
Johannes W. Fedderke* and Margaux Giannaros†

October 26, 2018

Abstract

Recent surprise election outcomes may be the result of nonlinearities in voting behaviour. We demonstrate that in a simple two issue, two party voting model the general case implies nonlinearities in voting due to interactions between the issues that influence voters. Empirical testing against an election characterized by a contest between two main parties, across two main issues is provided by the South African 2009 national election. We use a novel data set, which combines census and voting data at the ward level. Allowing for an interaction between income and race reduces the impact of racial identity for the major parties, and confirms the nonlinearity in voting behaviour predicted by our theory. Small changes in ward composition along the standard issues influencing voting may therefore contribute toward explaining recent surprises in electoral outcomes.

*Pennsylvania State University, Economic Research Southern Africa, jwf15@psu.edu
†University of the Witwatersrand.
1 Introduction

Recent elections have delivered surprising outcomes. In the United States and the United Kingdom political candidates or policy options that were not anticipated to win, did so nonetheless. In Europe, the political centre has been under significant challenge from populist alternatives from more extreme parts of the ideological spectrum. Developing countries too have seen an emergence of populism (Indonesia), a waning of historically dominant parties (South Africa), conventional or historical patterns of policy tensions (Brazil, Argentina), and strong challenges to historical elite monopoly control over power (the Arab spring), amongst other examples. The extent of political upheaval and dynamic change seems unusually high, irrespective of the level of economic development.

Plausible responses are to argue that the factors determining electoral outcomes have been subject to change, with a declining importance of the role of established party elites,\(^1\) increased partisanship,\(^2\) with new drivers in the form of specialist economic and cultural interests becoming important,\(^3\) declining trust in institutions, new fault lines such as urban-rural replacing historical ideological differences such as liberal-conservative,\(^4\) that changes in the role of advertising, news and social media result in novel means of information dispersal,\(^5\) that the impact of outside powers (eg. Russia on the US) on election outcomes has been magnified, even that voter decisions are irrational.\(^6,7\)

Such responses are plausible. This paper adopts another route, adding to the set of possible explanations. We present a simple theory of voting in a two issue (say culture and economic interests), two party electoral space, that predicts that in general voting outcomes in a proportional representation system at the electoral district level will be subject to nonlinearities. In particular, the two issues motivating voters will be subject to interaction. This implies that relatively small changes in electoral district characteristics over time, can result in marked changes in voting behaviour, not because the motivating drivers of voting behaviour have changed, but because of nonlinearities that result from their interaction.

\(^1\)Contra Cohen et al (2008), and see Azari (2016), Schlozman and Rosenfeld (2017), for new forms of party relevance, Maneto (2018), for anti-establishment sentiment.
\(^4\)Cramer (2016).
\(^6\)Achen and Bartels (2016), Gadarian and Albertson (2015).
\(^7\)See also the discussion in Gelman and Azari (2017).
The explanation of the apparent changes in political voting outcomes has Occam’s razor on its side - small changes in established explanatory variables account for the observed changes. In this it is consistent with accounts that emphasize continuity in the explanations of political voting outcomes, despite what appears like upheaval. The paper also presents an empirical test of the theory, employing South African 2009 national election data at the ward level. Choice of the South African case is determined by the fact that the principal electoral contest in 2009 was between two parties (the African National Congress (ANC) and the Democratic Alliance (DA)), and along two clearly identified voter interests (race and economic interests). We demonstrate that the historical platform differences (race, economics) remain important in 2009. But we also show that a set of electoral wards that show aberrant electoral outcomes are associated with interactions between race and income, and demonstrate that the implied nonlinearities suggest significant substantive variation in voting behaviour. The implication is that relatively small changes in electoral wards, along the two historical dimensions of voter interest, may account for significant changes in voting outcomes over time. The decline in the fortunes of the African National Congress, the relative rise of the Democratic Alliance and a party of the left (the Economic Freedom Fighters) may be due to these small reconfigurations of ward composition, rather than a fundamentally new politics.

Novelty of the paper is that it demonstrates theoretically that even in the most simple two issue, two party voting model, the general case creates an expectation of nonlinear associations between outcomes and voter interests. Further, it tests the theoretical prediction against a developing country data set that is well suited to a two issue, two party setting, and confirms the core prediction of nonlinear voting outcomes. In doing so, it also provides a novel extension of the understanding of voting outcomes in post-democratization South Africa, and provides some insight as to why the liberation movement party has lost ground to both ends of the ideological spectrum. Most generally, the claim of the paper is not that the framework we present can explain every one of the surprising political developments with which we opened the discussion. Instead we suggest simply that it may provide one of the perhaps multiple reasons why electoral outcomes have been experiencing such turbulence.

This paper represents an advance in the analysis of South African voting in creating a novel data set

---

8 See for instance Erikson and Wlezien (2016).
which combines census and voting data at the ward level for the 2009 national election. Few previous studies have undertaken an analysis of ward level voting behaviour linked to census demographic data.\textsuperscript{9} We link the proportion of the vote attained by the largest seven parties in the 2009 South African national election to a range of demographic characteristics covering race, income, labour market status, education, household size and status, as well as regional characteristics. Our analysis emphasizes the results for the ANC and DA. While we confirm conventional findings of the impact of race and income (rising proportion of Black ward population favours the ANC, rising proportion of rich ward population favours the DA), we also find associations that are surprising in that they overturn conventional wisdom. Specifically, we find that the DA loses disproportionately little support amongst poor Black voters, while the ANC does disproportionately worse than expected among the same class of voters. Instead, ANC support is concentrated among labour force participants that earn relatively high incomes, and thus might be considered a labour force "elite," though not among high income voters of any race group. While in 2009 the DA was the only credible voice of opposition, the emergence and relative success of the EFF, is potentially attributable to the weakness of the ANC in representing the interests of the poorest in society.\textsuperscript{10}

The paper is structured as follows. Section 2 presents our theory. Section 3 by providing background on South Africa clarifies why conventional explanations of its election outcomes are no longer adequate. In section 4 we detail the data set employed in the study, while section 5 reports results testing for the interactions implied by our theory. Section 6 provides conclusions and evaluations.

2 A simple theory of voting in two issue, two party political spaces

There are many theories of voting behaviour. Voting has been viewed as an expression of group identity,\textsuperscript{11} and expression of party identification,\textsuperscript{12} a response to specific issues of interest to voters,\textsuperscript{13} and as an expression

\textsuperscript{9}The exception is a new body of work, that does examine disaggregated data from census and survey sources, on South African voting behavior - see for instance De Kadt (2016), Sands and De Kadt (2016), De Kadt and Lieberman (2017), De Kadt and Larreguy (2016). Kroth et al (2016).

\textsuperscript{10}It needs to be remembered that electoral ANC support in 2009 was overwhelming. Under the South African proportional representation electoral system, the ANC won by an almost two thirds majority (65.8%), followed by the DA (16.8%) and COPE (7%). The evidence of this paper is that there is a disproportionately large relative gain or loss under the conditions identified.

\textsuperscript{11}See Lazarsfeld et al (1968).

\textsuperscript{12}See Campbell et al (1960).

\textsuperscript{13}See Key and Cummings (1966) and Nie et al (1999).
of rational choice.\textsuperscript{14} Our objective is to remain agnostic as to the source of voters’ interest. instead, we present a simple, hence general model in which two dimensions of importance to voters, map into voting behaviour. The underlying issues of importance in principle could be any of group, party affiliation, specific issues, or costs and benefits experienced by voters. As such, they could fit any of the substantive voting frameworks in the literature. For the sake of illustration, and to anticipate our empirical application, we term the two dimensions "identitarian" and "technocratic policy delivery," to illustrate that the model spans any of the above motivating forces of voting.

The literature also has many two issue, two party voting models. Beyond Downs (1957), these include models demonstrating that given voter preferences defined across two issues (eg. tax policy and race), with parties competing through the presentation of two-dimensional policy platforms, resultant equilibria will not guarantee the maximization of voters’ economic interests,\textsuperscript{15} models that endogenize voter interests,\textsuperscript{16} and numerous empirical applications of such models.\textsuperscript{17} Instead of the specificities of the equilibria explored in these contributions, the simplicity of our model provides a generality intended for wide application.

Toward this end, consider an electoral contest between two political parties, dominant in the political space. One is defined by a focus on identity politics, denoted party $A$.\textsuperscript{18} The other party is defined by a focus on technocratic policy delivery, denoted party $D$.\textsuperscript{19}

\section{The feasible electoral space for contestation}

The total electorate, hence the number of potential votes in any electoral district, $i$, is denoted $T_i$.

We allow for the existence of a range of fringe parties, defined either by narrow single issue interests, or by small minority identities (eg. small ethnic minorities). For the sake of analytical simplicity, we assume that fringe party voters are not open to defection to the two dominant parties. The remaining voters constitute

\textsuperscript{16}For instance, in Shayo (2009) in a voting model driven by identity and redistribution, identity becomes endogenous.
\textsuperscript{18}There is a large literature on the importance of identitarian issues on voting. See for instance Beck et al (2002), Conover (1984), Miller et al (1991). Our interest is not in the importance of identity issues per se, but in whether they interact with economic interests.
\textsuperscript{19}Again, there is ample literature that examines the impact of economic interests on voting outcomes. See for instance Edwards and Tabellini (1990), Evans (2000), Persson and Tabellini (2006, 2009), Glaeser and Ward (2006), Luttmer (2001). And again, our interest is not in the importance of economic policy issues per se, but in whether they interact with identitarian interests.
the feasible space over which the two dominant parties engage in their contest. Given the votes cast for fringe parties in electoral district $i$, $V_{F,i}$, this leaves contestation in district $i$ between parties $A$ and $D$ over the feasible electoral space $T_{F,i} = (T_i - V_{F,i})$.

