

Essays in Political Economy:
Elections, Public Finance and Service
Delivery in South Africa

Verena Ernestine Kroth *

A thesis submitted to the Department of Government of the London School of Economics and Political Science for the degree of Doctor of Philosophy. London, April 2014.

* London School of Economics and Political Science

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I can confirm that the idea for chapter 2 emerged from my previous study for Master of Public Administration at the London School of Economics and Political Science. Chapter 3 is based on a paper that was written jointly with Joachim Wehner and Valentino Larcinese. I certify that I was responsible for the data collection and spatial analysis. Each of us contributed equally to the data analysis and writing of the paper.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author. I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I can confirm that this thesis was copy edited for conventions of language, spelling, and grammar by Michael Wise.

Abstract

Who gets what, when and how? Each of the three papers in this thesis makes a distinct contribution to answering this question in the context of the political economy of South Africa. The first paper examines how South Africa's public financial management system distributes central government funds to its provinces. Using a unique panel dataset comprising all provinces and three elections over the period 1995-2010, I demonstrate that provinces where the national ruling party has higher vote margins receive higher per capita equitable shares in pre-election years. This result suggests that even in a dominant party framework, electoral competition can function as an incentive to implement political budget cycles. The second paper evaluates how the extension of the franchise affected the delivery of electricity to South African households. The dataset combines nightlight satellite imagery, census data and municipal election results, making it possible to exploit the heterogeneity in the share of newly enfranchised voters across nearly 800 municipalities with a difference-in-differences approach. The analysis demonstrates that enfranchisement has a significant positive effect on household electrification. Moreover, the findings show that political parties have a potential mediating role in accounting for service delivery patterns in new democracies. The third paper addresses the problem of measurement in studying public service delivery by examining a novel methodology for combining census-based data with satellite imagery of the world at night. Using cross-national data and South African census data, the paper provides a roadmap for how to navigate limitations and thus make the most of this technological advance in quantitative social science research.

Acknowledgments

There are many individuals to whom I am very grateful for guidance and support. First and foremost, I would like to thank my supervisor, Joachim Wehner, who inspired me to take on this project and provided invaluable ideas and feedback every step of the way. I would also like to thank my advisor, Valentino Larcinese, for his continued support, and Bob Mattes who provided invaluable guidance during my time at the Democracy in Africa Research Unit at the University of Cape Town. At Harvard University, where I spent the last year of my PhD, I benefited greatly from Jim Alt's mentorship. The spatial analysis in my thesis was possible thanks to the support from Harvard's Center for Geographic Analysis and Institute for Quantitative Social Science.

The first paper benefitted particularly from discussions with Joachim Wehner, Jim Alt, Albert van Zyl and Robert Cameron. The second paper was co-authored with Joachim Wehner and Valentino Larcinese and we received helpful feedback from Mark Borchers, Robert Cameron, Daniel de Kadtt, Torun Dewan, Taryn Dinkelman, Andy Eggers, Steven Friedman, Elliott Green, Dominik Hangartner, Simon Hix, Mareike Kleine, Johannes Lindvall, Brian Min, Ian Osgood, Pablo Querubin, David Soskice, Johannes Urpelainen, and Leonard Wantchekon. The idea for the third paper was inspired by Brian Min and greatly benefitted from written feedback from Anne-Marie Jeannet and Clara Rubincam. For support with data and spatial analysis relevant to all chapters, I wish to extend a special thank you to Simo Goshev at the Harvard-MIT Data Center, Kevin Perry at StatsSA, and Giovanni Zambotti and Stacy Bogan at the Harvard Center for Geographic Analysis.

Table of Contents

Chapter I		
Introduction	10	
1.1	Investigating the role of political competition as an incentive to redistribute public resources	11
1.2	Emphasizing the institutional environment as a determinant of available policy options	13
1.3	Explaining distributive patterns of public service delivery	15
1.4	Using South Africa as a testing ground for the politics of service delivery	17
1.5	Highlighting the importance of measurement and good data	20
Chapter II		
Political budget cycles and intergovernmental transfers in a dominant party framework: empirical evidence from South Africa	22	
2.1	Introduction	23
2.2	A distributive politics perspective on political budget cycles	26
2.3	Incentive and ability in the South African context	30
2.4	Empirical strategy	40
2.5	Estimation results	47
2.6	Evaluation and further discussion	53
2.7	Conclusion	57
Chapter III		
A better life for all? Democratization and electrification in post-apartheid South Africa	60	
3.1	Introduction	61
3.2	Enfranchisement, political parties, and electrification	67
3.3	Variables and data	73
3.4	Empirical strategy	84

3.5	Main results	85
3.6	The mediating role of political parties	95
3.7	Conclusion	102
Chapter IV		
How can we study socioeconomic outcomes from outer space? Opportunities and pitfalls of using nighttime lights data in quantitative social science research		106
4.1	Introduction	107
4.2	A social science perspective on nighttime light satellite imagery	112
4.3	Turning nighttime lights into proxies for socioeconomic outcomes	124
4.4	Cross-validation with cross-country and census data	128
4.5	Categorizing pitfalls and navigating them	141
4.6	Conclusion	152
Chapter V		155
Discussion and implications		155
References		166
Appendix		193
Appendix A		193
Appendix B		196
Appendix C		213

List of Tables

Table 1: Electoral outcomes in South African provinces, 1994, 1999, 2004 and 2009	33
Table 2: Definition and descriptive statistics for key variables	44
Table 3: The impact of vote margin on intergovernmental transfers in pre- election years.....	51
Table 4: The evolution of the pre-election vote margin effect over time ..	56
Table 5: The impact of enfranchisement on electrification (census data) .	88
Table 6: The impact of enfranchisement on electrification (satellite data, with placebo regressions).....	93
Table 7: The impact of enfranchisement on electrification (contiguous census tract pairs).....	94
Table 8: The role of the ANC's seat share on local councils.....	101
Table 9: Correlating different measures of socioeconomic outcomes in countries	133
Table 10: Correlating different measures of socioeconomic outcomes in South African municipalities	140
Table 11: Overview matrix of electrification measures	150
Appendix Table A1: Components of the equitable share formula between 1995 and 2010	193
Appendix Table A2: Description of variables	194
Appendix Table B1: Geographical hierarchies in census 1996 and 2001	205
Appendix Table B2: Descriptive statistics.....	209
Appendix Table B3: Robustness to excluding municipalities with high levels of electrification in 1996.....	210

Appendix Table B4: Robustness to excluding municipalities in individual provinces.....	211
Appendix Table C1: Selected empirical applications of nighttime light satellite imagery in descending order	218
Appendix Table C2: Tracing changes in socioeconomic outcomes over time	219

List of Figures

Figure 1: Breakdown of total provincial revenues into own receipts, conditional grants and equitable share, 1995-2010.....	38
Figure 2: South African provinces and their average share of the total population, 1995-2010	38
Figure 3: Counterfactual effect of vote margin on equitable shares to provinces relative to Limpopo.....	52
Figure 4: Overview of South Africa's census geography in 1996	75
Figure 5: Geographic distribution of key variables	76
Figure 6: Nighttime lights as a proxy for municipal household electrification rates	83
Figure 7: Enfranchisement conditional on ANC seat share quartile.....	98
Figure 8: Views of the African continent, Western Europe and South Africa by night (2001)	113
Figure 9: Showcasing differences between satellite and census based electrification rates	129
Figure 10: Visualizing changes in socioeconomic outcomes in South Africa between 1996 and 2011	134
Figure 11: Possible outcomes of observed pixels in nighttime light imagery	142
Appendix Figure B1: Municipal demarcation (green lines indicate 1996 and red lines 2001 boundaries)	197
Appendix Figure B2: Overlay of geographical hierarchies (green lines indicate 1996 municipalities and red lines 2001 subplaces)	200
Appendix Figure B3: Conditional effect with alternative measures of partisan representation.....	212

Chapter I

Introduction

Who gets what, when and how? This is one of the most basic questions in political economy. It also sets the scene for this thesis, which focuses on how government affects the distribution of financial resources and services. I build on the proposition that different types of government generate different policy outcomes by empirically examining the distribution of public goods and services as a country transitions to democracy. It is the fuzzy path along the spectrum from autocracy to democracy that lies at the heart of my thesis. When it comes to the distribution of public goods, the poles of the spectrum seem more clear-cut: there is a strong consensus that democracy is the type of government that is best at satisfying the needs of the people. In Sen's (1999: 16) words, "no famine has ever taken place in the history of the world in a functioning democracy." Indeed, there is a large literature that suggests democratic governments are better at providing public goods and

services than autocratic regimes (e.g., Acemoglu and Robinson 2000a, Lake and Baum 2001, Bueno de Mesquita et al. 2003, Stasavage 2005, Blaydes and Kayser 2011, Kudamatsu 2012).

However, this literature raises important questions that are fundamental to my thesis: what exactly is it about democratization that leads to redistribution? Is the implementation of universal suffrage the key mechanism driving the change? What roles do political competition and the institutional environment play in determining the redistribution of public resources? With the following three chapters, I aim to make a small contribution to answering specific aspects of these big questions. Each chapter presents separate tests of related propositions about elections, public finance and service delivery and was authored as independent, self-sufficient pieces of research with separate conclusions and implications. To offer a preview, this introduction presents a brief discussion of the core themes that unite the three chapters.

1.1 Investigating the role of political competition as an incentive to redistribute public resources

The aforementioned political economy literature posits that governments subject to electoral competition have an incentive to redistribute public goods and services in order to satisfy their electorate. The underlying assumption is that politicians in democracies are motivated by the desire to retain public office (Golden and Min 2013: 77). Empirically, this idea should be particularly noticeable when regimes transition from autocracy to democracy. Acemoglu and Robinson (2000b: 684) argue that the

process of democratization changes the identity of the median voter, thereby making a credible commitment to future redistribution. There is an extensive literature that supports this observation, suggesting shifts in the median voter due to the extension of the franchise improve redistribution towards the poor (e.g., Meltzer and Richard 1981, Husted and Kenny 1997, Boix 2003) and benefit the newly enfranchised (Miller 2008, Fujiwara 2013, Vernby 2013).

My thesis builds on this literature by examining the effect of democratization on the provision of basic public services. In line with the recent political economy literature, I emphasize the incentives of political parties “who may be opportunistic and implement policies so as to maximise their chances of re-election, or be partisan and so want to further the interests of their own support groups” (Arulampalam et al. 2009: 103). Indeed, when it comes to the provision of public goods and services, central governments cannot be expected “to behave as a benevolent social planner would do” (Brollo and Nannicini 2012: 742). The role of political competition and particularly political alignment of lower tiers of government therefore lie at the heart of my investigation.

