003751	0.0027
PCI	0.104824	0.022913	4.574923	0.0000
PD	6.877937	2.362946	2.910747	0.0036
PW	12.25706	8.557869	1.432256	0.1521

On backward selection; dropping the variables LL, IPL and PW the table below was obtained:

Table 7:   LRA output for most Important Variables

Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-10.79431	2.184910	-4.940389	0.0000
CC	8.733892	2.591544	3.370150	0.0008
PCI	0.064354	0.015654	4.110932	0.0000
PD	10.85327	2.005779	5.411000	0.0000

---

[22] This can be examined by the likelihood ratio tests and score tests (Peng *et al*, 2002). $p-values$ smaller than 0.05 indicate that the independent variables most likely influence the dependent variable. Consequently, all other variables were dropped from the model (Constant included in the regression).

(a) Interpretation of Analysis Results for Individual Regressors

1. For every unit increase in the number of persons conscious in civic duties (CC), the odds in favour of winning a constituency are increased by a multiplicative factor of 4.88. The reported standard error is 2.7070 and the statistical significance assessed by the Wald chi-square statistic is 2.613, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.0717). This would support the conclusion that CC is a useful predictor of a "positive" outcome although the probability value of 0.0717 is higher than what is acceptable (<0.05).

2. For every unit increase in the percentage of individuals surviving on less than One United States Dollar (USD 1.00) per day in a Constituency i.e. (below the poverty line), the odds in favour of winning a constituency are reduced by a multiplicative factor of 0.01. The reported standard error is 0.011143 and the statistical significance assessed by the Wald chi-square statistic is 1.467, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.2581 (>0.05)). This supports the conclusion that IPL is not a useful predictor of a "positive" outcome.

3. For every unit increase in the number of literate persons (LL), the odds in favour of winning a constituency are decreased by a multiplicative factor of 0.06. The reported standard error is 0.028192 and the statistical significance assessed by the Wald chi-square statistic is 4.498, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.0285).

4. For every unit increase in the average monthly income of household of individuals in a Constituency (PCI), the odds in favour of winning a constituency are increased by a

multiplicative factor of 0.07. The reported standard error is 0.019314 and the statistical significance assessed by the Wald chi-square statistic is 13.278, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.0002, which is <<0.05 in Table 3). This supports the conclusion that PCI is a useful predictor of a "positive" outcome.

5. For every unit increase in the Proportion of Local Government positions won in the Constituency by the overall winning party (the Dominance Factor, PD), the odds in favour of winning the constituency are increased by a multiplicative factor of 9.04. The reported standard error is 2.259686 and the statistical significance assessed by the Wald chi-square statistic is 16.276, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.0001). This supports the conclusion that PD is a useful predictor of a "positive" outcome.

6. For every unit increase in the proportion of registered voters to census population per Constituency (PW), the odds in favour of winning a constituency are increased by a multiplicative factor of 13.17. The reported standard error is 8.211668 and the statistical significance assessed by the Wald chi-square statistic is 2.302, i.e. with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed $p - value$ is reported to be 0.1087 which is much greater than the conventional value $p = 0.05$). This supports the conclusion that PW is not a useful predictor of a "positive" outcome.

<p style="text-align:center">The Regressor (Quantitative) Variables' Model</p>

The regression models for estimation of the "Win" or "Loss", WL, situations can be expressed as follows:

$$log(WL)_{Qi} = \beta_0 + \beta_1 CC + \beta_2 PCI + \beta_3 PD + \delta \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (5)$$

i.e.

$$log(WL)_{Qi} = -10.79 + 8.73 * Av\_CC + 0.06 * PCI + 10.85 * PD + \delta$$

Models for Dummy Variables: Calculation of Probabilities for Winning in a Given Constituency

Probability of a Winning an Election in a Selected Constituency

$$P = \frac{e^{(\beta_0+\beta_1 CC+\beta_2 PCI+\beta_3 PD+\delta)}}{1+e^{(\beta_0+\beta_1 CC+\beta_2 PCI+\beta_3 PD+\delta)}} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

i.e.