Voters in the feasible electoral space for the political contest are responsive to both identitarian issues and issues related to technocratic policy delivery, as defined above. Identity issues could be defined by race, gender, ethnicity, sexual orientation, or some combination of these and other characteristics deemed defining of identity. Policy delivery could be defined by economic, housing, education, welfare, foreign policy objectives, or some combination of these and other policy objectives. In both instances, allow a measure space to represent the extent to which identitarian or technocratic policy delivery issues are of importance in an electoral district, denoted $X_i \in [0, \infty)$ for identitarian issues, denoted $Y_i \in [0, \infty)$ for technocratic policy delivery issues. This generates the mapping into voting behaviour, $V_j (X_i, Y_i)$, $j \in (A, D)$, such that $\sum_j V_j (X_i, Y_i) = T_{F,i}$.

The proportion of the feasible electoral space that is partisan for a party focussed on identitarian policy delivery, is defined by a vote mapping, $\nu_A = V_A (X, Y) / (V_A (X, Y) + V_D (X, Y))$, such that $\partial V_A / \partial X > 0$, $\partial V_A / \partial X > \partial V_A / \partial Y$, and it does not preclude the possibility that $\partial V_A / \partial Y < 0$. The proportion of the feasible electoral space that is partisan for a party focussed on technocratic policy delivery, is defined by a vote mapping, $\nu_D = V_D (X, Y) / (V_A (X, Y) + V_D (X, Y))$, such that $\partial V_D / \partial Y > 0$, $\partial V_D / \partial Y > \partial V_D / \partial X$, and it does not preclude the possibility that $\partial V_D / \partial X < 0$.

Thus identitarian voters are "more" responsive to identitarian issues than technocratic policy delivery, while technocratic policy delivery voters reverse this ordering. Note that both identitarian and technocratic policy delivery partisan voters can in principle be "swing" voters, provided that $\partial V_A / \partial Y > 0$ and $\partial V_D / \partial X > 0$. For the sake of simplicity we exclude the existence of "pure" swing voters defined by indifference between identitarian and technocratic policy delivery questions, such that $\partial V_j / \partial X = \partial V_j / \partial Y > 0$. Thus $\nu_A + \nu_D = 1$, and the $\nu_A, \nu_D$, constitute the electoral base of the parties focussed on identitarian and technocratic policy delivery respectively.
2.2 The political parties

The two dominant political parties who are engaged in the main political contest, are defined by their policy platforms. One is defined by a focus on identity politics, denoted party $A$. The other party is defined by a focus on technocratic policy delivery, denoted party $D$.

Denote $P_{X,j}$, $P_{Y,j}$, $j \in (A, D)$, as the number of policies that address identity and technocratic policy delivery respectively.

The identitarian party, $A$, is defined as such by virtue of that fact that:

$$\pi_{X,A} = \frac{P_X}{P_X + P_Y} > 0.5 > \frac{P_Y}{P_X + P_Y} = \pi_{Y,A}, \quad 0 \leq \pi_{X,A}, \pi_{Y,A} \leq 1,$$
$$\sum_k \pi_{k,A} = 1, \quad k \in (X, Y)$$

i.e. that the majority of its policy platform is given by policies that address identitarian issues.

The technocratic policy delivery party, $D$, is defined as such by virtue of that fact that:

$$\pi_{X,D} = \frac{P_X}{P_X + P_Y} < 0.5 < \frac{P_Y}{P_X + P_Y} = \pi_{Y,D}, \quad 0 \leq \pi_{X,D}, \pi_{Y,D} \leq 1,$$
$$\sum_k \pi_{k,D} = 1, \quad k \in (X, Y)$$

i.e. that the majority of its policy platform is given by policies that address technocratic policy delivery.

Any choice $\pi_{X,A}$ defines $\pi_{Y,A}$, symmetrically for $\pi_{X,D}$, $\pi_{Y,D}$.

2.3 The realized vote of the dominant parties

The actual proportion of the feasible electoral space that the two parties will actually realize in any electoral district, $\rho_{j,i}$, $j \in (A, D)$, is determined by the orientation of their policy platform, and how this motivates
the feasible electoral space. Thus:

\[
\rho_{A,i} = R_{X,A} (\pi_{X,A}) v_{A,i} + R_{Y,A} (\pi_{X,A}) v_{D,i}, \quad 0 \leq R_{X,A} (\pi_{X,A}) \leq 1, \quad 0 \leq R_{Y,A} (\pi_{X,A}) \leq 1 \tag{3}
\]

\[
\rho_{D,i} = R_{X,D} (\pi_{Y,D}) v_{A,i} + R_{Y,D} (\pi_{Y,D}) v_{D,i}, \quad 0 \leq R_{X,D} (\pi_{Y,D}) \leq 1, \quad 0 \leq R_{Y,D} (\pi_{Y,D}) \leq 1 \tag{4}
\]

\[
\sum_{j} \rho_{j,i} \leq \sum_{\ell} v_{\ell,i} = 1, \quad j \in \{A, D\}, \quad \ell \in \{A, D, S\}
\]

where \(R_{X,A} (\pi_{X,A})\) and \(R_{Y,D} (\pi_{Y,D})\) is the vote mobilizing impact of parties \(A\) and \(D\) on their base, given their respective policy platforms, \(\pi_{X,A}\), \(\pi_{Y,D}\), while \(R_{Y,A} (\pi_{X,A})\), \(R_{X,D} (\pi_{Y,D})\), are the vote mobilizing impacts for voters that constitute the base for the rival party. In a limiting case where the electoral bases are highly energized, either \(R_{X,A} (\pi_{X,A}) \to 1 \Rightarrow R_{X,D} (\pi_{Y,D}) \to 0\), or \(R_{Y,D} (\pi_{Y,D}) \to 1 \Rightarrow R_{Y,A} (\pi_{X,A}) \to 0\), or both.20

Since voters in the feasible electoral space respond to both identitarian and technocratic policy questions, identifying a maximand interior to the \(\rho_{j,i} \in [0,1]\) interval, the general case precludes monotonicity of \(R_{X,A} (\pi_{X,A}), R_{Y,A} (\pi_{X,A}), R_{X,D} (\pi_{Y,D}), R_{Y,D} (\pi_{Y,D})\), which would produce corner solutions to party political platforms, whereby policy focusses only on identity questions, \(X\), or technocratic policy delivery \(Y\). For the sake of simplicity, we exclude the possibility that parties can choose policy platforms that suppress the base of the rival party (\(R_{Y,A} (\pi_{X,A}) < 0, R_{X,D} (\pi_{Y,D}) < 0\)).

### 2.4 Optimal policy platforms

In an electoral system of proportional representation, the decision problem for the two parties is then:

\[
\text{Party } A : \operatorname{arg\,max}_{\pi_{X,A}} \left( \sum_{i} \rho_{A,i} \right) \quad \text{Party } D : \operatorname{arg\,max}_{\pi_{Y,D}} \left( \sum_{i} \rho_{D,i} \right) \tag{5}
\]

---

20 \(R_{X,A} (\pi_{X,A}) < 1, R_{Y,D} (\pi_{Y,D}) < 1\), would occur with turn-out below 100%. 
with the optimal policy platform determined by the solution to the FOCs:

\[
\begin{align*}
\text{Party } A : & \quad \frac{\partial R_{X,A}(\pi_{X,A})}{\partial \pi_{X,A}} \sum_i v_{A,i} + \frac{\partial R_{Y,A}(\pi_{X,A})}{\partial \pi_{X,A}} \sum_i v_{D,i} = 0 \tag{6} \\
\text{Party } D : & \quad \frac{\partial R_{X,D}(\pi_{Y,D})}{\partial \pi_{Y,D}} \sum_i v_{A,i} + \frac{\partial R_{Y,D}(\pi_{Y,D})}{\partial \pi_{Y,D}} \sum_i v_{D,i} = 0 \tag{7}
\end{align*}
\]

Since:

\[
\begin{align*}
u_{D,i} = & \quad \frac{\rho_{A,i} - R_{X,A}(\pi_{X,A}) v_{A,i}}{R_{Y,A}(\pi_{X,A})} = \frac{\rho_{D,i} - R_{X,D}(\pi_{Y,D}) v_{A,i}}{R_{Y,D}(\pi_{Y,D})} \\
u_{A,i} = & \quad \frac{\rho_{A,i} - R_{Y,A}(\pi_{X,A}) v_{D,i}}{R_{X,A}(\pi_{X,A})} = \frac{\rho_{D,i} - R_{Y,D}(\pi_{Y,D}) v_{D,i}}{R_{X,D}(\pi_{Y,D})}
\end{align*}
\]

substitution renders clear that the optimizing behaviour of the two parties is strategic, with choice of the optimal policy platform conditional on the optimizing choice of the rival party. Our interest does not lie in the characteristics of this optimizing choice, but we note that the Nash equilibrium is specified by the simultaneous solution for pure or mixed strategy \(\pi^*_A, \pi^*_D\), by A and D. Henceforth we take these solutions as given.