In this context, I focus on the strategic role of political competition with respect to redistribution, as “an election that is expected to be ‘closer’ increases the value and thus the desirability of any action that increases the incumbent’s re-election prospects” (Alt and Rose 2007: 12). Such actions may include periodic fluctuations in fiscal policy variables in the vicinity of elections, which the literature refers to as political budget cycles (PBCs). There are two main theories that explain why political

budget cycles arise: partisan and opportunistic theories. Partisan theory, starting with Hibbs (1977) and Alesina (1987), states that PBCs are predetermined by the ideology of the incumbent government. However, this explanation is unlikely to be useful for studying PBCs in developing countries, where political parties generally do not exhibit the typical Western left-right pattern (Block 2002b: 209, Shi and Svensson 2003: 68). Since South Africa provides the testing ground for my investigation, I focus on the opportunistic theory pioneered by Nordhaus (1975)¹ and later developed by Rogoff and Sibert (1988), and Rogoff (1990). The main proposition that I test is that incumbents signal their competence to the electorate and maximise their chance for re-election through the manipulation of fiscal variables. Political competition functions as the key incentive for such manipulations and it is in this context that I carry out empirical tests to determine how electoral competition affects the distribution of public resources.

1.2 Emphasizing the institutional environment as a determinant of available policy options

Drawing on Tufte's (1978) motive, opportunity and weapon analogy Alt and Rose (2005) argue that incumbents must not only have an incentive to manipulate policy, but also the ability to do so. While political competition may function as an incentive to redistribute public resources, the institutional environment determines the policy options available to

¹ The subject of Nordhaus' theory is the political business cycles, which refers to electorally timed distortions in macroeconomic variables, such as inflation and unemployment. This is different from PBCs, which refer to periodic fluctuations in fiscal policy variables (i.e. variables affecting the budget).

the incumbent. That explains why Alt and Rose's (2007) theory of context-conditional budget cycles is based on two conditions: incentive and ability. The first condition relates to the incentivising role of political competition as discussed above. The second condition relates to the institutional environment. If the institutional environment makes it difficult to manipulate policy instruments, politicians are less able to resort to electioneering. In many countries, independent central banks have almost entirely eliminated the possibility for politicians to influence monetary policy for electoral purposes. Alt and Rose (2007: 6) cite this as one of the reasons why the literature has shifted its focus from outcomes, such as unemployment and inflation, to policy instruments, such as spending. While incumbent politicians tend to have more control over fiscal policy than over monetary policy, the institutional environment often limits political discretion over spending, for example through fiscal rules or budget transparency initiatives. However, it is difficult to fully shelter the distribution of public resources from political distortions, as "incumbent politicians have a lot of discretion in using them to tease swing voters, reward core supporters, or build alliances with other politicians" (Brollo and Nannicini 2012: 742).

Given the importance of the institutional environment in curtailing political discretion, it is not surprising that the magnitude of political budget cycles has been found to be far greater in developing countries (Shi and Svensson 2002) and new democracies (Brender and Drazen 2005). Striking examples include the 400 per cent increase in money supply by Zambian President Kaunda prior to the 1991 election, or the confiscation and redistribution of foreign currency bank accounts of Zimbabwean

President Mugabe prior to the 2008 election (Chauvet and Collier 2009: 514). These albeit extreme examples emphasize the importance of a country's institutional environment: it is easier to implement political budget cycles undetected in countries where institutions are fragile and impose few checks and balances on executive authority, and most media outlets are state-owned. While South Africa's constitutional oversight is stronger and the media landscape more vibrant than in most other Sub-Saharan African countries, such factors are a main motivation for why my thesis examines political budget cycles in the South African context.

1.3 Explaining distributive patterns of public service delivery

The distributive politics literature offers useful insights into the spatial distribution of public resources. This literature explains how policies that involve taxes and transfers, and in particular the decision about allocations of government goods and services are distributed to identifiable localities or groups (Golden and Min 2013: 74). There are two dominant models, which make predictions regarding the incentives for resource allocation across different regions based on their level of competitiveness. On the one hand, Cox and McCubbins' (1986) core supporter model predicts that incumbent governments allocate more resources to their political support groups in order to maximize the return of votes. In the case of Sub-Saharan Africa one may view this as a form of patronage i.e., when an incumbent selectively distributes "favours and material benefits to loyal followers" (Bratton and van de Walle 1997: 62). The basic idea is thus to reward loyalty and punish

disloyalty, an incentive that is independent of electoral competition. This model provides a possible explanation for why incumbents may choose to manipulate the distribution of resources even if they do not face electoral competition.

Empirical evidence is mixed. Some endorse the core supporter model, e.g., Larcinese, Snyder and Testa (2013) use exit polls to measure voter preferences in the United States (US) over the period 1978-2002 and find that politicians favour areas that contain a large percentage of their core supporters. They find no association between the amount of federal funds and the degree of partisan balance in a state. Miguel and Zaidi (2003) use a regression discontinuity design to show that in Ghana, education expenditure per student was 27 per cent higher in districts in which the ruling party had an overwhelming majority of votes when compared to other districts in 1989. Similarly, Case (2001) finds that the size of the block grants in Albanian districts depends on the district's political leanings and increases with the president's vote share in the previous election.

On the other hand, Lindbeck and Weibull (1993), and Dixit and Londregan (1996, 1998), argue that rational incumbents target swing regions in which they have a higher chance of impacting vote choices. Empirical evidence in support of this model includes Dahlberg and Johansson (2002) who find that a Swedish municipal grant programme has been used to distribute transfers to regions where there are many swing voters. Moreover, Kwon (2005: 324) finds that national subsidies in South Korea were allocated to swing provinces in which electoral

contests were competitive, while no such allocations were made to non-competitive provinces. In an investigation of intergovernmental transfers in Ghana, Banful (2010) shows that certain grants were higher in districts where vote margins in the previous presidential election were lower, also suggesting that swing districts were targeted.

It is in this context that my thesis differentiates between core and swing provinces to investigate the theory of context-conditional political budget cycles. My thesis also highlights the essential complementarity of distributive politics and median voter accounts in order to interpret spatial patterns of service delivery following a democratic transition.

1.4 Using South Africa as a testing ground for the politics of service delivery

My thesis focuses on the politics of service delivery following the historic 1994 South African election. Under apartheid, black South Africans were not represented in Parliament, but through separate Black Local Authorities. Indian and coloured population groups had limited representation in the Tricameral Parliament, but the system lacked legitimacy. Effectively, the election in 1994 thus increased the electorate by over 500 per cent, enabling all South Africans aged 18 or above to participate in a credible, democratic election for the first time. South Africa thus represents a case of contemporary democratization on a unique scale. This is a key advantage over comparable studies, such as Larcinese (2014) who analyses the effect of Italy's franchise extension in 1912 to 5 million voters or Berlinski and Dewan (2011) who examine the

effect of 1 million newly registered voters following the United Kingdom's Second Reform Act in 1867. The modern-day South African setting with 19 million enfranchised voters makes for an ideal testing ground.

At the same time, South Africa is a country that is highly heterogeneous on the subnational level, which lends itself well for empirical analyses: while South Africa has a dominant party system, subnational units vary greatly in terms of electoral competition faced by the ruling party, the number of newly enfranchised voters, as well as the extent of backlog in the provision of basic public services. In the following chapters, I exploit the observed spatial and temporal heterogeneity of different outcomes in South Africa's subnational units.

Another advantage of the South African setting is that the three-tiered system with national, provincial and local levels of government offers the possibility to compare across subnational units. The availability of high quality data reduces the risk of measurement error, particularly compared to other developing countries: with a high level of fiscal transparency, even in comparison to most developed countries, South Africa's Treasury makes extensive subnational budget information publicly available on an annual basis.² In addition to this, South Africa has carried out three censuses in the past two decades, in 1996, 2001, and 2011. Since even most industrialised countries only carry out one census per decade, the frequency of the data, coupled with its high quality, present a strong

² South Africa ranks second out of 100 countries in the Open Budget Index 2012, which measures budget transparency and accountability (Open Budget Index 2012).

comparative advantage for studying elections, public finance, and service delivery in South Africa.

While the South African context offers many advantages, it also poses particular challenges to disentangling the effect of democratization on public service delivery. Indeed, South Africa's "negotiated revolution" went far beyond the implementation of universal suffrage (Sparks 1995), which increases the risk of compound treatment effects. Indeed, the release of political prisoners and liberalization of previously banned political organizations in 1990 paved the way for major transformations, including the implementation of two democratic constitutions, the creation of new national, provincial and local spheres of government, the complete overhaul of fiscal structures and new assignment of functions and resources to the different spheres (Wehner 2000: 47-48). Disentangling the effect of franchise extension from the other transformations that took place at the same time is therefore a major challenge in the South African setting. A cleaner identification of the effect of enfranchisement is possible in the cases of the United Kingdom in 1867 (Berlinski and Dewan 2011) or Italy in 1912 (Larcinese 2014) where the extension of voting rights to unskilled population groups was the only major change during the periods under review. In order to take a step towards addressing this compound treatment problem, chapter 3 disentangles the effect of franchise extension from political representation. However, this does not fully eliminate the possibility of a compounding effect of the other transformations associated with the date of South Africa's first democratic election, as further discussed in chapter 3.

1.5 Highlighting the importance of measurement and good data

In order to empirically determine the effect of democracy on redistribution, the quantification of both measures is an important precondition, which I argue should not be underestimated. Most existing papers use standard indices of democracy (Munck and Verkuilen 2002). The use of such indices in particular, and generally the reliance on data based on multiple national sources with different data collection methodologies, gives rise to concerns about measurement error. Put simply, bad data produce bad results. Good quantitative research requires good data. Subnational level analyses reduce concerns of comparability across units, but that in itself does not guarantee data quality, particularly in Africa (Jerven 2013). That is why technological advances in the areas of geographic information systems and remote sensing are powerful tools for quantitative social science research. Nighttime light satellite imagery has the potential to address the growing need for high-resolution data. However, there are important risks associated with any application of these data to a given context. The objective of the third paper in this thesis is to enhance the usefulness of nighttime light satellite imagery to measure socioeconomic outcomes for social science research.