$$P = \frac{e^{(-10.79+8.73*Av_{CC}+0.06*PCI+10.85*PD)}}{1 + e^{(-10.79+8.73*Av_{CC}+0.06*PCI+10.85*PD)}}$$

Table 8:   **Results for Model Evaluation and Calculated Probabilities**

Outcome	Constituency	Variables				Model Evaluation	Probability
WL=1	Lebel	Av_CC	PCI	PD	β_0	$log(WL)_{Qi}$	P
		8.73	0.06	10.85	-10.79		
	84	0.63	40.13	0.77		5.2874	0.995
	107	0.66	40.13	0.33		0.8809	0.707
	108	0.59	40.13	0.50		2.0735	0.888
	164	0.53	40.54	0.28		-0.7599	0.319
	176	0.67	40.13	0.44		2.1353	0.894
WL=0							
	61	0.64	72.48	0.00		-3.8990	0.020
	64	0.65	72.48	0.00		-8.1642	0.000
	106	0.74	40.13	0.56		-1.5036	0.182
	121	0.54	43.54	0.25		-5.6967	0.003
	159	0.58	40.13	0.40		-3.9244	0.019
	160	0.70	40.13	0.33		-4.1247	0.016
	69	0.70	72.48	0.00		-7.6260	0.000
	67	0.67	72.48	0.00		-7.7616	0.000

Combining the equations (4) and (5) for the Dummy and Quantitative variables, the overall

model for estimating the probability for a dichotomous outcome, $P(WL)$, is

$$P(WL) = \frac{e^{\sum_{\forall_i}[\log(WL_{Di}) + \log(WL_{Qi})]}}{1 + e^{\sum_{\forall_i}[\log(WL_{Di}) + \log(WL_{Qi})]}}$$

Where,

WL :   Dichotomous outcome; "1" for a "Win" and "0" "loss" of election contest.

WL_{Di} :   Outcome for Dummy Variables in which "1" is a "Win" and "0," a "loss."

WL_{Qi} :   Outcome for Quantitative Variables in which "1" is a "Win", "0," a "loss."

Table 9:   **Statistical Significance of Individual Regression Coefficients**

Wald Test:			
Equation: Untitled			
Null Hypothesis: C(1)=0			
	C(2)=0		
	C(3)=0		
	C(4)=0		
F-statistic	10.93725	Probability	0.000000
Chi-square	43.74900	Probability	0.000000

The reported statistic is 43.74900, with a $p-value \leq 0.000000$ (Durbin-Watson Statistics)

indicating that the null hypothesis that $C(1) = C(2) = C(3) = C(4) = 0$ is strongly rejected.

Table 10:   Goodness-of-Fit Test

Dependent Variable: WL
Method: ML - Binary Logit
Date: 08/26/11   Time: 17:47
Sample: 1 214
Included observations: 214
Andrews and Hosmer-Lemeshow Goodness-of-Fit Tests
Grouping based upon predicted risk (randomize ties)

| | Quantile of Risk | | Dep=0 | | Dep=1 | | Total | H-L |
	Low	High	Actual	Expect	Actual	Expect	Obs	Value
1	0.0591	0.3399	17	16.0321	4	4.96791	21	0.24702
2	0.3404	0.4337	12	12.7880	9	8.21197	21	0.12418
3	0.4341	0.5705	11	11.2947	11	10.7053	22	0.01580
4	0.5956	0.6535	6	7.86942	15	13.1306	21	0.71024
5	0.6570	0.7286	10	6.61796	12	15.3820	22	2.47197
6	0.7313	0.8421	4	4.44879	17	16.5512	21	0.05744
7	0.8494	0.9167	2	2.27483	19	18.7252	21	0.03724
8	0.9174	0.9679	0	1.15439	22	20.8456	22	1.21832
9	0.9689	0.9909	1	0.42279	20	20.5772	21	0.80424
10	0.9912	0.9998	0	0.09699	22	21.9030	22	0.09742
		Total	63	63.0000	151	151.000	214	5.78388

H-L Statistic:	5.7839	Prob. Chi-Sq(8)	0.6714
Andrews Statistic:	42.9743	Prob. Chi-Sq(10)	0.0000

The goodness-of-fit test was carried out using the Hosmer-Lemeshow (1984) and Andrew's (1988) tests to compare the fitted expected values to the actual value by group.