Critically, since the electoral system is proportional representation, the optimizing pure or mixed strategy \(\pi^*_A, \pi^*_D\), is determined nationally, not with reference to individual electoral districts.

Finally, note that in general the implication of (6) and (7) is not that the two parties are forced to choose the same optimal policy platform - analogously to the median voter theorem. This would be the case if and only if \(\frac{\partial R_{X,A}(\pi_{X,A})}{\partial \pi_{X,A}} = \frac{\partial R_{X,D}(\pi_{Y,D})}{\partial \pi_{Y,D}}\), \(\frac{\partial R_{Y,A}(\pi_{X,A})}{\partial \pi_{X,A}} = \frac{\partial R_{Y,D}(\pi_{Y,D})}{\partial \pi_{Y,D}}\). In general there is no reason to make this assumption. The history of political parties alone creates hysteresis effects, that render the response of different electoral bases to political party platforms distinct.\(^{21}\)

### 2.5 Nonlinear voting outcomes in electoral district characteristics

The core point of this simple voting model for our purposes is that in a context where voters are concerned about more than one issue, in general voting outcomes will be nonlinearly associated with the underlying

\(^{21}\)In our empirical application, for instance, the history of the ANC in South Africa in representing the Black interests, makes it effectively impossible for any other party to have the same impact on Black voters, no matter what its policy platform, at least in the short to medium run.
characteristics of electoral districts.

Given (3), we have:

$$
\frac{\partial \rho_{A,i}}{\partial X} = R_{X,A}(\pi_{X,A}) \frac{\partial u_{A,i}}{\partial X} + R_{Y,A}(\pi_{X,A}) \frac{\partial u_{D,i}}{\partial X}
$$

$$
= R_{X,A}(\pi_{X,A}) \frac{\partial V_A(X,Y)}{\partial X} - \frac{\partial V_A(X,Y)}{\partial X} \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2} + R_{Y,A}(\pi_{X,A}) \frac{\partial V_A(X,Y)}{\partial X} - \frac{\partial V_A(X,Y)}{\partial X} \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2}
$$

$$
= (R_{X,A}(\pi_{X,A}) - R_{Y,A}(\pi_{X,A})) \left( \frac{\partial V_A(X,Y)}{\partial X} - \frac{\partial V_A(X,Y)}{\partial X} \right) \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2}
$$

(8)

and:

$$
\frac{\partial \rho_{A,i}}{\partial Y} = R_{X,A}(\pi_{X,A}) \frac{\partial u_{A,i}}{\partial Y} + R_{Y,A}(\pi_{X,A}) \frac{\partial u_{D,i}}{\partial Y}
$$

$$
= (R_{X,A}(\pi_{X,A}) - R_{Y,A}(\pi_{X,A})) \left( \frac{\partial V_A(X,Y)}{\partial Y} - \frac{\partial V_A(X,Y)}{\partial Y} \right) \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2}
$$

(9)

Symmetrically, from (4):

$$
\frac{\partial \rho_{D,i}}{\partial X} = R_{X,D}(\pi_{Y,D}) \frac{\partial u_{A,i}}{\partial X} + R_{Y,D}(\pi_{Y,D}) \frac{\partial u_{D,i}}{\partial X}
$$

$$
= (R_{X,D}(\pi_{Y,D}) - R_{Y,D}(\pi_{Y,D})) \left( \frac{\partial V_A(X,Y)}{\partial X} - \frac{\partial V_A(X,Y)}{\partial X} \right) \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2}
$$

(10)

and:

$$
\frac{\partial \rho_{D,i}}{\partial Y} = R_{X,D}(\pi_{Y,D}) \frac{\partial u_{A,i}}{\partial Y} + R_{Y,D}(\pi_{Y,D}) \frac{\partial u_{D,i}}{\partial Y}
$$

$$
= (R_{X,D}(\pi_{Y,D}) - R_{Y,D}(\pi_{Y,D})) \left( \frac{\partial V_A(X,Y)}{\partial Y} - \frac{\partial V_A(X,Y)}{\partial Y} \right) \frac{V_A(X,Y) + V_D(X,Y))^2}{(V_A(X,Y) + V_D(X,Y))^2}
$$

(11)

From this it follows immediately that for the general case, in which $V_A(X,Y)$, $V_D(X,Y)$, are non-linear, it follows that both $\frac{\partial V_A(X,Y)}{\partial X} - \frac{\partial V_A(X,Y)}{\partial X}$, and $\frac{\partial V_A(X,Y)}{\partial Y} - \frac{\partial V_A(X,Y)}{\partial Y}$, would generate interactions between the $X$ and $Y$ dimensions of voter interests, thereby generating nonlinear association between $X$, $Y$, and the final vote outcome, as captured by the $\frac{\partial \rho_{A,i}}{\partial X}$, $\frac{\partial \rho_{A,i}}{\partial Y}$, $\frac{\partial \rho_{D,i}}{\partial X}$, $\frac{\partial \rho_{D,i}}{\partial Y}$. 

10
3 South Africa: conventional explanations of political election outcomes no longer serve

South Africa’s political transition in 1994 took place in the most advanced economy of Africa, and was led by one of the oldest African liberation movements. The transition took place by means of a peaceful transfer of power from a small elite to the general population defined principally along racial lines (White to Black), under an economic policy framework adopted by the liberation movement political party that did not engage in extreme redistributive interventions despite facing the world’s highest level of income inequality. As a result it was praised by many commentators for its sustainability and long-term orientation. The policy framework did, however, explicitly target poverty reduction and service delivery to the historically disadvantaged (Black) majority. Strikingly, the municipal elections of 2016 have given an indication that the stranglehold that the party of political liberation, the African National Congress (ANC), had on political power, is now under threat, much sooner than was the case in most of the rest of the African continent’s former colonies. Remarkably, this challenge is emerging both from the "right," by a party historically associated with the rich White elite, the Democratic Alliance (DA), and by a more radical populist "left" party claiming to represent the most marginalized members of South African society, the Economic Freedom Fighters (EFF).

Which raises the question of why the decline of the political fortunes of the ANC could have been so swift, and why a challenge from the left and the right could have gained such credibility?

Conventional explanations of voting behaviour in South Africa do not help to account for the shift in party support. Generally, South African voting allegiances are attributed to markers of race and income, (as well as the strong covariates of income given by employment and education status). For good reason, since race and income strongly correlate with party support. In Table 1 we show that the proportion of the

---

22 The ANC, founded in 1912, long forced to represent its political platform through extra-parliamentary mechanisms due to the discriminatory policies of Apartheid, was explicitly created to represent the interests of the Black majority of the population, as well as Coloured and Asian minority groups. As such, identitarian questions are defined into the political purpose and platform of the party.

23 The DA has historically been understood as the party of the affluent, suburban white minority. It presented itself as a party with a liberal political agenda in contrast to "conservative" white parties that defended racial political exclusion. As such, it was held to stand for liberalizing policy aimed at maximizing economic welfare, as distinct from the identitarian politics of "conservative" white parties.

24 On the changing voting patterns in South Africa see Schmitz (2017).
Table 1: Proportion of the Ward Population in Each Racial Category Across Specified Party Share of the Vote

<table>
<thead>
<tr>
<th></th>
<th>ANC Proportion of the Vote</th>
<th>DA Proportion of the Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 - 30 %</td>
<td>30 - 50 %</td>
</tr>
<tr>
<td>Black</td>
<td>31.76</td>
<td>62.15</td>
</tr>
<tr>
<td>White</td>
<td>34.29</td>
<td>12.30</td>
</tr>
<tr>
<td>Coloured</td>
<td>30.52</td>
<td>22.98</td>
</tr>
<tr>
<td>Asian</td>
<td>3.80</td>
<td>3.23</td>
</tr>
</tbody>
</table>

Columns sum to 100

Table 2: Mean Income in ZAR in Wards by Party Share of the Vote

<table>
<thead>
<tr>
<th></th>
<th>Party Proportion of the Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 - 30 %</td>
</tr>
<tr>
<td>ANC</td>
<td>78,190</td>
</tr>
<tr>
<td>DA</td>
<td>16,831</td>
</tr>
</tbody>
</table>

vote gained by the two main parties (ANC, DA) in the 2009 South African national election across all 4,276 electoral wards, varies systematically and strongly with the proportion of the ward population constituted by racial groups exactly as conventional wisdom predicts. The proportion of the vote gained by the ANC rises in the proportion of ward populations that is Black, the proportion of the vote gained by the DA rises in the proportion of the ward population that is White (and to a lesser degree Coloured and Asian). The standard income correlates of party support are also easily verified for the 2009 national election. Table 2 shows that with a rising proportion of the vote gained by the ANC in a ward, mean income of the ward falls, while for the DA the reverse is true. The standard view of the ANC as the representative of the poor, Black population of South Africa, and the DA as the representative of the rich, White population, appears justified.

But this leaves a puzzle. Since neither income composition of electoral wards, nor the racial composition of the wards has been subject to dynamic change over the 2009-16 period, what would account for the dramatic shift in political party support?