Nonetheless, studying outcomes that are related to government decision-making is a challenging endeavour, particularly cross-nationally (Golden and Min 2013: 84). Since subnational units tend to be uniform in terms of electoral rules and scope for distributive policy making, this level of

analysis naturally offers a higher degree of comparability than cross-national studies. Cross-national studies face the difficulty of having to control for different institutional regimes, however, any such attempt is necessarily imperfect (Kneebone and McKenzie 2006: 755). This greatly increases the risk of omitted variable bias, reverse causality, and sample selection bias, making it very difficult to convincingly establish causality in such studies. My thesis contributes to the still limited inventory of sub-national analyses of distributive outcomes: the next chapter uses South African provinces as the unit of analysis, the third chapter analyses municipalities, and the fourth demonstrates how to examine local socioeconomic outcomes at a resolution as fine grained as one square kilometre. My thesis makes a methodological contribution to the growing body of literature using nighttime lights, by highlighting potential Type I and Type II errors and thus showing how to enhance the social science toolkit with this new data source.

The remainder of this thesis is structured as follows: each of the three following chapters includes one paper, authored as a self-contained piece of research. Each addresses one aspect of the political economy in South Africa as introduced above. In the final chapter, I discuss the substantive and methodological contributions of the three papers, address their external validity, and highlight future areas of research.

Chapter II

Political budget cycles and intergovernmental transfers in a dominant party framework: empirical evidence from South Africa

Abstract

This chapter tests the theory of context-conditional political budget cycles in South Africa's dominant party framework, and demonstrates that the central government distributes intergovernmental transfers according to electoral motivation. Using a unique panel dataset comprising South Africa's nine provinces over the period 1995-2010, I find that provinces where the national ruling party faces less electoral competition receive a higher per capita equitable share in pre-election years. Political influence on the intergovernmental transfer system is possible despite the fact that

the equitable share is based on a formula. However, the impact of vote margin on the equitable share in pre-election years decreases over time in line with existing accounts of political budget cycles in new democracies. Overall, the results suggest that even in a dominant party framework, electoral competition can trigger the political targeting of intergovernmental transfers in the vicinity of elections.

2.1 Introduction

The popular perception that incumbent politicians manipulate fiscal policies to increase their chances of re-election is given formal expression through political budget cycle theory. Building on the pioneering work of Nordhaus (1975), the literature explains the phenomenon in terms of a moral hazard model and asymmetric information. Incumbent politicians are predicted to take advantage of informational asymmetries to signal their competence towards the electorate, for example, by demonstrating their ability to produce public goods without raising taxes (Shi and Svensson 2002: 70). The result of this behaviour is a “periodic fluctuation in a government’s fiscal policies induced by the cycle of elections” (Alt and Rose 2007: 1), which is the generally accepted definition of PBCs.

In the past, most empirical studies have tested PBC theory in the context of developed countries and have provided evidence on how the magnitude and composition of cycles varies with respect to different fiscal variables.³ More recent work has started to analyse PBCs in developing countries and generally finds that cycles there are more pronounced than

³See Drazen (2000) or Franzese (2002) for a comprehensive review of the empirical literature.

in developed countries (Brender and Drazen 2005, Block et al. 2003, Shi and Svensson 2002, Schuknecht 2000). However, most of these studies use country-level panels, which means that they cannot provide insight into the mechanisms through which PBCs are implemented at the subnational level. While some papers have carried out subnational analyses, they have generally focused on traditional PBC variables, such as revenue and tax, but not intergovernmental transfers.⁴ These are likely to be important channels for PBCs in countries where the subnational level of government has limited revenue-generating power. Moreover, while an emphasis has been placed on studying PBCs in new democracies, the literature has not explicitly applied PBC theory to the context of a dominant party framework. Hence there is little understanding about a dominant incumbent's motivation to use fiscal policy to win votes in such a context.

Motivated by this gap in the literature, this chapter examines the presence of opportunistic PBCs in South Africa's intergovernmental transfer system. South Africa is an ideal case to consider as the federal structure offers the possibility to compare across provinces. Moreover, even though South Africa has a dominant party system, the provinces vary greatly in terms of electoral success. In other words, the main party, the African National Congress (ANC), does not win everywhere all the time, although it controls the national level. This variation in subnational electoral competition can be exploited to identify its effect on intergovernmental

⁴ Among others, Akhmedov and Zhuravskaya (2004) carry out a subnational analysis on Russian regions. Moreover, Cerdá and Vergara (2007) analyse Chilean municipalities, Gonzales (2002) focuses on Mexican regions, and Khemani (2004) on Indian Provinces. To date, no subnational PBC studies have been carried out in Sub-Saharan Africa, for which lack of quality data seems to be an obvious explanation.

transfers. While provinces are responsible for independently implementing their own budgets, they are almost entirely financed by the central government. Indeed, the provinces' own receipts are typically less than 5 per cent. It is thus intuitive to analyse whether the dominant party can use the centralized intergovernmental transfer system to distribute funds according to political motivations.

This chapter finds that provinces where the ANC has a higher vote margin receive a higher equitable share through the provincial intergovernmental transfer system. This increase is possible despite the presence of a formula-based mechanism, as the ANC controls two important levers: formula revisions and data updates. Overall, the findings suggest that electoral competition on the subnational level is a salient feature when it comes to distributing national funds through South Africa's intergovernmental transfer system.

The remainder of this chapter is organized as follows. The next section provides a distributive politics perspective on political budget cycle theory and its empirical predictions. Section 2.3 assesses the role of incentive and ability to manipulate fiscal variables in the South African context. This is followed by a description of the empirical strategy and presentation of the estimation results in sections 2.4. and 2.5 respectively. In section 2.6 I further discuss the results, particularly regarding the evolution of the vote margin effect over time. Finally, the conclusion summarises the results and discusses policy implications.

2.2 A distributive politics perspective on political budget cycles

Originally, most empirical studies on PBCs focused on developed countries, which is the context in which PBC theory was conceived (Block 2002a: 2). An increasing number of empirical studies now concentrates on developing countries. Empirical evidence often confirms the presence of opportunistic PBCs, although a publication bias against null results may distort the real picture. Schuknecht (2000) conducts a study of 24 developing countries for the 1973-1992 period and finds evidence for PBCs, particularly in the form of public investment. In a study of sub-Saharan Africa, Block et al. (2003) discover significant election-year increases in public expenditure and net claims on the government. Shi and Svensson (2002) find that in a panel of selected developing countries, government spending increases and revenues fall before elections, leading to larger deficits in election years. Crucially, the latter study also finds that the magnitude of PBCs is far greater in developing countries than it is in developed ones. Other scholars have extended this finding to argue that in comparison to developed countries, PBCs are not only more pronounced in developing countries, but also in new democracies (see Brender and Drazen 2005, Akhmedov and Zhuravskaya 2004).

Evidently, opportunistic PBC theory does not apply equally under all circumstances. If the extent of informational asymmetry is a function of political institutions and voter characteristics, then the possibility to implement PBCs is also a function of these factors. Alt and Rose (2007) use this logic to argue that PBCs are context-conditional. In particular,

they employ Tufte's (1978) motive, opportunity and weapon analogy to argue that two conditions must hold: incumbents must have both an incentive and the ability to manipulate policy.

The first condition relates to the question of why an incumbent would choose to manipulate public finances. In the political budget cycle literature, the competitiveness of an election has often been used as an explanation. Indeed, Alt and Rose (2007: 12) trace the idea back to Wright (1974) and Tufte (1978) and explain how competitiveness functions as a strategic variable, as "an election that is expected to be 'closer' increases the value and thus the desirability of any action that increases the incumbent's re-election prospects". But exactly what type of action maximises the incumbent's re-election prospects? The distributive politics literature provides two models that make predictions regarding the incentives for resource allocation across different regions based on their level of competitiveness. Neither one focuses on the timing of fiscal distortions in the way that PBC theory does, but both add a perspective on distribution.

Lindbeck and Weilbull's (1993) and Dixit and Londregan's (1996, 1998) swing voter model predicts that parties strategically allocate public resources to swing voters to maximise their chances for re-election. This makes sense as "voters who are predisposed in favour of [a party] on partisan or programmatic grounds cannot credibly threaten to punish their favoured party if it withdraws rewards" (Stokes 2005: 317). On the other hand, Cox and McCubbins' (1986) core supporter model suggests that incumbent governments allocate more resources to political support

groups in order to maximize the return of votes and reward loyalty. Stokes (2005) and Golden and Min (2013) provide insightful reviews of the core versus swing voter debate, which I omit here to focus on their empirical implications.

As indicated in the introduction, empirical evidence on distributive politics has been mixed.⁵ If there is one theme, then it is that patterns of distribution and favouritism vary across countries and outcomes (Kramon and Posner 2013: 467). Nonetheless, the empirical predictions of the two distributive politics models add an interesting nuance to political budget cycle theory: periodic fluctuations in fiscal policy variables in the vicinity of elections may be conditional on the level of electoral competition in a given region. This paper builds on the existing literature by empirically examining this question in the context of South Africa's intergovernmental transfer system.

Alt and Rose's second condition, the ability to manipulate policy, relates to the institutional environment. If the institutional environment makes it difficult to manipulate policy instruments, politicians are less able to resort to electioneering. While politicians tend to have more control over fiscal policy than over monetary policy, the institutional environment often limits political discretion over spending, for example through fiscal rules or budget transparency initiatives.

⁵ See Golden and Min (2013) and Kramon and Posner (2013) for a review of the existing empirical literature on distributive politics.

Such factors may constrain an incumbent's ability to implement electoral fiscal effects. In a sample of 19 OECD countries in the 1990s, Alt and Lassen (2005) find that higher transparency countries have lower deficits and debt accumulation, while lower transparency countries display a persistent pattern of electoral cycles. In this vein, Rose (2006) shows that fiscal rules can mitigate PBCs by limiting the incumbent's ability to increase spending in the vicinity of an election. Benito et al. (2013) show how the balanced budget rule has succeeded in limiting the budget deficit in a sample of the largest Spanish municipalities in 1994-2009. Khemani (2007) analyses fiscal transfers to Indian states in 1972-1995 and finds electoral fiscal effects only in states where transfers are determined by the central political executive rather than by an independent agency.

To constrain political manipulations of fiscal policy, many countries (e.g., Austria, Brazil, Japan, Ghana and South Africa) have implemented intergovernmental transfer formulae based on objective and quantifiable criteria. While such attempts are not likely to be equally effective, empirical evaluations are scarce. Banful (2011) is an exception and in her study on Ghana's intergovernmental transfer system she finds that the formula-based mechanism does not eliminate politically motivated targeting of groups. Instead, Banful (2011: 289) finds that districts with more swing voters have systematically benefitted from changes to the formula. Her finding that the mere presence of a formula-based mechanism cannot be assumed to fully remove political influence over the distribution of funds, motivates this paper.