Four variables were entered into the model, viz. AC, GC, PL and PW. The $p-values$ of the Hosmer-Lemeshow (H-L) and Andrews Test statistics are indicated at the bottom of the results table. The p-value for the H-L test was large while that for the Andrews test statistic was small, providing mixed evidence for significance. Therefore caution in interpreting the results was necessary.

Table 11: **Expectation - Prediction (Classification) Table**

Dependent Variable: WL
Method: ML - Binary Logit
Date: 08/26/11   Time: 17:47
Sample: 1 214
Included observations: 214
Prediction Evaluation (success cutoff C = 0.5)

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=C	36	18	54	0	0	0
P(Dep=1)>C	27	133	160	63	151	214
Total	63	151	214	63	151	214
Correct	36	133	169	0	151	151
% Correct	57.14	88.08	78.97	0.00	100.00	70.56
% Incorrect	42.86	11.92	21.03	100.00	0.00	29.44
Total Gain*	57.14	-11.92	8.41			
Percent Gain**	57.14	NA	28.57			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	32.34	30.66	63.00	18.55	44.45	63.00
E(# of Dep=1)	30.66	120.34	151.00	44.45	106.55	151.00
Total	63.00	151.00	214.00	63.00	151.00	214.00
Correct	32.34	120.34	152.67	18.55	106.55	125.09
% Correct	51.33	79.69	71.34	29.44	70.56	58.45
% Incorrect	48.67	20.31	28.66	70.56	29.44	41.55
Total Gain*	21.89	9.13	12.89			
Percent Gain**	31.02	31.02	31.02			

*Change in "% Correct…"[23]

The gain in the number of correct predictions was obtained by moving from the right of the table to the left in determining the measure of the predictive ability of the model. The gain measures are reported in absolute percentage increases (Total Gain) and as a percentage of the incorrect classifications in the constant probability model (Percent Gain). 63 records have $Dep = 0$. This prediction is correct for the 63 (y=0) observations, and 151 for the ($Dep = 1$) observations. Thus, the restricted model predicts 70.56% (12.89 *Total gain* and 31.02 *Percent Gain*).

---

[23] *Change in "% Correct"[23] from default (constant probability) specification; **Percent of incorrect (default) prediction corrected by equation.