The most immediate response is to question the correlation between race and income on the one hand, and voting outcomes on the other on the grounds of the ecological fallacy. Aggregate voting outcomes hide

---

25 For ease of reference, throughout we employ the official South African designation of racial groups: Indian or Asian (we shorten to Asian); black, coloured (loosely hence imprecisely this maps into mixed race), white.

26 Throughout we specify income values in South African rand (ZAR) units. In 2009 the average exchange rate was ZAR8.5 : US$1, per capita GDP US$5,831 in current US$ terms, or US$11,302 in purchasing power parity terms.

27 Results for education and employment status are symmetrical: ANC support is skewed toward wards with lower educational outcomes, and formal labor force participation, while DA support is skewed toward better educational outcomes and stronger formal labor market participation. Results available from the authors on request.
heterogenous behaviour across the racial and income categories we employ. The concern is present wherever individual level behaviour is inferred from aggregate data.28 However, in the South African case it is unlikely that this is the source of the difficulty. Whenever grouped data is treated as being randomly clustered rather than location specific, the ecological fallacy is avoided under the ‘constancy assumption,’ such that context is immaterial.29 In South Africa the homogeneity requirements for wards is satisfied to an unusually high degree, for both race and income.

Decades of Apartheid policies has rendered the characteristics of individuals and households in wards relatively homogenous across demographic, income and educational dimensions. Moreover, while intra-South Africa migration is significant, migration is non-random, and strongly subject to selection effects by education, income, economic and social infrastructure, such that homogeneity of ward identity is not significantly compromised even over the medium term. This is readily verified for South African voting wards.

In Table 3, we report the proportion of all voting wards, for which the four official racial categories constitute the specified proportion of the ward population. Unsurprisingly given South Africa’s demographics, 74% of all wards have more than 80% of the ward population enumerated as Black, and between 87% and 98% of wards have less than 20% of the ward population enumerated as White, Colored or Asian. Moreover, as reported in Table 4 in income terms wards are also strikingly homogeneous. Overwhelmingly, voting wards are poor, with 31% of wards having mean incomes below ZAR9,600, and 61% of wards report mean incomes of ZAR9,601 - ZAR76,800. In fact, these data underreport the degree of homogeneity, since 53% of all wards report a mean income in the ZAR9,601 - ZAR38,400 range. It is always possible that the ecological inference fallacy may hide important surprises. But given the level of homogeneity of South African voting wards, this also seems unlikely to be the case.

The literature on South African voting is not of much help in resolving the puzzle of changing party support. Explanations of voting in South Africa have identified four main categories of explanatory variables. Voting patterns in South Africa have been most frequently linked to racial identity structure, either directly or indirectly as a signal of credibility, competency and trustworthiness, and social differences within

---

Proportion of Wards with Specified Proportion of Population in the Indicated Racial Category:

<table>
<thead>
<tr>
<th>Racial Category</th>
<th>&lt; 20%</th>
<th>20% - 40%</th>
<th>40% - 60%</th>
<th>60% - 80%</th>
<th>&gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>0.09</td>
<td>0.06</td>
<td>0.04</td>
<td>0.06</td>
<td>0.74</td>
</tr>
<tr>
<td>White</td>
<td>0.87</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Colored</td>
<td>0.88</td>
<td>0.10</td>
<td>0.10</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>Asian</td>
<td>0.98</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Rows sum to 1.00

Table 3: Proportion of Wards for which the Specified Racial Category Constitutes the Specified Proportion of the Ward Population

<table>
<thead>
<tr>
<th>Income Range</th>
<th>Poor</th>
<th>Middle</th>
<th>Rich</th>
<th>Wealthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 9,600</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9,601 - 76,800</td>
<td></td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76,801 - 307,200</td>
<td></td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>&gt; 307,200</td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
</tbody>
</table>

Rows sum to 1.00

Table 4: Proportion of Wards with Mean Income in the Specified Range

the electorate.\(^{30}\) Due to hysteresis, persistence of racial (and other) identity markers over time in voting behaviour has also been advanced.\(^{31}\) Table 1 has already confirmed the prima facie importance of race in South African elections.

Since none of the racial categories in South Africa are ethnically homogenous, an alternative plausible explanation of voting behaviour is ethnic identity.\(^{32}\) For the 2009 election, parties that had an identifiable ethnic platform (even under a very liberal interpretation of an ethnicity) would include at most the Inkatha Freedom Party (Zulu), the Vryheids Front (Afrikaans), the United Democratic Movement (Xhosa), and the Independent Democrats (Coloured), which collectively realized a combined share of 7.18% of the national vote - see Figure 1. This suggests that while ethnicity may play some role in South African electoral outcomes, ethnicity as a principal driver of broad electoral patterns is more difficult to sustain. For good reason: mobilization along ethnic identity is not an obviously rational strategy for political parties, since the largest single ethnolinguistic group in South Africa (Zulu speakers) comprises only 22.7% of South Africans (Stats SA, 2014). Building a stable ruling coalition of voters relying on ethnic identity is difficult as a result, especially since using linguistic ties to divide people politically can be unreliable.\(^{33}\)

\(^{30}\)The racial consensus view of SA politics has a long history. See Lodge (1999); Horowitz (1991); Reynolds (1994, 1999); Mattes (1995); Popkin (1995); Johnson and Schlemmer (1996); Eldridge and Seekings (1996); Seekings (1997); Mattes, Taylor and Africa (1999); Giliomee, Myburgh and Schlemmer (2001); Ferree (2004, 2006); Habib and Naidu (2006); Sylvester (2009); Everatt and Mbeki (2011); Habib and Schulz-Herzenberg (2011).

\(^{31}\)See De Kadt (2016).

\(^{32}\)See Horowitz, (1991); Johnson and Schlemmer (1996); Palmberg (1999); Bekker, Dodds and Khosa (2001); Giliomee, Myburgh and Schlemmer (2001); Norris and Mattes (2003); Ferree (2004); McLaughlin (2007). Where ethnicity does gain salience, this might also open the door for a political role for traditional leaders (De Kadt and Larreguy, 2016).

\(^{33}\)See Piombo (2005); Posner (2004); Norris (2003).
The literature has also attributed voting behaviour to income\textsuperscript{34} - which Table 2 has already shown to be a plausible correlate of voting behaviour. By way of extension, income status also maps into the nature of the association of households with the formal labour market, union membership, or unemployment status,\textsuperscript{35} and education,\textsuperscript{36} both of which can further define the nature of voter interests.

Finally, a newly emergent literature has pointed to the potential importance of service delivery, welfare policies and inefficiency in service delivery as particularly important in South Africa, given the context of high inequality, poverty, and unemployment.\textsuperscript{37} Accordingly the impact of service delivery on voting outcomes has begun to be researched for South Africa.\textsuperscript{38}

Yet these standard explanations of voting outcomes in South Africa are not plausible explanatory candidates of the change in electoral results in South Africa. For all four categories of variables, service delivery, the racial, ethnic, and income status composition of both the population as a whole, and of wards, there is little evidence to suggest that any of these categories has been subject to strong dynamics. Racial and ethnic identity, income status thus are unlikely to be able to account for the declining support for the ANC.\textsuperscript{39}

How then are we to explain the changing patterns of voting behaviour in South African elections? We explore the possibility suggested by our theoretical framework: that at the individual ward level we should anticipate a nonlinear variation in voting outcomes in the dimensions of interest to voters. Empirically we do so by testing whether interactions between race and income prove to be important correlates of South African voting outcomes at the ward level.\textsuperscript{40} If such nonlinearities are confirmed and verified to be of substantively significant magnitude, small rather than large changes in ward composition along the conventional dimensions held to explain South African voting behaviour might account for at least some, potentially for a significant proportion of the changes in South African voting behaviour.

\textsuperscript{34}See Seekings, (1997); Norris (2003); Seekings and Nattrass (2005); Habib and Naidu (2006).
\textsuperscript{35}See Idasa (1998); Mattes, Taylor and Africa (1999); Giliomee, Myburgh, and Schlemmer (2001); Nattrass and Seekings (2001); Norris (2003); Ferree (2004); Seekings and Nattrass (2005); Schulz-Herzenberg (2006).
\textsuperscript{36}See Mattes (1995); Norris (2003); Norris and Mattes (2003); Everett and Mbeki (2011).
\textsuperscript{37}See Stevens (1993); Eldridge and Seekings (1996); Chandra (2004); Habib and Schulz-Herzenberg (2011).
\textsuperscript{38}See Kroth et al (2016), De Kadt and Lieberman (2017) for first attempts.
\textsuperscript{39}Ideally, to exclude change across the four standard explanations of South African voting behaviour, would require a difference-in-difference analysis, to identify changing demographic and other environmental conditions in South African voting wards that correlate with changing voting behavior. This route is not yet available, since no South African census data is available after the 2011 census already being employed for the present study.
\textsuperscript{40}The literature has begun to explore this possibility elsewhere also, for instance in suggesting the importance of race in South African voting, but conditioned on the extent to which voters are segregated - suggesting an underlying information asymmetry explanation for the salience of race. See Sands and De Kadt (2016).
This study merged two distinct data sets: Census Data on the demographics of the population, and Independent Electoral Commission (IEC) data on voting behaviour at ward level.\textsuperscript{41}

Only the most recent (2011) South African Census was used, given the close temporal proximity to the election date. The national statistics agency of South Africa (Stats SA) and the Municipal Demarcation Board (MDB) provided the data.