Regardless of the policy instrument, we may expect political budget cycles to weaken over time as young democracies gain experience with the electoral cycle and develop their institutions and checks and balances. Indeed, Brender and Drazen (2005: 10) demonstrate that cycles become less strong in new democracies as voters grow suspect of fiscal manipulations in the vicinity of elections and punish rather than reward such activities at the polls. Their point is not that the voting population in new democracies is naïve, but rather that in “countries with less of an electoral history, and hence less exposure to pre-electoral fiscal manipulations, a political cycle is more likely to occur” (Brender and Drazen 2005: 18).

The empirical analysis in this paper builds on the existing literature, by determining whether South Africa’s intergovernmental transfer system, which has both a formula-based and non-formula-based component, is subject to fiscal manipulations and if so, whether such manipulations display distributive patterns and whether this effect changes over time.

2.3 Incentive and ability in the South African context

There are four main reasons why South Africa is a useful case to consider for this analysis. First, its federal structure offers the possibility to make comparisons across provinces. A main problem with country level panels is the inherent difficulty to control for cross-country differences, such as the rule of law or certain institutions that could be endogenous to fiscal policy. An examination of fiscal variables in the South African provinces

avoids many of these problems, as the provinces are characterized by high uniformity in electoral rules, institutional framework, and governance structures in general. Second, although South Africa's political landscape is dominated by one party on the national level, the subnational level displays high variation in the level of electoral competition, whereby at least two of nine provinces may be considered swing provinces. This variation makes it possible to identify the effect of electoral competition on the political budget cycle. Third, provincial revenues derive almost entirely from central government transfers, which implies that investigating intergovernmental transfers to the provinces means investigating the only possible source of PBCs on the subnational level. This circumstance renders the South African setting particularly relevant for the study of PBCs and intergovernmental grants. Finally, South Africa has the highest level of fiscal transparency among Sub-Saharan African countries. This is a significant advantage because extensive subnational budget information is publicly available, which renders this chapter's methodology feasible. Based on the discussion on PBC theory and distributive politics, two questions arise with respect to the South African context: does the South African government have an incentive to instigate political budget cycles? And if so, does it have the ability to implement them?

Incentives. With respect to the first question it seems unlikely that the ANC would have an incentive to implement PBCs on the national level, where it hardly faces any electoral competition. Winning 63, 66, and 70 per cent in the 1994, 1999, and 2004 elections respectively, the ANC has held a large and increasing majority (Independent Electoral Commission

2012). On the subnational level, South Africa's electoral system, which is based on proportional representation with minimum thresholds, has given parties other than the ANC chances of controlling provincial governments, such as the National Party (NP), the Inkatha Freedom Party (IFP), and more recently the Democratic Alliance (DA). Table 1 illustrates the variation of provincial electoral outcomes, by presenting the share of votes achieved by the first and second party in each province in the 1994, 1999, 2004, and 2009 elections. While the ANC won Limpopo province with as much as 92.7 per cent of votes in the 1994 election, it came second in KwaZulu-Natal (KZN) and the Western Cape with merely 31.6 and 33.6 per cent respectively.

Table 1: Electoral outcomes in South African provinces, 1994, 1999, 2004 and 2009

Province	1994		1999		2004		2009	
	Party	Vote share*						
Eastern Cape	ANC	84.4	ANC	73.9	ANC	79.3	ANC	68.8
	NP	10.6	NNP	12.9	UDM	8.9	COPE	13.7
Free State	ANC	77.7	ANC	81.0	ANC	82.1	ANC	71.1
	NP	14.5	DP	5.9	DA	8.9	COPE	11.6
Gauteng	ANC	59.1	ANC	68.2	ANC	68.7	ANC	64.0
	NP	27.6	DP	17.7	DA	20.3	DA	21.9
KwaZulu-Natal	IFP	48.6	IFP	40.5	ANC	47.5	ANC	63.0
	ANC	31.6	ANC	39.8	IFP	34.9	IFP	22.4
Limpopo	ANC	92.7	ANC	89.3	ANC	89.7	ANC	84.9
	NP	3.6	UDM	2.6	DA	3.8	COPE	7.5
Mpumalanga	ANC	81.9	ANC	85.3	ANC	86.3	ANC	85.6
	NP	10.3	DA	5.0	DA	7.2	DA	7.5
Northern Cape	ANC	49.8	ANC	64.4	ANC	68.8	ANC	60.8
	NP	41.9	NNP	22.5	DA	11.6	COPE	16.7
North West	ANC	83.5	ANC	80.5	ANC	81.8	ANC	72.9
	NP	10.1	UCDP	7.5	UCDP	6.5	COPE	8.3
Western Cape	NP	56.2	ANC	42.6	ANC	46.3	DA	51.5
	ANC	33.6	NNP	34.4	DA	26.9	ANC	31.6

Note: ANC = African National Congress; COPE = Congress of the People; DA = Democratic Alliance; DP = Democratic Party; IFP = Inkatha Freedom Party; NP = National Party; UDM = United Democratic Movement; UNCDP = United Christian Democratic Party. Source: Independent Electoral Commission 2012; * denotes per cent as unit of measurement.

The closeness of the political race on the provincial level may serve as an incentive to implement PBCs.⁶ From the point of view of the ANC, the main strategic question is whether to target provinces where it has a high vote margin or where it faces strong political opposition, or both. On the one hand, it would make sense to spend more money where it could potentially make a difference in the year before an election, for example in KZN or the Western Cape, but not in a province like Limpopo where the ANC can expect a vote share in excess of 90 per cent. On the other hand, this rationale is contingent on the assumption that voters attribute increases in spending to the ANC, rather than the provincial government. If the two are not the same, as has been the case in swing provinces such as the Western Cape or KwaZulu-Natal, this assumption seems unrealistic. Voters may attribute an increase in spending or the provision of certain public goods to the competence or bargaining success of the provincial government, which is *de facto* responsible for service provision.

Similarly to Brollo and Nannicini's (2012: 746) model of the politics of federal transfers in Brazil, shared responsibilities between central and provincial governments in South Africa mean that there are "political credit spillovers". I only expect the ANC to implement PBCs where it can claim political credit. This argument strengthens the core supporter prediction, as claiming credit is straightforward in the ANC's core provinces, where revenue and spending powers are aligned within the

⁶The ethnic census argument, which suggests that the variation in political competition mirrors ethnic patterns, could be interpreted as a disincentive. However, while ethnic voting patterns can explain some of the variation in South African political competition, Mattes (1995) shows that the nature of elections and political preferences are much more than a racial census. Therefore there is scope for vote-purchasing behaviour.

same party. In this context, Stokes (2005: 316) emphasizes that the crucial feature of the core supporter model is that the party is more certain about how core groups will respond to rewards than it is about other groups, making the targeting of swing voters less likely. Given the South African context, the empirical analysis in this chapter is guided by the core support model's prediction about the distribution effect of the political budget cycle: in the year before an election, I expect higher vote margins in a province to be associated with higher intergovernmental transfers. The model also evaluates the alternative outcomes, i.e. whether the opposite relationship can be observed or whether vote margins do not have a conditioning effect on intergovernmental transfers. These expectations are fully aligned with the theoretical models developed in Brollo and Nannicini's (2012) assessment of centre-municipality transfers in Brazil and Arulampalam et al.'s (2009) analysis of centre-state transfers in India.

Ability. Even if we assume the ANC has an incentive to implement PBCs through the intergovernmental transfer system, would the party be able to do so given the policy instruments available? In order to answer this question, it is important to understand South Africa's budgeting system, which changed substantially during the overhaul of government structures after the end of apartheid. Constitutional arrangements made the nine newly created provinces responsible for independently drafting and implementing their own budgets (SA Treasury 2000: 21). However, Wehner (2000: 71) shows that despite these decentralization efforts the system has operated in a highly centralised manner in practice. As illustrated in Figure 1, the provinces' own revenues have constituted less

than 5 per cent of total provincial revenues between 1995 and 2010, which means that provinces rely almost entirely on transfers from the national government to fund their activities.

The ability to implement political budget cycles on the provincial level is thus contingent on the ability to manipulate the transfer from the national government, which is made up of two components: an equitable share and conditional grants. With a range of between 80 per cent and 90 per cent of the total transfer between 1995 and 2010, the equitable share is the largest component of the total transfer. The main idea behind it is that “each level of government shall have a constitutional right to an equitable share of revenue collected nationally so as to ensure that provinces and local governments are able to provide basic services and execute the functions allocated to them” (SA Treasury 1999: 22). Unlike the equitable share, the second component of the total transfer to provinces, the conditional grant, is not determined by a formula. By definition, it provides for national priorities in provincial budgets (SA Treasury 1999: 15) and “is voted in the budget of a national department and reflected as a revenue item in provincial accounts” (SA Treasury 1999: 38). Wehner (2000) provides a detailed account of how the two components are integrated into South Africa’s budget cycle.

The decision making process around the conditional grant has often been criticized for lack of transparency and the former Head of Intergovernmental Relations in the National Treasury has argued that many conditional grants “lack a clear purpose and measurable objective, and are poorly designed” (Momoniat 1999: 12). While these features may

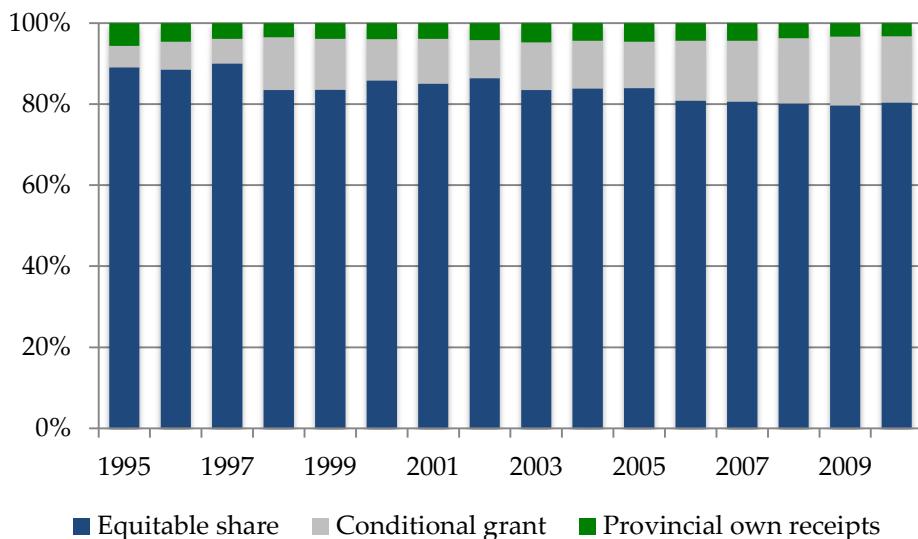
make it easier to manipulate conditional grants, the small share of the total transfer render it an unlikely target for the implementation of electorally motivated fiscal distortions. In 1995, 1 per cent of the equitable share was equivalent to 17 per cent of the conditional grant. Hence, a small distortion of the equitable share offers a larger effect in monetary terms compared to a similar distortion of the conditional grant in percentage terms. In this sense, politically motivated distortions of the equitable share generate higher impact at a lower risk of getting caught. But how would the ANC be able to get its hands on the lever of the equitable share? The equitable share is not only determined by a formula, but the division of national revenues is also overseen by an independent body, the Financial and Fiscal Commission (FFC).⁷ I argue that neither is sufficient in safeguarding the equitable share from political influence.