PROBABILITY CURVES
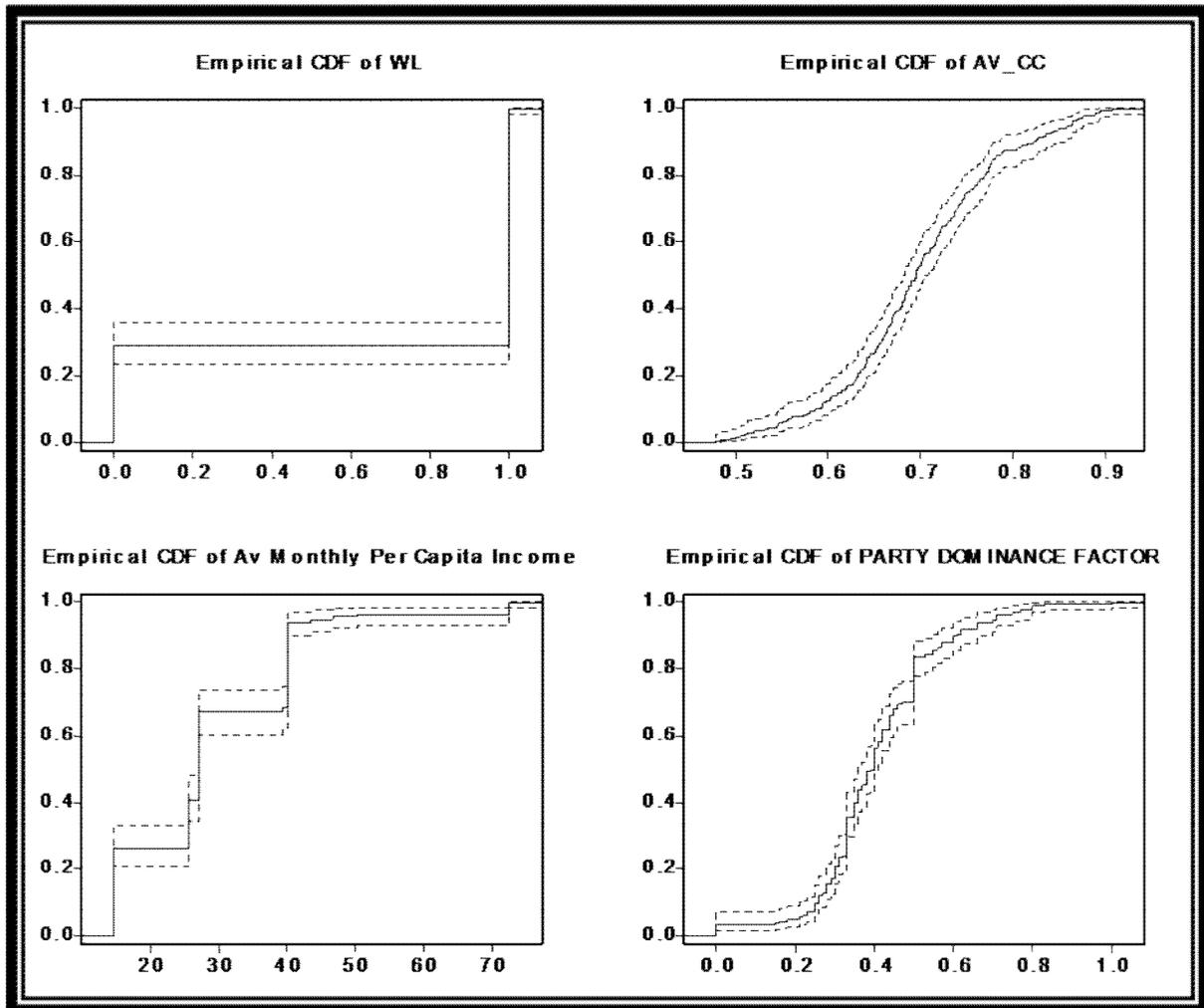
Figure 2: **Plots of the relationships between winning or losing, and the variables *AV_CC, PCI* and *PD*.**

The probability curves suggest the probability of winning the election at the Presidential level increases with the number of Constituencies. However, it can be seen that the relationship is nonlinear and that the probabilities hang much more at the high or low extreme levels. These patterns are typical since probabilities cannot lie outside the range 0 to 1. The relationships can be described as following an 'S'-shaped curve.

Key theoretical underpinnings on the findings

- From political economy literature the findings suggest that redistributive spending by government should be higher in the more unequal constituencies due to median voter preferences. Although voters do not look exclusively at economic issues, they generally weigh those more heavily than any others, regardless of the democracy they vote in.

- On the scale of political participation, the data reliably captures the behaviour across a range of statistically relevant and theoretically meaningful civic activities or observations. The key understanding here is that high levels of civic consciousness and literacy are necessary to sustain widespread political participation.

- A contextual effect where income inequality is related to the weight that voters place on income is not explicitly depicted in the voting decisions but is depicted in the analysis. That is, rewarding or punishing, especially, the incumbent party at the ballot box.

Strategic and Technical Implications for Policy

For strategic and technical policy implications, political actors and coalition analysts can help strategic voters make informed choices in multiparty systems through developing further insights from the findings in this research. Similarly, governments could optimise their vote chances or decisions regarding initiation of people centred policies. Coalition or governments-in-waiting could consider articulating alternative policy matters in their campaign strategies needed to win elections.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

From the foregoing discussion, it can be concluded that a predictive baseline Constituency characteristic could have led to a potentially important increase or decrease in the likelihood of a certain result during the 2006 Presidential poll. Key theoretical and practical causal claims leading to the outcomes were analysed, paying particular attention to the pre-election environment in the Constituencies.