Voting data is for the 2009 National Election, obtained from the Independent Electoral Commission (IEC). National elections have the advantage of better turnout than municipal and local elections. Moreover in national elections every ballot has the same political parties listed on it, which provides consistency across voting districts (in the municipal elections, party selection varies between regions and wards). On average, 82\% of the people living in a ward are registered to vote and the average voter turnout among the registered is 73\%, higher than the international average registered voter turnout for presidential elections, 66.33\% (IDEA, 2014).

Given available data, non-disclosure and confidentiality policies over Census and Voting Data in South

\textsuperscript{41}The data merge is detailed in Appendix 1.
Table 5: Seven Largest Political Parties in the 2009 National South African Elections

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANC</td>
<td>African National Congress</td>
</tr>
<tr>
<td>DA</td>
<td>Democratic Alliance</td>
</tr>
<tr>
<td>VP</td>
<td>Freedom Front Plus</td>
</tr>
<tr>
<td>IFP</td>
<td>Inkatha Freedom Party</td>
</tr>
<tr>
<td>UDM</td>
<td>United Democratic Movement</td>
</tr>
<tr>
<td>ID</td>
<td>Independent Democrats</td>
</tr>
<tr>
<td>COPE</td>
<td>Congress of the People</td>
</tr>
</tbody>
</table>

Africa, the ward level of disaggregation is the lowest level of aggregation possible without infringing on individuals’ privacy.

One concern with the data is that the census data applies to 2011, the voting data to 2009. The temporal mismatch raises the possibility that dynamics in preferences, demographics and voting patterns render statistical associations subject to error. However, since the only alternative census is that of 2001, the 2011 census minimizes this type of error in the South African context. Moreover, a two year mismatch likely minimizes the importance of any dynamics to the validity of our inferences.

The data set provides observations of the proportion of votes received by all political parties and demographic characteristics for 4276 wards. The voting data provides the proportion of the vote attained by the officially registered political parties, and from the census data we obtain 75 demographic measures. We consider only voting proportions for the largest seven political parties as listed in Table 5—our discussion prioritizes the largest two among them, the ANC and DA. Independent variables used include categories on race, gender, age, education, employment, wealth and income (earnings). A detailed list of the variables can be found in Table 6. Education variables relate to individuals who have completed some or all of the described level of education, except for the higher education variable which only includes individuals who have completed a tertiary education such as a degree or a diploma. Income includes all sources of income such as employment income, social grants, unemployment benefits, remittances, rentals and income from investments, sales or services.

We measure voting behaviour of each ward by vote proportions (rather than absolute number of votes),

---

42 In total of 26 political parties were registered for the national election; list available from authors on request.
43 A perennial concern when using census data is that misrepresentation of earnings and occupations generates significant measurement error. For this reason, we supplemented the census household earnings data by additional proxies of wealth status, given by the size of the household dwelling (number of rooms). Our results are generally consistent across the alternative measures of income and wealth. Note that earnings data is available only in range format, not as point estimates.
<table>
<thead>
<tr>
<th>Variable Category</th>
<th>Dimension Measured</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Black</td>
<td>Black</td>
</tr>
<tr>
<td></td>
<td>Coloured</td>
<td>Coloured</td>
</tr>
<tr>
<td></td>
<td>Indian or Asian</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>White</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td>Age</td>
<td>Youth (aged 0 - 19 years)</td>
<td>Youth</td>
</tr>
<tr>
<td></td>
<td>Young adults (aged 20 - 34 years)</td>
<td>Young_ad</td>
</tr>
<tr>
<td></td>
<td>Middle-aged adults (aged 35 - 49 years)</td>
<td>Middle_aged_ad</td>
</tr>
<tr>
<td></td>
<td>Senior adults (aged 50 - 64 years)</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td></td>
<td>Retired adults (aged 65 years and older)</td>
<td>Retired_ad</td>
</tr>
<tr>
<td>Education</td>
<td>No education</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td></td>
<td>Primary education</td>
<td>Primary_ed</td>
</tr>
<tr>
<td></td>
<td>Secondary education</td>
<td>Secondary_ed</td>
</tr>
<tr>
<td></td>
<td>Higher education</td>
<td>Higher_ed</td>
</tr>
<tr>
<td>Employment</td>
<td>Employed</td>
<td>Employed</td>
</tr>
<tr>
<td></td>
<td>Unemployed</td>
<td>Unemployed</td>
</tr>
<tr>
<td></td>
<td>Discouraged workers</td>
<td>Discouraged</td>
</tr>
<tr>
<td></td>
<td>Not economically active workers</td>
<td>NEA</td>
</tr>
<tr>
<td>People per household</td>
<td>Singular household (1 person)</td>
<td>Singular_hh</td>
</tr>
<tr>
<td></td>
<td>Small household (2 - 3 people)</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td></td>
<td>Medium household (4 - 5 people)</td>
<td>Medium_hh</td>
</tr>
<tr>
<td></td>
<td>Large household (6 or more people)</td>
<td>Large_hh</td>
</tr>
<tr>
<td>Rooms per household</td>
<td>Small house (1 - 4 rooms)</td>
<td>Small_house</td>
</tr>
<tr>
<td></td>
<td>Medium house (5 - 8 rooms)</td>
<td>Medium_house</td>
</tr>
<tr>
<td></td>
<td>Large house (9 or more rooms)</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td>Annual income</td>
<td>Poorest (R 0)</td>
<td>R0</td>
</tr>
<tr>
<td></td>
<td>Very poor (R 1 - R 9,600)</td>
<td>R1_R9600</td>
</tr>
<tr>
<td></td>
<td>Poor (R 9,601 - R 38,400)</td>
<td>R9601_R38400</td>
</tr>
<tr>
<td></td>
<td>Low income (R38,401 - R 76,800)</td>
<td>Ref. Cat.</td>
</tr>
<tr>
<td></td>
<td>Middle income (R 76,801 - R 153,600)</td>
<td>R76801_R151600</td>
</tr>
<tr>
<td></td>
<td>Rich (R 153,601 - R 307,200)</td>
<td>R151601_R307200</td>
</tr>
<tr>
<td></td>
<td>Richest (R 307,200 or more)</td>
<td>R307201_plus</td>
</tr>
</tbody>
</table>

Ref. Cat denotes the Reference Category where appropriate.

Table 6: Independent Variables Employed in Study

to provide comparability across wards which have different sized populations.\(^{44}\)

5 Results

The standard explanation of South African politics is that poor, uneducated Black folk, marginalized in the labour force or in blue-collar jobs, vote for the ANC. Rich, educated, professional Whites vote for the DA.

\(^{44}\)Since we employ a linear specification linking the vote proportion to demographics, there is a potential violation of the upper (100%) and lower (0%) bound values feasible for the vote proportion. However, given our interpretation of the impact of independent variables in small intervals surrounding observed vote proportions (which are distributed away from the upper and lower bounds), this is not a critical consideration in our application. An alternative would be to employ a logistic transform of the dependent variable.
All other parties are ethnic fillers. In addition, our theory predicts the possibility of nonlinearities arising from interactions between fundamental drivers of voting behaviour such as race and income.

To explore this we estimate:

\[
V_{p,r} = \beta_0 + \sum_j \alpha_j R_{j,r} + \gamma G_r + \sum_k \delta_k A_{k,r} + \sum_{\ell} \beta_\ell E_{\ell,r} + \sum_m \eta_m L_{m,r} \\
+ \sum_n \theta_n H_{n,r} + \sum_q \kappa_q S_{q,r} + \sum_u \lambda_u Y_{u,r} + \sum_v \mu_v W_{v,r} \\
+ \sum_j \nu_j R_{j,r} Y_{poorest,r} + \sum_j \xi_j R_{j,r} Y_{richest,r} + \varepsilon_r
\]  

(12)

where \( V_p \) represents the proportion of votes for a given party, \( p; R, G, A, E, L, H, S, Y, W \), represent the vectors of explanatory variables for ward \( r \) as defined below, and \( \beta_0 \) and \( \varepsilon_r \) represent the constant and the error term, respectively. As proportions were used, the variable with the smallest proportion was employed as the excluded category.\(^{45}\) Controls employed included race (\( R \)), income (\( Y \)), education (\( E \)), employment status (\( L \)), gender (\( G \)), number of people recorded in households (\( H \)), rooms per household (proxy for assets or wealth: \( S \)), the age structure of the ward population (\( A \)),\(^{46}\) as defined in Table 6, as well as geographical region as defined by the nine provinces (\( W \)).\(^{47,48}\)

Finally, we include two sets of interaction terms between race (\( R \)), and the proportion of the ward population that falls into the poorest income category (\( Y_{poorest,r} \)), and the proportion of the ward population that falls into the richest income category (\( Y_{richest,r} \)). This allows for the examination of the impact on the party vote proportion (\( V_p \)) as the proportion of the ward population that falls into the poorest income category varies for any specific race group \( j \), \( (\alpha_j + \nu_j Y_{poorest,r}) \), symmetrically for the richest income group, \( (\alpha_j + \xi_j Y_{richest,r}) \), or the impact on the party vote proportion (\( V_p \)) as both the proportion of the ward population belonging to the two income categories, and the proportion of the ward population belonging to  

\(^{45}\)The constant will therefore identify the mean impact of excluded categories: Asian households; females; senior adults; no education; low income; small households; with large houses.