Since its implementation in 1997, the formula has had a number of components: an education share, a health share, a basic share, a backlog component, an economic output share and an institutional component.⁸ Definitions of the individual components are provided in Table A1 in the Appendix. Prior to the formula, the allocations to provinces have been described as a system where “all arrangements were subject to arbitrary political intervention” (Simkins in van Zyl 2003: 9). The formula was intended to curtail such interventions by largely deriving the distribution

⁷ The FFC was established with section 220 of the Constitution of the Republic of South Africa as an independent, impartial body subject only to the Constitution and the law (SA Constitution 1996).

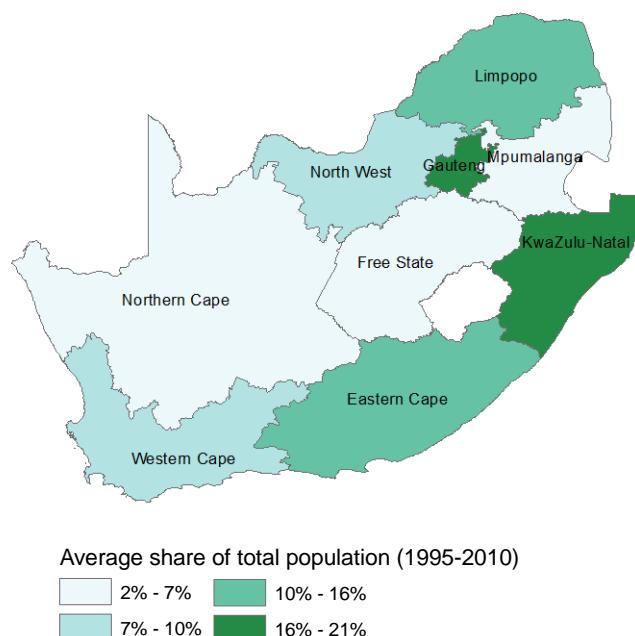
⁸ Two main components have historically been (and continue to be) education and health, having made up over half of the total annual transfer since 1995. Between 1997 and 2005 the formula also had a social welfare component, but this was removed in 2006 when responsibility for social security grants shifted from the provincial to the national sphere of government (SA Treasury, 2006: 49).

Figure 1: Breakdown of total provincial revenues into own receipts, conditional grants and equitable share, 1995-2010



Source: Author's calculations with data from SA Treasury 1999-2011.

Figure 2: South African provinces and their average share of the total population, 1995-2010



Source: Author's calculations with population data and shape file from Statistics South Africa 2011 Community Profiles and GIS database.

across provinces based on the provincial share of the total population. To illustrate, Figure 2 maps each of the nine South African provinces, such that darker shades indicate higher shares of total population and thus higher shares of total intergovernmental transfers.

Under the formula, there are two main discretionary elements available for manipulation: data revisions and formula revisions. Regarding the data, the equitable share formula is subject to change every year as information is reviewed or improved, which is particularly challenging given the “number of informal settlements and high levels of illiteracy and poverty” (Momoniat 1999: 7). One possibility for political intervention would be to strategically delay information updates to the formula. For the 1998 budget, for example, the figures used to determine the equitable share were based on the ‘October Household Survey 1995’, rather than on the preliminary results from the 1996 census. In retrospect, this decision had important consequences: provinces such as Gauteng and the Western Cape were two main destinations of migratory movements, which meant that their true population counts were greatly underestimated and budget allocations based on the 1995 figures put them at a distinct disadvantage (van Zyl 2003: 11).

Regarding revisions to the formula, one main problem is that the FFC lacks teeth and the Department of Finance has often ignored its recommendations (Wehner 2000: 69). The decision not to follow or delay the FFC’s recommendations is another lever for political control over the equitable share. For the budget year prior to the 1999 election, the FFC’s recommendations for the horizontal division of revenue between

provinces were set aside and replaced by the national Department of Finance's formula (Folscher et al. 1999: 31). Instead of implementing the FFC's formula in the pre-election year 1998, the Department of Finance promised that the issues raised by the FFC "will be addressed in the budget process for 1999/2000" (SA Treasury 1998: E10). The decision not to follow the FFC's recommendation to increase the subnational percentage of shared revenues reinforced the government's bias towards the centralized control of provincial revenue streams.

That revisions to an intergovernmental formula can be used to achieve political targeting of certain regions is not an anomaly: Banful (2011) finds empirical evidence of politically motivated revisions in Ghana's District Assemblies Common Fund (DACF) formula. The empirical strategy tests whether such manipulations occur in the South African case.

2.4 Empirical strategy

The empirical analysis is based on a unique panel data set, consisting of annual observations of South Africa's nine provinces for the period 1995 - 2010.⁹ The data on the fiscal variables to be tested for PBCs has been provided by the South African Treasury and are based on the Intergovernmental Fiscal Reviews (1999 - 2011). The fiscal variables include all provincial revenue variables, in terms of the actual allocations of total transfers, equitable shares, and conditional grants. The dataset therefore captures all of the provinces' annual receipts from national

⁹The time period has been determined by the availability of data; 1995 being the earliest and 2010 being the most recent year for which data were available at the time of writing.

government. This is a considerable advantage over other empirical studies (e.g., Banful 2011), which are selectively based on one type of transfer to subnational governments.

The time span of the data set includes three elections in 1999, 2004 and 2009. As is the case in most empirical studies of PBCs, the timing of these elections is taken to be exogenous to fiscal policies.¹⁰ In the case of South Africa, this is an appropriate assumption as the timing of these elections was fixed by constitutionally predetermined five-year intervals. The elections were not strategically delayed or advanced, and it would have been extremely difficult to do so. The national and provincial government fiscal year starts on 1 April and ends on 31 March the following year.¹¹ The elections took place during the first quarters of the fiscal year, on 2 June 1999, 14 April 2004 and 22 April 2009. This means that the fiscal year before the election is most relevant when it comes to the implementation of PBCs. In order to estimate the effect of these elections on the fiscal variables, the dummy *pre-election* takes the value one in the year before an election (i.e. in 1998, 2003 and 2008) and zero otherwise. While this variable changes over time, it is constant across provinces as elections are held at the same time in all provinces.

¹⁰ Shi and Svensson (2002), among others, relax this assumption with respect to countries in which the timing of elections is set strategically, for example at the time of an economic boom. However, they do not find a significant effect on their results when they exclude the countries in which they classify elections to have been endogenous (Shi and Svensson, 2002: 10). Khemani (2004) applies an instrumental variable to distinguish early and scheduled elections.

¹¹ The first quarter starts on 01 April and ends 30 June; the second quarter starts 01 July and ends 30 September; the third quarter starts on 01 October and ends on 31 December; the fourth quarter starts on 01 January and ends on 31 March.

Most empirical applications use voting data in order to measure regional electoral competition. Schultz (1995) uses opinion polls and Alt and Rose (2007) use governors' job approval ratings. The most common measure in subnational analyses is the vote margin of the incumbent government in the previous election (see Cingermayer and Wood 1995, Case 2001, Dahlberg and Johansson 2002, and Banful 2011). In line with this literature, I define the variable *vote margin* (1994, 1999, 2004, 2009) as the difference between the share of votes won by the first party and the second party in the previous provincial election. This quantitative measure of electoral competition changes over time and across provinces: Based on the election outcomes reported in Table 1, the vote margin between the ANC and the NP in the 1994 election is 63.2 per cent in the Free State, while it is 31.5 per cent in Gauteng. This indicates that the Free State is less competitive than Gauteng. Anecdotal evidence regarding the competitiveness of the provinces is in line with the ranking of provinces by this measure. In terms of interpretation, a higher vote margin percentage indicates less political competition and vice versa.

However, the main criticism of using this measure is that voting behaviour is endogenous to the policy variables of interest. Larcinese et al. (2006) address this issue by using exit polls to measure voter preferences and partisanship. While this may be a more appropriate measure, such data are not available for South African provinces. I therefore define an alternative measure of electoral competition in order to maximise the plausibility of exogeneity: *vote margin* (1994) captures electoral competition based on 1994 election results, rather than the results of

elections during the sample period. While it changes across provinces, it remains constant over time as I use 1994 vote margins for all time periods. The 1994 election represented the transition from apartheid to democracy and extended the franchise to approximately 20 million South Africans who had never voted before. It is highly unlikely that vote margins in this first democratic election were endogenous to provincial spending after the end of apartheid. Fixing the competition variable to vote margins from the 1994 election therefore greatly helps to reduce the risk of endogeneity. However, the downside of this measure is that it gets less precise over time. Indeed, the ANC results in 2009 do not bear a close resemblance to those in 1994 in all provinces. For this reason, I use both vote margin measures in the empirical analysis.

To provide an overview of key variables, Table 2 presents definitions and descriptive statistics. Table A2 in the Appendix provides further information on the characteristics, construction and sources of these variables.