5.1   SUMMARY OF FINDINGS

The main purpose of the research was to apply the methodology for analysis of dichotomous outcomes to study the case of the 2006 Presidential poll in Uganda. The characteristics were depicted by data for all 214 Constituencies. The approach of Hosmer and Lemeshow (1989) and Agresi and Finlay (1986) utilizing the stepwise logistic regression procedure was used as an effective data analysis technique.

The specific objectives that dealt with Constituency characteristics entailing poverty, income and literacy levels were found not to have influenced the outcome of the Presidential election. Participation of the Constituency voters and the extent to which a party dominated politics at the Constituency level influenced the result. Similarly, candidate incumbency was a largely responsible factor in influencing the results as were regional locations and the rural-urban nature of the Constituencies. These were deemed complementary factors that could have delivered victory.

5.2   CONCLUSION

The study was motivated by the need to model polls' outcomes based on constituency characteristics as likely determinants for winning or losing an election in Uganda. This followed as a backdrop to contested elections whose validity could only be ascertained through ballot re-counts or circumstantial evidence. The researcher applies the methodology for analysis of dichotomous outcomes in order to predict the polls' outcome using constituency characteristics' data. An attempt is made to ascertain which variables were central to influencing the poll's outcome. On analyzing the variables, all other quantitative predictors, except participation of the Constituency voters, average monthly per Capita incomes of voters and dominance by a political party, were eliminated. These were found to be statistically reliable for the restricted model in correctly predicting 70.56% of the data.

5.3   RECOMMENDATIONS

For the political actors, it is important they align their campaign strategies in addressing or countering those influences on polls' outcomes with what is or without their control. Coalition candidate should consider newer people-centered strategies as the study seems to confirm winning advantage on the incumbent.

Future research should examine the influences of factors that may emerge in subsequent elections. For example these factors may include legal reforms, money in politics (or money power) and other external influences such as regional stability and the need to align to global economic (or market forces) and political demands.

REFERENCES

1. Cox, D.R., & Snell, E.J. (1989). <u>*The analysis of binary data*</u> (2nd ed.); London:Chapman and Hall.

2. Christensen, Ronald (1990). *Log-Linear Models.* Springer-Verlag. New York, New York.

3. Garson, G.D. (2009). "Logistic Regression" from *Statnotes: Topics in Multivariate Analysis*; <u>http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm. 6/11/2009.</u>

4. Harrell, F.E. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis* .UK, Springer, 2001

5. Henry Milner (2007), *Civic Literacy in Comparative Context – Why Canadians Must be Concerned*; Enjeux Publics, Vol.2, No.2, July 2001.

6. Hosmer, D.W. and Lemeshow, S. *Applied logistic regression.* US, Wiley-Interscience, 2000

7. McCullagh, P. and J.A. Nelder, (1983). *Generalized Linear Models.* Second Edition. Chapman and Hall, New York, New York.

8. Peng, Chao-Ying Joann; Lee, Kuk Kida & Ingersoll, Gary M. (2002). *An introduction to logistic regression analysis and reporting*. Journal of Educational Research 96(1): 3-13.

9. Spicer, J. (2004). *Making sense of multivariate data analysis*. Sage Publications, California, 123-151.

10. Strömbergsson S., (Spring, 2009). *Binary logistic regression and its application to data from a study of children's recognition of their own recorded voices*: Term paper in Statistical Methods.