\(^{46}\)Given the relatively high fertility rate in South Africa, and the significance of the 1994 political transition from white minority rule, which divides the electorate into "not-frees" and "frees" by virtue of either having, or not having personal experience of Apartheid, age structure may be of particular importance in the South African case (though in the 2009 election its impact should be limited, given that only 15 years had elapsed since the transition). See Schulz-Herzenberg (2006, 2013); Everatt and Mbeki (2011).

\(^{47}\)Geographic region is significant in South Africa since a number of parties explicitly claim regional identity: the IFP gains more than 90% of its votes from KwaZulu Natal, while the UDM gains more than 60% of its votes from the Eastern Cape.

\(^{48}\)South Africans are also subject to socioeconomic issues such as high crime rates - see Ferree (2004); Seekings and Nattrass (2005); Idasa (1998); Mattes, Taylor and Africa (1999); Southall (2013) - making related policies potentially relevant. Unfortunately we do not have crime statistics at the electoral ward level, and cannot control for this.
a specific race category \( j \) vary, by \((\alpha_j + \nu_j Y_{poorest,r}) R_{j,r} \), \((\alpha_j + \xi_j Y_{richest,r}) R_{j,r} \), respectively.

Estimation is by OLS, separately for each of the seven major parties. Statistically our examination of patterns of voting behaviour at the ward level in terms of the demographic characteristics of the wards is correlational rather than causal. Baseline estimation is both under a zero restriction on the interaction terms, \( \nu_j = 0 = \xi_j \), as reported in Table 7, and the removal of these restrictions, as reported in Table 8. In what follows, we discuss estimation results by independent variable category.\(^{49}\)

### Table 7: Baseline Regressions

<table>
<thead>
<tr>
<th>Race</th>
<th>Constant</th>
<th>R307201_plus</th>
<th>R307201_plus_ln</th>
<th>Middle_aged_ad</th>
<th>Retired_ad</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
<th>R307201_plus Ln</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.016**</td>
<td>-0.165***</td>
<td>-0.005</td>
<td>-0.207***</td>
<td>0.12</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>Black</td>
<td>0.016**</td>
<td>0.165***</td>
<td>-0.005</td>
<td>-0.207***</td>
<td>0.12</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>Coloured</td>
<td>0.105***</td>
<td>-0.129***</td>
<td>-0.010***</td>
<td>0.181***</td>
<td>0.001</td>
<td>0.112***</td>
<td>0.202***</td>
<td>0.001</td>
<td>0.112***</td>
<td>0.202***</td>
<td>0.001</td>
<td>0.112***</td>
<td>0.202***</td>
</tr>
</tbody>
</table>

Figures in round parentheses denote standard errors.

\(^{49}\)We also considered at length the impact of possible outliers on final estimation results. Out of a total of 4275 wards, outliers were found in only 38 wards. 9 of these wards contained outliers across several demographic variables and several parties. This could be due to the enumerators’ capturing techniques in these wards. As a proportion of the total number of wards, these outliers proved to be immaterial, and did not influence the results when controlled for.
### Table 8: Regressions with Interaction Terms

<table>
<thead>
<tr>
<th></th>
<th>Adj. R²</th>
<th>Df</th>
<th>Adj. Df</th>
<th>Coeff.</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted Observations</td>
<td>0.093</td>
<td>941</td>
<td>0.673</td>
<td>0.668</td>
<td>0.222</td>
<td>0.535</td>
<td>0.424</td>
</tr>
</tbody>
</table>

Figures in round parentheses denote standard errors.

*p < 0.05, **p < 0.01, ***p < 0.001
the significant homogeneity of South African wards moderates the ecological fallacy concern, in fact renders it unlikely to be an important limitation in the South African case. In any event, there is no alternative in South Africa. While Afrobarometer conducts an individual survey (2–3000 individuals), it records only opinion, and cannot verify voting behaviour.\(^{50}\) Ward level data is the only recourse to offer insight into South African voting patterns.

5.1 Accounting for race

Results confirm the obvious: race matters in South Africa. For all political parties the race variables are statistically significant.

As shown in Table 7, for the ANC, the voting proportions increase as, relative to the proportion of Asians, the proportion of Black and Coloured voters increase, and decrease as the proportion of White people living in a ward increases, with a 1% increase in each group translating to a 0.647% increase, 0.105% increase and 0.116% decrease in the share of the vote received by the ANC. The DA shows the anticipated mirror image, with a 1% increase in the proportion of Black, Coloured and White population proportion, relative to the proportion of Asians, generating a 0.456% decrease, 0.129% decrease and a 0.373% increase in the DA voting proportion, respectively. The result is entirely unsurprising, and conforms to the standard perception of voting splitting along racial lines for the ANC and the DA, as noted in the section 3.\(^{51}\)

But there are anomalies, evident from Figures 2 and 3. Figure 2 illustrates the relationship between the proportion of votes received by the ANC and the interaction between the proportion of the ward population that is White and the proportion of households that is wealthy (earning an annual income of more than R307,200), and the interaction between the proportion of the ward population that is Black and the proportion of households that falls into the lowest income category (no income), respectively. Figure 3 considers the relationship between the proportion of DA votes and the interaction between the proportion of the ward population that is White and the proportion of households that is wealthy (earning an annual income of more than R307,200). From the illustration, for the preponderance of wards, as the proportion of rich Whites

\(^{50}\)Note also that individual level data poses its own issues (Park, 2008).

\(^{51}\)This is further corroborated by the findings for the other "major" parties: the share of the VP increases in the proportion of the electorate that is White, the ID from the proportion of the electorate that is Coloured, the UDM from the proportion of voters that are Black, and COPE gains support from both Black and Coloured electorate proportions.
in the population increases, or the proportion of poor Blacks decreases, the share of the vote for the ANC declines. By contrast, again for a preponderance of wards a rising proportion of rich Whites rises in the population increases the share of the vote for the DA. This pattern is consistent with prior expectations concerning the impact of race and income on ANC and DA vote proportions.

Despite this preponderant pattern, a significant number of wards do not conform.

Wards corresponding to area D1 in Figure 2 (550 wards), indicate the possibility of a low proportion of the vote for the ANC, despite the fact that the proportion of rich Whites in the population in the ward is low - and recall that the full range of variables across race, gender, education, income, employment status, age and household size have been controlled for. In the wards identified by area D2a in Figure 2 (171 wards), despite a low proportion of poor Blacks in the ward population, the ANC nevertheless achieves a high proportion of the total vote, while in wards identified by area D2b (142 wards) the ANC realizes a low proportion of the vote despite the ward recording a high proportion of the population as poor and Black.

For the DA there are also anomalies. Wards corresponding to area D3 of Figure 3 (129 wards), the DA obtains a high proportion of the vote, despite the fact that the proportion of the ward population that is rich and White is moderate in size at best.

Wards that fall in the anomalous areas are all differentiated from sample averages:

- Area D1 wards have relatively larger proportions of Coloureds, relatively lower proportions of Blacks, and a higher proportion of households with more than 6 members, and that are not economically active. Moreover 73% of these wards are located in either KwaZulu-Natal or in the Western Cape.

- Area D2a wards have a higher proportion of Coloureds, those who are formally employed, and earning between R9,600 and R38,400 annually, and a lower proportion of Blacks and those who are not economically active. Geography again matters, with 86% of the wards located in either the Northern Cape or the Western Cape.

- Area D2b wards report higher proportions of Blacks, individuals younger than 20, individuals without education, large households of more than 6 members, earn between R0 and R9,600 per annum, are not economically active, and are discouraged from the labour force. Of area D2b wards, 84% are in
Figure 2: Anomalies in ANC Voting Patterns.

KwaZulu-Natal.

- Area D3 wards have lower proportions of Blacks, single person households, incomes between R0 and R9600 per year, and larger proportions of Coloureds, of formally employed, and household sizes of 4-5 persons. Regionally, 65% of the wards are located in the Western Cape.

The anomalous wards are thus differentiated by geography, family structure, income and employment status (particularly the not economically active and working class), and Coloured population proportions.

5.2 The interaction between race and income

While for some of the anomalous associations the proportion of the wards may seem relatively small, across the four exceptional associations a total of 992 wards are identified - 23% of the sample of wards (not population). To account for the impact of these wards, we reestimate specification of (12), by removing the zero restriction on the interaction terms between each race variable and the two extreme income brackets: the poorest households which earn no income; and the richest households with an annual income of more than R307,200. We also control for provincial dummies, to account for the importance of the KwaZulu-Natal, Western Cape, and Northern Cape exceptionalisms noted for the D1, D2a, D2b, D3 cases.

Results are reported in Table 8. Goodness-of-fit improves dramatically for the ANC (0.77→0.90), though less markedly for the DA (0.90→0.94). The impact of controlling for the interaction terms and geographical dummies is to modulate the impact of racial identity, particularly for the ANC. While the ANC’s proportion
of the vote increases with the proportion of Blacks in a ward, the response is weaker than in the absence of controlling for race and income interactions (0.65 → 0.32), the proportion of the population that is Coloured is rendered insignificant, though the White population proportion impact is effectively unchanged (0.12 → 0.13).

For the DA, voting proportions still decrease in the proportion of the population that is Coloured and Black, but more aggressively so (−0.46 → −1.01, −0.13 → −1.25), but no longer respond statistically significantly to changes in the proportion of Whites (0.37 → 0).