Table 2: Definition and descriptive statistics for key variables

Variables	Definition	N	Mean	Stand. dev.	Min	Max
Total transfer	Total transfer from national government to provinces per capita, 1995-2010.	144	2,084.97	425.84	1,296.10	3,515.43
Equitable share	Equitable share per capita, 1995-2010.	144	1,823.78	346.97	1,066.37	2,699.46
Conditional grant	Conditional grants per capita, 1995-2010.	144	263.08	149.14	66.14	815.97
Vote margin (1994, 1999, 2004, 2009)	Vote margin is the absolute difference between the share of votes of the 1 st and 2 nd party in the previous provincial election.	144	53.89	26.72	0.70	89.10
Vote margin (1994)	Vote margin is the absolute difference between the share of votes of the 1 st and 2 nd party in the 1994 provincial elections.	144	50.01	28.46	7.90	89.10
Pre-election	Election dummy equal to 1 one year before an election and 0 otherwise.	144	0.19	0.39	0.00	1.00
Provincial GDP	Provincial GDP per capita, 1995-2010.	144	2.62E+04	1.27E+04	1.91E+03	5.86E+04
Population < 14	Number of people aged under 14 years by province, 1995-2010.	144	1.44E+06	9.16E+05	1.68E+05	3.70E+06
Population > 60	Number of people aged above 60 years by province, 1995-2010.	144	3.56E+05	1.94E+05	6.09E+04	8.22E+05

Note: All fiscal variables are measured in inflation adjusted ZAR per capita.

The aim of the identification strategy is to test the validity of the empirical predictions formulated above, i.e. whether the ANC implements PBCs conditional on vote margin, using the equitable share. For a given fiscal variable per capita¹² (fiscal) the baseline specification takes the following form:

$$\begin{aligned} \log(fiscal_{it}) = & \alpha + \beta_1 \text{ vote margin}_{it} + \beta_2 (\text{vote margin}_{it} \times \text{pre-election}_t) + \gamma Z_{it} \\ & + \sigma_i + \tau_t + u_{it} \end{aligned} \quad (1)$$

In this specification, fiscal is the dependent variable, which corresponds to each of the three fiscal variables to be tested for evidence of PBCs. Subscript i indexes the nine provinces ($i = 1,2,3\dots 9$) and t indexes the years (1995, 1996\dots 2010). The variable vote margin_i measures electoral competition as discussed in the previous section. The variable pre-election is a dummy equal to one in the year before an election and zero otherwise. Z_{it} is a vector of control variables, GDP per capita and demographic variables, σ_i represents province fixed effects and τ_t represents year fixed effects. Finally, u_{it} is the error term, which is estimated using autocorrelation and heteroskedasticity robust standard errors.¹³

¹²Since provinces differ largely in terms of population, and since spending and revenue are highly correlated with population, per capita measures are used to make comparisons across provinces.

¹³As fiscal variables for each province are likely correlated over time, it is generally advisable to use clustered standard errors. However, with only nine possible clusters, robust standard errors are used instead, following Nichols and Schaffer (2007: 7) who argue that with a number of clusters less than 50, it is generally argued that “the cure would be worse than the disease”.

The rationale behind equation (1) is that the dependent variable in each specification can be tested for evidence of fiscal manipulations in pre-election years, following the standard specifications in the PBC literature (see for example Faal 2007).¹⁴ The coefficient on vote margin_i, β_1 , shows the direct effect of vote margin on the dependent variable. The coefficient on the interaction term between pre-election_t and vote margin_i, β_2 , is used to test whether the electoral cycle depends on electoral competition; it is thus the key coefficient of interest. As per the empirical predictions formulated above, β_2 is expected to be positive and significant with respect to the equitable share.

Province fixed effects, σ_i , control for time-invariant omitted variables and absorb much of the effect of any slowly changing variables such as the level of development or the ratio of the economically active population in a province. These variables are important determinants of how much funding a province receives, while at the same time also determining electoral competition. For example, low levels of education in a province require higher educational spending, while also being associated with a higher share of votes for the ANC in the 1994 election (Johnson, 1996: 126). Year fixed effects, τ_t , control for aggregate shocks and the national business cycle effect, which have been found to exacerbate the electoral cycle in a given year, thus leading to an overestimation of PBCs (Kwon, 2005: 331). Since pre-election_t only varies across time but not across provinces, it is absorbed by the year fixed effects estimator, τ_t .

¹⁴ In the literature, lagged dependent variables are often used to control for fiscal inertia. However, as the lagged dependent variable in combination with fixed effects introduces a bias of magnitude $1/t$ (where t is only 16), I will not use this method in this context.

2.5 Estimation results

Table 3 reports the estimation results for the three main fiscal variables, the total transfer, the equitable share and conditional grants, based on equation (1). The positive and statistically significant coefficient on the interaction term between vote margin and pre-election in column (2) shows that as predicted provinces with higher vote margins received a higher per capita equitable share in pre-election years. This finding suggests that the ANC strategically channels a higher share of the main component of the total intergovernmental transfer to its core provinces.

For each percentage point vote margin, the equitable share is 0.053 per cent higher in pre-election years. This seems small, but given that the average vote margin over the time period under review is 50 per cent (see Table 2) the average effect is multiplied: in a province with a political competition profile like Mpumalanga, the equitable share is 2.65 per cent higher than in a province where the ANC only achieved 50 per cent of the vote in 1994. In terms of the mechanism, the ANC appears to successfully control the discretionary levers of the equitable share formula to channel additional funds to core support provinces in pre-election years.

Interestingly, this effect is not detected in the total per capita transfer to the provinces. The coefficient of the interaction term in column (1) of Table 3 is positive and thus points in the same direction as with respect to the equitable share, but it is statistically insignificant. If fiscal manipulation of the equitable share is implemented in pre-election years, why is this not detected in the total transfer to provinces? Although the equitable share is the main component, the total transfer also consists of

conditional grants. The latter is a very noisy measure, with a coefficient of variation, i.e. ratio of the standard deviation to the mean, in excess of 50 per cent. This helps explain the large standard error in column (1) relative to column (2). By the same token, the size of the coefficient in column (1) is smaller as it absorbs the opposing force from the conditional grant: vote margin has a strong negative, but far from statistically significant effect on the conditional grant in the year before an election (see column (3)). Despite the aforementioned lack of transparency in this component of the intergovernmental transfer, the conditional grant does not appear to be used systematically as an instrument for politically motivated fiscal distortions in the vicinity of elections. Given the evidence, we cannot accept the hypothesis that per capita total transfers or conditional grants increase with the vote margin in pre-election years.

In the first three columns of Table 3, vote margin is defined in terms of previous election results, analogous to Banful's (2011) analysis of a formula-based component of Ghana's intergovernmental transfer system. In comparison, the magnitude of the coefficient (.05 per cent) is slightly less than half compared to what Banful finds regarding the DACF in Ghana, which is also based on a formula. However, in monetary terms, the impact on South Africa's equitable share is far greater since the DACF only corresponds to approximately 5 per cent of national revenue in Ghana, whereas the equitable share consumes over 40 per cent of annual national revenue in South Africa.

However, a serious concern about the coefficients reported in columns (1) - (3) of Table 3 is that voting behaviour and thus vote margin is

endogenous to the fiscal variables under review if it varies over time. A useful exercise is thus to re-run the regressions with a measure of electoral competition that is fixed over time: columns (4) to (6) report regression results with vote margin (1994) (see Table 2 for a definition and descriptive statistics). Compared to column (1), the coefficient of the interaction term in column (4) is slightly smaller, but still far from statistically significant. Regarding the impact on the equitable share, the coefficient of the interaction term is significantly weaker, with a decrease of over 40 per cent, but still statistically significant at the 5 per cent level. Since vote margin only varies across provinces but not over time, its direct effect on fiscal variables in columns (4) - (6) is absorbed by the province fixed effects estimator and therefore does not enter separately in the equation. Since fixing the alternative vote margins to the 1994 election results greatly helps to reduce the risk of endogeneity, I consider the results reported using vote margin (1994) the most demanding and parsimonious.

In order to get a better sense of the effect of vote margin on the equitable share, it is useful to consider a counterfactual scenario in which the swing province KwaZulu-Natal were a core support province of the ANC, such as Limpopo province. Based on the measure vote margin, Limpopo province scores 89 per cent compared to 17 per cent in KZN. According to my results, KZN would receive an additional 2.7 per cent of equitable share payments per capita from the national government in pre-election years, if KZN had the same vote margin Limpopo in 1994. In monetary terms this would have amounted to an additional ZAR 1.18 billion

(approximately USD 115 million) in the year 2008. Figure 3 illustrates this example systematically for all provinces relative to Limpopo.

The idea for providing this counterfactual scenario is inspired by Kwon (2005: 338). In his example, the hypothetical additional allocation of national subsidies to South Cholla province if it were as contested as Seoul, is estimated at approximately USD 490,000. The distortion Kwon detects in South Korea is thus only a small fraction of the result presented here, even though the examples are based on roughly similar vote margins in the respective provinces. This benchmarking exercise suggests that the extent to which core provinces in South Africa have benefitted from the equitable share is relatively large.

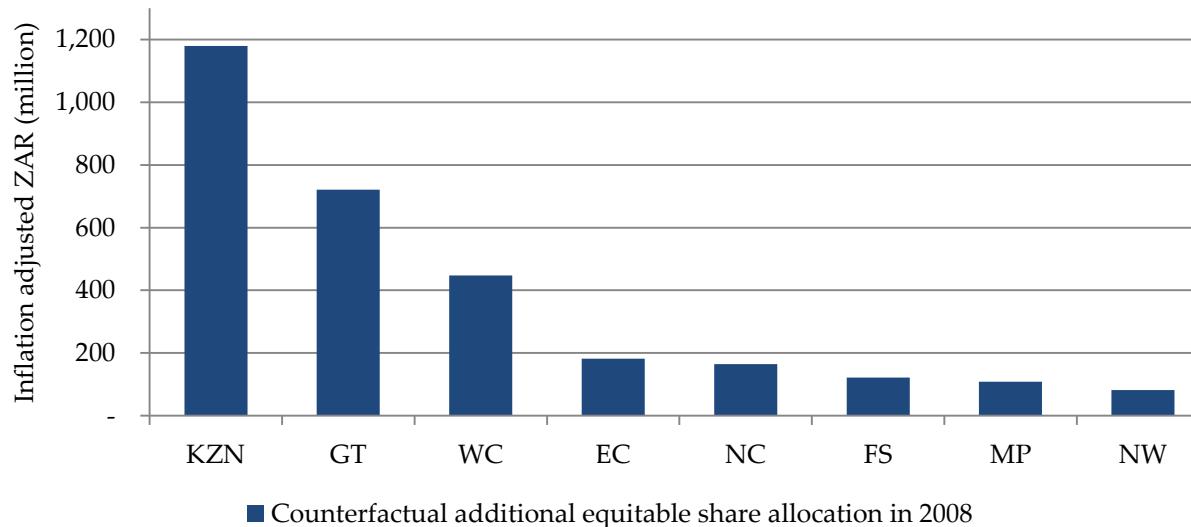
This is particularly worrying given that the coefficient in column (5) is a conservative estimate: in the main regression, vote margin is defined in terms of 1994 election results because one might argue that subsequent voting behaviour is endogenous to the fiscal variables under review. While fixing the competition variable to vote margins from 1994 helps to reduce the risk of endogeneity, the quality of the indicator is likely to decrease over time as it becomes more removed from the true level of electoral competition. Indeed, the ANC results in 2009 do not bear a close resemblance to those in 1994. This is particularly noticeable with respect to KZN where the ANC won just over 30 per cent of votes in the 1994 election and then more than doubled its share to over 60 per cent in the 2009 election (see Table 1).