APPENDIX : DEFINITION OF VARIABLES

No.	Variable ID	Description	Variable Type	Sampling Technique	Source
1	WL	Dichotomous outcome for which "1" implies a "Win" and "0" is "loss" of electoral contest.	Dependent; characterizing Win or Loss	Scored as "1" for Win or "0" for Loss depending on whether the overall winning party won or lost in the EA.	2006 Election data
2	Av_CC	Civic participation; an attribute of the voters determined from routine participation during General Elections.	Independent	Average of the ratios of total votes cast to the number of registered voters in each Constituency compared for 2001/6 with respective means (0.699; 0.694) and standard deviations (0.1185; 0.775). The resultant average has mean and standard deviation 0.697, 0.088 respectively.	2006 Post - election data, Uganda Electoral Commission.
3	LL	Literacy Level is defined in the source document as comprising ability to read, write and carry out simple numerical tasks.	Independent	A national aggregate minimum literacy level considered Primary two (P.2) or 10 year olds because then most children went to school or began schooling late. However, in this study, the researcher is concerned with adult literacy, i.e. from 18 years on. UNHS/REPORT Pg. 15.	"POST ENUMERATION SURVEY: 2002 Uganda Population and Housing Census," 2005

4	PD	This is the party dominance factor defined as the aggregate (total) of seats won by the overall winning party. This is the infrastructure upon which the party thrives in the constituency.	Independent	Proportion of elective seats won in the Constituency by the overall winning party.	2006 Election data, Uganda Electoral Commission
5	PW	A dynamic, momentous, characteristic for the proportion of voters in the Constituency who may turn out to vote.	Independent Variable	Proportion of registered voters to census population per Constituency for each of the 214 EAs.	2006 Election data vis-à-vis national census data
6	PCI	Per Capita Income; This has been computed basing on average monthly household incomes provided per region. Also average household sizes are provided in survey data at the regional levels.	Independent Variable	Report gives regional data on average monthly income; further average household size data is obtained for use in computing per average capita income.	Uganda Human Development Report 2007 *Rediscovering Agriculture for Human Development.*
7	IPL	Percentage of Individuals Living below poverty line in the constituency (EA)	Independent Variable	The poverty estimates were derived for each county in separate analyses.	*Uganda Rural-Urban Poverty Rates by County, 1992*
8	RU	Constituency Characteristic; Rural - Urban	Dummy	A Constituency characterized as RURAL or URBAN is assigned the value dummy value "1" if so, else "0."	2006 Election data, Uganda Electoral Commission
9	RC	Characteristic; Rural Constituency	Dummy	Constituencies characterized as Rural is identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission

10	UC	Characteristic; Urban Constituency	Dummy	Constituencies characterized as Rural is identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission
11	PUC	Characteristic; Peri-Urban Constituency: Area immediately adjoining an urban area; between the suburbs and the countryside. They are characterized by rural and urban features and may consist of highly heterogeneous and rapidly changing socio-economic groups[24].	Dummy	The selection of the Constituencies comprised those Electoral Areas within which there are Town Councils.	2006 Election data, Uganda Electoral Commission
12	DC	Region Characteristic for Central Uganda	Dummy	Districts comprising the region are identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission
13	DE	Region Characteristic for Eastern Uganda	Dummy	Districts comprising the region are identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission
14	DW	Region Characteristic for Western Uganda	Dummy	Districts comprising the region are identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission
16	DN	Region Characteristic for Northern Uganda	Dummy	Districts comprising the region are identified and assigned a dummy value "1."	2006 Election data, Uganda Electoral Commission

---

[24] Governance of Water and Sanitation for the Peri-Urban Poor (Adriana *et al, 2006*)

17	DLG	Incumbency Effect due to Local Government seats won by Party local council candidates likely aiding Incumbent for seat of President	Dummy	Local Council seats Won/lost in 2006 assigning "1," else "0" for winning President's party.	2006 Election data, Uganda Electoral Commission
18	DMP	Incumbency Effect due to Parliamentary Constituency won by Party Candidate likely aiding Incumbent for seat of President	Dummy	Constituency for MP Won/lost in 2006 assigning "1," else "0" for winning President's party.	2006 Election data, Uganda Electoral Commission
19	DIP	Incumbency Effect for President	Dummy	Compared wins/losses in winning EAs in 2001/6 assigning "1," else "0."	2001/6 Election data, Uganda Electoral Commission