To demonstrate the impact of controlling for the interaction between income and race and geography, consider the voting outcomes for the wards reported in Table 9. Note that the baseline model, which controls for race, income, education, age structure of the population, House and Household size, both underpredicts (Wards 1-4) and overpredicts (Wards 5-9) actual voting outcomes. Controlling for the interaction of race with income significantly improves the fit of the model; controlling for geography and interaction terms improves the fit of the model further still. Noteworthy is the fact that the wards cover both rural and urban cases, and are geographically widely spread, not confined to the Western Cape and KwaZulu-Natal exceptional areas.

At the very least, the common view that the ANC is a "Black" party, and that it is actively shunned
Actual Vote % | Baseline Pred. Vote % | Pred. Vote % with Interactions | Name of Ward | Provincial Location
---|---|---|---|---
1 | 93 | 72 | 81 | Nyandeni | Eastern Cape
2 | 92 | 84 | 88 | Ephraim Mogale | Limpopo
3 | 89 | 86 | 87 | Drakenstein | Western Cape
4 | 84 | 82 | 86 | Drakenstein | Western Cape
5 | 59 | 68 | 62 | Johannesburg | Gauteng
6 | 56 | 67 | 60 | Tshwane | Gauteng
7 | 12 | 19 | 5 | Tlokwe | North West
8 | 14 | 24 | 28 | Drakenstein | Western Cape
9 | 15 | 22 | 23 | Drakenstein | Western Cape

Table 9: Impact of Allowing for Interaction Nonlinearities

by Whites, while the DA is predominantly a "White" party actively avoided by Blacks is thus brought into question. Here we focus specifically on the implications for the ANC and DA that arise from the interactions between the two extreme income categories (→R0 and above R307,200) and population racial proportions. In all figures that follow, the vertical axis represents the change in the proportion of votes received by the party; Race represents the proportion of people in a ward belonging to the specified race; and Income represents the proportion of people in a ward earning the specified income category. The two income categories we consider are the lowest recorded income per annum, termed Poorest, and income above R307,200 per annum, which we termed Richest. Our interest is particularly in the impact on the party vote proportion \( \pi \) as both the proportion of the ward population belonging to the two income categories, and the proportion of the ward population belonging to a specific race category \( j \) vary, by \((\alpha_j + \nu_j Y_{poorest,r}) R_{j,r}, (\alpha_j + \xi_j Y_{richest,r}) R_{j,r}, \) in (12) respectively.

The general insight to emerge for our purposes, is that the interaction between race and income in South Africa, for both the ANC and the DA, implies that relatively small changes in both racial ward population proportions, and the proportion of the ward population that is rich or poor, can trigger relatively strong variation in the proportion of the vote that the party realizes.

For the ANC we find statistical significance for the interaction terms for Blacks, Whites and Coloureds.\(^{52}\) Specifically, we find that the ANC vote proportion declines as the proportion of the population that is Coloured and that is Richest or Poorest increases, moderately for the poor, but dramatically for the rich.\(^{53}\)

---

\(^{52}\) A visual representation of these results is summarized in Figure 7 of Appendix 2.

\(^{53}\) See Figure 7 Panel A.
Further, the proportion of the vote that the ANC receives also declines with a rising proportion of both poor Whites and rich Blacks in the population of a ward, though statistically significantly only for Whites. A rising proportion of the population that is rich and White also lowers the share of the vote attained by the ANC, with a 1% increase (not a 1 percentage point increase) in the rich White population (off any arbitrary base) share generating a 0.2 percentage point decline in the proportion of votes realized by the party. The inference is therefore that the ANC, while in general gathering support from Black voters (the variable for the proportion of the ward population that is Black remains positive), loses support amongst the richest Black voters, and poor and rich Coloureds and Whites, though more dramatically amongst Coloureds than Whites. By contrast, middle income Whites and Coloureds matter less for the ANC voting share.

For the DA we find significance of the interaction terms for Blacks, Whites and Coloureds. The proportion of votes received by the DA rises with the proportion of the ward population that is Black and poor and that is Black and rich - even more surprisingly, more strongly for the poorest than for the richest Blacks. Symmetrically, the DA vote proportion also rises with the poorest and richest Coloured ward population proportion, though now more strongly for the richest than the poorest, and more strongly than for Blacks. A rising proportion of poor Whites in ward populations also raises the proportion of votes the DA obtains, with the size of the impact slightly higher than that for poor Blacks, but of a similar order of magnitude. For the richest Whites, the result is a mirror image of that obtained for the ANC, with a 1% increase (not a 1 percentage point increase) in the rich White population (off any arbitrary base) share generating a 0.3 percentage point increase in the proportion of votes realized by the DA. The inference is that the DA consistently gathers support from the highest income bracket across all races (Black, White, Coloured), most dramatically amongst Coloureds. Surprisingly, DA voter support also appears to increase disproportionately with a rising proportion of the ward population that is in the poorest income bracket (zero income), irrespective of race. While the poor Black population proportion impact on the DA voting proportion is not large in substantive terms, it is nevertheless a startling finding, and runs counter to the

---

54 See Figure 7 Panel B.
55 Note that this implies a non-linear elasticity conditional on the level of the voting share.
56 A visual representation of these results is summarized in Figure 8 of Appendix 2
57 See Figure 8 Panel A.
58 See Figure 8 Panel B.
59 See Figure 8 Panel C.
conventional wisdom which views the support base of the DA as strongly biased away from poor, Black constituencies. It is worth emphasizing that this result is not driven by ward outliers. What is more, the result is not a regional effect since the positive impact of the proportion of poor Blacks on the DA vote proportion is present for every province.

One interpretation is that the ANC has not actively represented the interests of the poorest in South African society, and while at face value it is surprising that the political support would have emerged for the DA, in 2009 it also represented the largest, most organized, and most vocal opposition to the ANC, that was emphasizing its ability to deliver services. It also points to the credibility of a populist party platform identifying the interests of the most disadvantaged Black voters in South Africa - essentially to the "left" of the ANC. This effectively presents an alternative to the possibility that the poor voters that have not been served well by ANC policy, have ceased to be active voters - see Everatt (2016) - though the two explanations are not mutually exclusive.

In the specifications controlling for income and race interaction terms, we also controlled for provincial dummies, as identified by the anomalous wards under Figures 2 and 3. As a result we controlled for the Western Cape, the Northern Cape, and KwaZulu-Natal in the interaction term regressions. Results confirm a positive, statistically significant positive coefficient on the Western Cape for the DA, while the Western Cape and KwaZulu-Natal are negative and statistically significant for the ANC, and the Northern Cape positive.

5.3 Income, employment and education

What of the impact of income, employment and education status under the presence of interaction effects as reported in Table 8?60

The ANC experiences statistically significant decreasing vote proportions both at incomes above and below the R38,401 - R76,800 reference income category. Over the entire R0-R38,400 income range, ANC support declines relative to the R38,401-R76,800 reference category range. It also declines above the reference

---

60 The remainder of our controls in estimation are of less importance, and we comment only briefly. We find that in terms of gender, ANC support is positively linked to male population proportions, the DA to female population proportions. In terms of age, ANC support is biased toward younger voters, and pensioners. ANC support is concentrated in mid-sized households (4-5 members), and declines amongst the largest household (>6 members), while the DA support is concentrated in nuclear family structures (2-3 members) and those living in medium-sized houses.
category range, up to R153,600 income per annum. For the DA, we find statistically significant declining support at the lowest income levels (R0 - R38,400), and the highest income categories (above R153,600). However, while the DA appears to lose votes among the rich, this is offset by the significant gains among rich Black, Coloured and White voters observed from the interaction terms discussed in the preceding subsection of the paper.

The implication is thus that both the ANC and DA support peaks for relatively privileged working class voters (incomes of R38,401 - R76,800), and falls off both below and above this income bracket. This is consistent with our finding from the interaction term analysis that the ANC incompletely represents the interests of the very poor, creating an opening for the emergence of a populist "left" party.61

Associated with income status, we also controlled for the impact of labour market outcomes. For the ANC, voting support increases in more marginal labour force participation, increasing as the proportion of voters that are unemployed, discouraged or not economically active (predominantly populations that are in school or pensioned) rises. The proportion of the population employed is nonlinearly, concavely associated with the ANC voting proportion, implying a maximum ANC support level where 76% of the population is employed. Figure 4 illustrates the plausibility of the nonlinearity. Note that the low ANC vote proportion at low employment level wards (the left downward spike), are characterized by disproportionate numbers of young people, females, poor (R0-R9,600 income), low education, large households, small houses, not economically active and discouraged from the labour force, located in KwaZulu-Natal.

We note that the declining ANC support at low employment levels, is again consistent with the finding of declining ANC support among the most marginalized Black population, as reported under the interaction term findings, and the income findings of Table 8.

The DA effectively returns a mirror image of the ANC results, with falling support as the proportion of the population that is marginalized in the labour force (unemployed, discouraged, not active) rises, and with a convex nonlinear association with respect to employment. The implied minimum level of support occurs where 0.4 of the population is employed. We note, however, that while the nonlinearity finds support for the

---

61 For the remaining "major" parties: the VP gains support among voters below the reference category (R38,401 - R76,800) and loses support for incomes above the reference category; IFP support increases among the poorest and most affluent relative to the reference category, and declines in low income working class income categories (R1-R38,400); the UDM loses support below the reference category, gains it amongst high income voters, while COPE gains at all income levels relative to the reference category. The ID voting support is invariant in income.
The ANC loses support as the proportion of the voting population becomes more educated, and gains support as the proportion of the population has zero education increases. By contrast, DA voting support is rises statistically significantly with prevalence of secondary education, but not otherwise. Results thus suggest that ANC support lies with uneducated voters, while education appears to play only a moderate role in the support for most other parties, with the exception of COPE whose support rises strongly in education.