Table 3: The impact of vote margin on intergovernmental transfers in pre-election years

Dependent variable	Log of real per capita total transfer	Log of real per capita equitable share	Log of real per conditional grant	Log of real per capita total transfer	Log of real per capita equitable share	Log of real per conditional grant
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Vote margin (1994, 1999, 2004, 2009): changes over time, previous election results time</i>						
Vote margin	.00070 (.00109)	.00010 (.00140)	.00056 (.00214)			
Pre-election × vote margin	0.00018 (0.00031)	0.00053** (0.00024)	-0.00139 (0.00108)	0.00012 (0.00030)	0.00037** (0.00017)	-0.00145 (0.00140)
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.87	0.79	0.84	0.87	0.79	0.84
Observations	144	144	144	144	144	144

Note: Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls are provincial GDP per capita, provincial population aged less than 14 years and provincial population aged 60 years or above. In columns (4) – (6) the variable vote margin varies across provinces, but not over time and is therefore absorbed by the province fixed effects and does not enter the regression separately. All regressions include a constant.

Figure 3: Counterfactual effect of vote margin on equitable shares to provinces relative to Limpopo



Note: this Figure is based on a counterfactual scenario in which the provinces of KwaZulu-Natal (KZN), Gauteng (GT), Western Cape (WC), Eastern Cape (EC), Northern Cape (NC), Free State (FS), Mpumalanga (MP) and North West (NW) are assumed to have the same vote margin as Limpopo province. The difference between each province's vote margin and Limpopo's vote margin is calculated and then multiplied with the vote margin effect in pre-election years (i.e. the coefficient in Table 3, column (2)). This percentage determines the counterfactual monetary effect on each province, in this case derived from each province's total equitable share in the year 2008.

2.6 Evaluation and further discussion

Based on this discussion, two factors are particularly worrying: first, the impact of the pre-electoral distortion of the main source of provincial revenues is large and significant. Second, the distortion is possible despite the implementation of a formula, which has been designed to ensure the fair distribution of funds, and despite the creation of the FFC. Other countries such as India have had positive experiences with the creation of an independent body to oversee their intergovernmental transfer system. Khemani (2007: 465) shows that, while the transfers that are determined by the central political executive are distributed to favour those Indian states that are politically important for the central ruling party, the transfers that are delegated to an independent agency serve to constrain such partisan impact. However, in South Africa the FFC lacks the teeth to curtail politically motivated targeting: “whereas the Constitution envisaged a key role for the Finance and Fiscal Commission in ensuring agreement on the assignment of revenue, it never gained enough influence to fulfil its function as an independent advising body” (Folscher et al. 1999: 31).

From a distributive politics point of view, the results presented above point in the opposite direction of what Banful (2011) finds in Ghana, where additional DACF funds are channelled to swing voter districts. In South Africa, a different strategy seems to be at play as provinces where the ANC has higher vote margins receive preferential treatment through the equitable share in pre-election years. Since the swing voter model hinges on attribution, it is unlikely to be an effective strategy in the context of South Africa’s political economy. Anecdotal evidence corroborates this

view: in an interview with van Zyl (2003: 10) Hennie Bester, the former leader of the Democratic Party in the Western Cape, suggests that “the budget of the province is reduced without any input from its side ... the province is nevertheless perceived as responsible for the decline in services that follows the reduction of its allocation. This would be a very effective strategy by the ANC to control the Western Cape, being a province that escapes from its direct political control”.

Brender and Drazen (2005) offer some insight into what we might expect from the evolution of politically motivated fiscal distortions in pre-election years over time. Their paper distinguishes between new and established democracies and suggests that electoral fiscal effects fade out over time as new democracies gain experience with elections (Brender and Drazen 2005: 10). In their sample, the pre-election effect on national fiscal balances loses statistical significance after the fourth election in new democracies and is generally insignificant for elections in established democracies. While South Africa has yet to hold its fifth election as it celebrates the 20th anniversary of the first democratic elections this year, the temporal analysis in Table 4 can still provide useful insights.

Columns (1), (3) and (5) of Table 4 report the coefficients on the interaction term between pre-election and vote margin (1994) for the period 1995 to 2007, covering only the first three elections. Again, the direct effect of vote margin on fiscal variables is absorbed by the province fixed effects estimator and therefore does not enter separately in the equation. The impact on both the total transfer and the equitable share is larger than in the full sample, suggesting that the effect is decreasing over time. In this

sense, the results confirm Brender and Drazen's (2005) finding on a subnational level, at least as far as intergovernmental transfers to South African provinces are concerned. Whether this is because of institutional improvements that constrain the ability to control the levers on the equitable share or because voters have become more experienced, reducing the incentive to do so is an interesting question for future research, particularly once election results on the fifth democratic election become available. Reassuringly, the results are robust to both vote margin measures; whether they are fixed to 1994 electoral outcomes or vary over time.

Table 4: The evolution of the pre-election vote margin effect over time

Dependent variable	Log of real per capita total transfer		Log of real per capita equitable share		Log of real per conditional grant	
	1995-2007		1995-2010		1995-2010	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-election × vote margin (1994)	0.00050 (0.00047)	0.00012 (0.00030)	0.00074** (0.00031)	0.00037** (0.00017)	-0.00127 (0.00228)	-0.00145 (0.00140)
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.72	0.87	0.67	0.79	0.74	0.84
Observations	117	144	117	144	117	144

Note: Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls are provincial GDP per capita, provincial population aged less than 14 years and provincial population aged 60 years or above. All regressions include a constant.

2.7 Conclusion

This chapter builds on the context-conditional political budget cycle literature to present a first analysis of subnational PBCs in South Africa. It demonstrates that the national government has both an incentive and the ability to implement political budget cycles in provincial revenues. In line with PBC theory and the core supporter model from the distributive politics literature, the empirical analysis suggests that less politically competitive provinces receive higher transfers from the central government. This increase is driven by the equitable share, which is the main component of total intergovernmental transfers. While the equitable share is determined by a formula, the government has discretion over two important levers: data revisions and formula revisions, the timing of which can be strategically manipulated to favour core support provinces in pre-election years. No presence of electorally motivated spending is found in the conditional grant. Overall, the results indicate that even in a dominant party system, electoral competition can trigger the implementation of PBCs.

With respect to South Africa's intergovernmental transfer system, these findings have important implications. The equitable share has been shown to be vulnerable to electorally motivated manipulations, in particular the targeting of less competitive provinces in pre-election years. This is possible despite the fact that the equitable share is determined by a formula and part of a comprehensive budget process. This substantiates existing qualitative evidence of the FFC's inability to ensure the equal division of revenues. In line with Brender and Drazen (2005), the findings do suggest that the impact of vote margin on the equitable share has

decreased over time, but the magnitude of the effect in the full sample is nonetheless disconcerting.

However, it is also clear that the South African budget system faces a number of other forms of misallocations, in the form of unauthorized payments, contracts without competitive bidding, and manipulation of tenders. While these issues are beyond the scope of this chapter, it should be noted that compared to the overall level of corruption in South Africa's financial management system, the distortion of the equitable share is admittedly small.¹⁵ From this perspective, reforms addressing the public procurement system require more immediate attention than the manipulation of the equitable share uncovered here. Nonetheless, it is important to point out that in addition to off-the-book malpractices, official channels may be subject to manipulation, too.

With respect to the existing political budget cycle literature, the results suggest that intergovernmental grant systems can function as potential channels through which the central government can distribute electorally motivated funds across regions. Empirical predictions from the distributive politics literature offer useful insights into the incentives that may govern the political decision to target specific groups of voters or regions. In this context, the empirical findings build on Miguel and Zaidi (2003) and Case (2001), who also find evidence of the core supporter model. By analysing the relationship between intergovernmental transfers and electoral competition in other settings, future studies can

¹⁵ The South African government is estimated to lose approximately USD 2.7 billion to procurement corruption each year (Corruption Watch 2013).

contribute to a better understanding of the dynamics between central and subnational governments and the role of strategic distributions of central government funds. The results presented here suggest that PBCs may be found in an environment with little or no electoral competition on the national level. The analysis of political budget cycles in countries that have previously been excluded from cross-country panels due to little or no electoral competition on the national level is thus encouraged. Indeed, more research should focus on the subnational level and disaggregated fiscal instruments, to identify fiscal manipulations that would otherwise remain undiscovered.

Chapter III

A better life for all? Democratization and electrification in post-apartheid South Africa

Abstract

Does democracy affect basic service delivery? If yes, which elements of democracy matter – enfranchisement, the liberalization of political organization, or both? In 1994, 19 million South Africans gained the right to vote. The previously banned African National Congress won the elections promising “a better life for all”. Using a difference-in-differences approach, we exploit heterogeneity in the share of newly enfranchised voters across municipalities to evaluate how franchise

extension affected household electrification. Our dataset combines geo-referenced nightlight satellite imagery, census data, and municipal election results. We control for a range of covariates and perform multiple robustness checks, including placebo regressions and an examination of contiguous census tracts from different municipalities. We find that enfranchisement has a significant positive effect on electrification, and that the national electricity company prioritized core constituencies of the ANC.

3.1 Introduction

Does democracy affect the delivery of essential basic services? And if yes, which elements of democracy trigger changes in implemented policies – enfranchisement, the liberalization of political organization, or both? It is crucial to assess the socio-economic consequences of democracy, not only because it is today's main form of government, but also because countries across the globe continue to transition towards democracy. A number of studies in the political economy literature find that democracies are better at providing public services than autocratic regimes (e.g., Lake and Baum 2001, Bueno de Mesquita et al. 2003, Diaz-Cayeros et al. 2013). A related literature suggests that the extension of the franchise can shift the median voter in a way that affects redistribution towards the poor (e.g., Meltzer and Richard 1981, Husted and Kenny 1997, Boix 2003) and the provision of services benefiting the newly enfranchised (Miller 2008, Vernby 2013).