\footnote{For the remainder of the "major" parties, formal labor market participation, either in the form of employment or unemployment, or being not economically active (in schooling, pensioned), lower IFP support, while discouraged status in the labor force raises IFP support. COPE support responds positively to unemployment, but declines in employment and being discouraged from the labor force. UDM support declines in discouraged labor force status, and rises in employed status and not economically active status. Neither the ID nor the VP is statistically conditional on the employment status of voters.}

\footnote{For the remainder of the parties, the IFP loses voter support with rising primary and secondary levels of schooling. COPE gains support as all forms of education increase, the ID only in secondary education. While UDM support does not appear to respond to education at all, the VP has falling support in secondary, but rising support in higher education.}
Conclusions and evaluation

This paper presents a simple model of voting in a two issue, two party electoral space, that predicts that in general voting outcomes in a proportional representation system at the electoral district level will be subject to nonlinearities. The two issues motivating voters will be subject to interaction, which implies that relatively small changes in electoral district characteristics over time, can result in marked changes in voting outcomes, not because the motivating drivers of voting behaviour have changed, but because of the nonlinearities that result from their interaction.

The empirical test of the theory is on a case which has the requisite structure in its electoral contests. South Africa in its 2009 national elections had two principal political parties contesting the election, and South African politics is also characterized by a two issue contest over race and income.

Conventional wisdom surrounding South African voting behaviour is not in contention in our reported results. Race matters, with Blacks more likely to vote for the African National Congress, Whites more likely to vote for the Democratic Alliance. Rising incomes and greater security in the labour market predict a switch in allegiance from the African National Congress to the Democratic Alliance.

But we also confirm that there are important nonlinearities hidden in these broad patterns. Allowing for an interaction between income and race, reduces the impact of racial identity for the major parties. This is particularly so for the African National Congress, for which voting proportions no longer respond statistically significantly to White and Coloured population proportions. It also eliminates the statistical significance of the White population proportion for the Democratic Alliance. This calls into question the common view that the African National Congress is a "Black" party actively shunned by Whites, while the Democratic Alliance is predominantly a "White" party actively avoided by Blacks.

Amongst the interaction terms, the striking finding is that while the Democratic Alliance consistently gathers support from the highest income bracket across all races (Black, White, Coloured), though most dramatically amongst Coloureds, Democratic Alliance voter support is disproportionately strong amongst poor Black (and other race group) voters. While the poor Black population proportion impact on the Democratic Alliance voting proportion is not large in substantive terms, it nevertheless runs counter to the conventional wisdom which views the support base of the Democratic Alliance as strongly biased away from
poor, Black constituencies. The result is neither driven by ward outliers, nor is it a regional effect since the positive impact of the proportion of poor Blacks on the Democratic Alliance vote proportion is present for every province.

These results are consistent with the findings we report for the impact of income and the nature of labour force participation. Here we find that African National Congress support peaks for relatively privileged working class voters (incomes of R38,401 - R76,800), and falls off both below and above this income bracket. We also find that the proportion of the population employed is nonlinearly, concavely associated with the African National Congress voting proportion, implying a maximum African National Congress support level where 0.76 of the population is employed. Declining African National Congress support at low employment levels, is again consistent with the finding of declining African National Congress support among the most marginalized Black population.

These findings carry broader implications for an understanding of South African political contests. It provides one explanation of why the support of the African National Congress has tailed off since the 2000s, despite relatively unchanging racial, income, education, employment conditions. Importantly, while the aggregate statistics on race, income, education and employment characteristics demonstrates little by way of dynamics for South Africa, there is evidence that within the aggregates there is significant churn - see for instance with respect to employment (Banerjee et al, 2008), and with respect to poverty and inequality (Fedderke et al., 2004). Thus while in aggregate racial proportions, and income category proportions remain relatively stable, the proportion of the Black (and other race group) population that has fallen into the richest and poorest income categories, may have been subject to dynamic change. Our results suggest that even quite modest changes in either income or racial proportions in wards can trigger relatively strong changes in party voting proportions.

A further interpretation of these findings is that the African National Congress has not actively represented the interests of the poorest in South African society, and while at face value it is surprising that the political support would have emerged for the Democratic Alliance, in 2009 it also represented the largest, most organized, and most vocal opposition to the African National Congress. It also points to the credibility of a populistic party platform identifying the interests of the most disadvantaged Black voters in South
Africa - essentially to the "left" of the African National Congress, explaining the relative rise of the radical Economic Freedom Fighters after 2009. The voting outcome of the 2016 municipal elections in South Africa, dramatically lowering the vote proportion gained by the African National Congress, is thus not entirely surprising.

Finally, at the most general level, the results of the paper imply that relatively small changes in electoral district characteristics over time, can result in marked changes in voting outcomes, due to the interaction between the motivating drivers of voting behaviour. Dynamic change in voting outcomes therefore do not necessarily require dramatic changes in the drivers of voting behaviour, but can be due to nonlinearities that result from relatively modest changes in well established drivers of voting interacting. Confirmatory evidence for South Africa provides support.

References


Appendix 1: Data Merges

Merging the two data sets required several processes. The Census data obtained from Stats SA was aggregated up to ward level, while the voting data obtained from the IEC was aggregated up to voting district level. Voting districts are smaller than wards, and the boundaries of voting districts and wards do not always coincide, which means that a voting district can be shared between two wards as illustrated in Figure 5.

The reason for this lies in the determination process of the boundaries for the voting districts and the wards. This is laid out in various South African statutes.

As Apartheid drew to an end in 1994, the laws of the time were repealed and a restructuring of the laws took place. Parliament drafted a new Constitution, the Constitution of the Republic of South Africa, 1996, which laid the foundation of new laws which were to govern the Republic of South Africa. This constitution formalized new provincial borders (s.103) changing the number of provinces from four to nine and most importantly, changed the National and Provincial election process from Constituency-based elections to Proportional Representation Elections.

The former involved subdividing the complete territory of each province into municipalities, which are the local sphere of government (s.151) and the latter meant that each registered voter could cast a vote for the political party of his/her choice, while each political party would choose its own Members of Parliament. Once the votes are tallied nationally, each party obtains the number of seats in Government that is proportional to the number of votes it won – each percent of the national vote obtains 4 members of parliament.

7.1 Determination of Voting Districts

In 1996, the Electoral Commission Act (51 of 1996) established an independent electoral commission, which is responsible for managing and promoting free, fair and orderly elections (ss.3-5). There was, however, no statute in place regarding the formation of voting districts. The first Act that addressed this issue was the Electoral Act of 1998. According to ss.60-61 and 66, the Commission and electoral officer must establish voting districts for the whole territory of the country, taking into consideration the number and distribution of voters; the availability of voting stations; telecommunications facilities; geographical features and municipal and provincial boundaries.
7.2 Determination of Wards

The boundaries of municipalities are determined and implemented by the Municipal Demarcation Board subject to ss. 21, 24-25 in the Municipal Demarcation Act 27 of 1998 and the boundaries of the wards are determined according to s.22 and ss. 2 – 4 in Schedule 1 of the Municipal Structure Act 117 of 1998. The number of wards depends on the number of ward councillors in a municipality (indirectly determined using a formula) as well as the number of registered voters in the area. The Municipal Demarcation Board consults with the Electoral Commission and ensures that each ward has the same number of registered voters with 15% leeway either way, and that communities are not fragmented.

7.3 Merging the Data sets

Each voting district was given a unique identifier and then allocated to the ward in which it lies. Where a voting district was shared across two wards, it was allocated to both wards. In order to determine how the population in the voting district was distributed and in which ward the people resided, GIS technology at the MDB was utilized. Dots representing the population within a voting district were placed on the map of voting districts. Using these dots, a weighting measurement was created for each voting district, such that the percentage of votes within a voting district that lay within the ward would correspond to the percentage of the population residing in both voting district and ward. This is illustrated in Figure 6 below. In Figure 6, 75% of the residents living in the voting district labelled VDc and 100% of the residents living in the voting districts labelled Ada, VDb, VDd and VDe live in the ward labelled Wa. This means that 75% of the votes for each political party in the voting district, VDc would be allocated to ward Wa, and the remaining 25% allocated to ward Wb. Voting districts VDa, VDb, VDd and VDe do not need to be split and hence 100% of the votes for each political party in those voting districts would be allocated to the ward Wa.

Using this technique and summing the amount obtained from the proportions of votes per party in each voting district, the number of votes per party could be aggregated to ward level. This process was applied to all the variables in the voting data set.
Figure 5: Wards and Voting Districts

Figure 6: Population Distribution in Wards and Voting Districts
8 Appendix 2: Interaction Terms
Figure 7: Panel A: Coloured ANC: Poorest (Red-Blue), Richest (Blue-Gray). Panel B: White and Black ANC: Poorest White (Red-Blue), Richest Black (Blue-Gray).