However, several challenges make it difficult to quantitatively assess the direct effect of democratization on the lives of the poor. Many empirical

studies are based on cross-country comparisons from which it is hard to convincingly establish causality due to the possibilities of omitted variable bias, reverse causality, and sample selection bias (Ross 2006, Hollyer et al. 2011). Reliance on data collated by international organizations from diverse sources of uneven quality, and sometimes based on inconsistent definitions, also gives rise to concerns about measurement error. Other studies examine the impact of democracy on resource allocation rather than services delivered (Stasavage 2005). Yet, especially in poor countries, funds are often not spent as intended, which undermines service provision (Stasavage and Moyo 2000, Reinikka and Svensson 2004). These are some of the reasons why our knowledge about the effect of democratization on public service delivery is still limited.

An additional challenge is that the process of democratization typically entails a bundle of different changes, including franchise extension as well as the lifting of barriers to political activity and organization. Yet, existing empirical work, especially when based on standard indices of democracy (Munck and Verkuilen 2002), often leaves unresolved what precisely it is about democracy that accounts for the effect of interest. We argue that tackling such “compound treatment” problems requires a joint consideration of the effect of franchise extension via its effect on the median voter as well as an analysis of the mediating role of political parties in service provisions. The most common conjecture in the distributive politics literature, the “swing voter” hypothesis, posits that rational incumbents target swing regions in which they have a higher chance of affecting vote choice. In other words, “parties woo the groups that are politically central, and most willing to switch their votes in

response to economic favors" (Dixit and Londregan 1996: 1132). On the other hand, Cox and McCubbins' (1986) core supporter model predicts that incumbent governments allocate more resources to core support groups to maximize votes. In Sub-Saharan Africa, this has been linked to the notion of patronage, where an incumbent selectively distributes "favors and material benefits to loyal followers" (Bratton and van de Walle 1997: 62). This reward of loyalty and punishment of disloyalty is independent of political competition. Our approach highlights the essential complementarity of median voters and partisan accounts in order to interpret spatial patterns of service delivery following a democratic transition.

This chapter examines the effect of democratization on the provision of basic public services in post-apartheid South Africa. We estimate the effect of extending voting rights to the non-white population on electrification during the first period of their democratic local government. Our subnational approach helps to mitigate several of the concerns raised above. In particular, municipalities within South Africa are more comparable units than countries with different historical, cultural, and institutional circumstances, which should help to mitigate omitted variable bias. We use perhaps the most reliable data on service delivery, collected without involvement by the units we examine. Moreover, we directly measure actual service delivery outcomes, not resource allocation. Finally, by differentiating franchise extension and political representation, we take a step towards addressing "compound treatment" problems that afflict cross-national studies.

There are few prior studies of the political determinants of electricity provision. Brown and Mobarak (2009) document a positive effect of democratization on the relative share of electricity going to residential consumers in poorer countries. This cross-national study is vulnerable to the various threats to valid inference noted above, and it suffers from the “compound treatment” problem. Moreover, we are interested in whether democracy improves absolute levels of electricity access amongst disadvantaged groups. Most closely related to our work is the study by Min (2010), who traces partisan effects on electricity use with subnational data from India. However, unlike Min, we examine electrification in the context of democratization, involving franchise extension and the liberalization of political organizations. At the same time, we draw on Min’s pioneering work with satellite nighttime light imagery in our robustness checks, thus expanding our understanding of the use of these images in conjunction with other data.

South Africa’s post-apartheid experience in the 1990s provides a particularly good opportunity to evaluate the effect of democracy on service delivery. With two decades since South Africa’s first democratic election, we are examining a case of contemporary democratization on a unique scale: in 1994 suffrage was extended to about 19 million previously disenfranchised South Africans, representing 84 per cent of the estimated 22.7 million citizens and permanent residents aged 18 or above who had the right to vote (Southall 1994: 639). At the same time, the backlog in the provision of basic public services reflected the apartheid era’s political inequality. Two-thirds of households were estimated to lack access to electricity (African National Congress 1994:

para. 2.7.1). These conditions make for an ideal setting to investigate the effect of democratization on electrification. Our identification strategy exploits the heterogeneity in the share of the newly enfranchised population across South African municipalities following the unbanning of the African National Congress led by Nelson Mandela and the first fully democratic local government elections. Our quantitative estimates show that municipalities with larger shares of newly enfranchised voters, and thus a clear change in the identity of the median voter, experienced larger improvements in household access to electricity in the period 1996–2001.

We also show that the ANC's representation plays an important mediating role. In municipalities directly responsible for electricity distribution to households, our results are compatible with standard theories of electoral competition that predict policy convergence on the median voter (Downs 1957): in this case the effect of enfranchisement does not appear to be conditioned by the party political composition of the local council. In other parts of the country, the state-owned electricity company Eskom carried out electrification. Here, we find evidence that newly enfranchised voters in the ANC's core constituencies were prioritized. This suggests that the dominant party at the national level might have exerted pressure on the company to prioritize its base support, in line with theories of distributive politics that policy-makers reward their core constituents. Hence, a combination of median-voter and core-constituency accounts is reflected in the spatial patterns of service delivery that we observe. This finding makes a contribution towards clarifying the mechanism by which democratization affects the

provision of essential basic services: the effect of franchise extension is mediated by political parties, which control the levers available to deliver essential services. Our findings highlight the importance of pinpointing the actors who control service provisions, which is crucial for the interpretation of the results.

The remainder of this chapter is structured as follows. Section 3.2 provides the relevant background for our case and develops our expectations. Section 3.3 discusses the main variables, data sources, and related methodological issues. This is followed by a discussion of our empirical approach in section 3.4, and our analysis of the effect of franchise extension on household electrification in section 3.5. Section 3.6 investigates the conditioning role of political parties. The conclusion discusses the wider relevance of our results, and highlights several implications for further theoretical and empirical work on the relationship between democratization and public service delivery.

3.2 Enfranchisement, political parties, and electrification

South Africa’s “negotiated revolution” (Sparks 1995) culminated in democratic elections in April 1994, where for the first time all adult South Africans had the right to vote (Mattes 1995).¹⁶ The country’s mass electrification campaign was inherently linked to the transition to democracy, which had fundamental implications for the composition of the electorate, as well as South Africa’s party political landscape. In this section, we develop our expectations as to how these developments affected electrification patterns.

Enfranchisement. Under apartheid, formal political representation was mainly reserved for the white population. The Independent Electoral Commission estimated a total of 22.7 million eligible voters in 1994, consisting of South African citizens and permanent residents who were 18 years or older and who possessed at least one of five documents demonstrating their eligibility. Of these, 72 per cent were estimated to be “black African,” 9 per cent “coloured,”¹⁷ and 3 per cent “Indian or Asian.” The “white” population previously enfranchised in the territory of the

¹⁶ Black people were stripped of their South African citizenship in 1970. In 1982, the apartheid government established Black Local Authorities in South Africa, but they lacked resources and legitimacy. From 1984 to 1994, the Indian and coloured population groups had limited representation in separate chambers of the Tricameral Parliament, but this system lacked legitimacy and elections were widely boycotted. None of these pre-democracy forms of representation enabled credible participation and influence.

¹⁷ Under apartheid, people regarded to be of mixed descent were classified as “coloured” and distinguished from the other groups listed here, often with arbitrary rules such as the notorious “pencil test.” The continued use of the classification is widespread but not uncontroversial.

Republic of South Africa constituted a mere 16 per cent of the estimated electorate (Southall 1994: 637). This population classification remains in use to this day for census purposes and in public policy debates, and forms the basis for our analysis of franchise extension.

Results from the 1996 census indicate a striking correlation between these population groups and access to basic services, reflecting their unequal treatment by the apartheid regime. Merely 44 per cent of black and 84 per cent of coloured households used electricity for lighting (Statistics South Africa 2005: 145). As a result, the extension of the franchise to these historically disadvantaged groups also implied a substantial shift in the identity of the median voter, increasing demand for basic services such as electricity. Hence, we expect a positive effect of enfranchisement on electrification, particularly for black but also coloured households. In order to establish causality and not mere correlation, we need to overcome some important empirical concerns that will be addressed in the chapter.

Political parties. The transition to democracy also brought fundamental shifts in political party control. In 1994, the ANC won 62.6 per cent of the national vote in the first national elections after its unbanning, giving the party 252 out of 400 seats in the National Assembly. The party also won a majority of seats in six out of nine newly constituted provincial legislatures (Southall 1994). In the Western Cape, the previously ruling National Party (NP) won 53.2 per cent of the vote, while in the Northern Cape the ANC was the largest party but fell short of a majority with 49.7 per cent of the vote, and in KwaZulu-Natal the Zulu-nationalist Inkatha

Freedom Party (IFP) narrowly achieved victory with 50.3 per cent of the vote.

Following a transition period and the redemarcation of municipal boundaries, the first democratic elections of local councils took place in 1995 and 1996. For the first time in its history, South Africa had municipal councils elected under universal suffrage which were responsible for delivering essential services to all residents (Cameron 1996). According to the Elections Task Group (1996: 230-232), the ANC won 6,032 out of 11,368 seats on municipal councils, or about 53 per cent. While the party attracted the majority of black voters, these elections were not purely determined along racial lines (Mattes 1995). In several provinces, dominant population groups were divided across parties, notably coloured voters in the Western Cape and the Northern Cape, as well as the Zulu-speaking population in KwaZulu-Natal. Moreover, local elections combined proportional representation with wards in a way that biased representation in favour of local minorities and reduced proportionality (Cameron 1996: 30-31).

Shortly after its unbanning in 1990, the ANC adopted mass electrification as a central political goal. The National Electrification Conference in 1992 started a process that led to an agreement between Eskom, the country's state-owned electricity company, and the ANC over a set of electrification targets (Bekker et al. 2008: 3128).¹⁸ These became part of a National

¹⁸Eskom had initiated a more limited "Electrification for All" program in the late 1980s (Conradie and Messerschmidt 2000: 265-309). Previously, Eskom had not been permitted to supply black communities with electricity. In the empirical section, we use placebo regressions to investigate electrification trends prior to democratization.

Electrification Program (NEP) and the ANC's manifesto for the 1994 liberation elections, the Reconstruction and Development Program (RDP). The ANC promised "electricity for all" and pledged: "An accelerated and sustainable electrification programme must provide access to electricity for an additional 2.5 million households by the year 2000, thereby increasing the level of access to electricity to about 72 per cent of all households (double the present number)" (ANC 1994: para. 2.7.7). This high-level commitment meant that the ANC's performance would be measured in no small part by whether it would be able to achieve these ambitious targets.

At the local government level, the RDP remained a focal point of the ANC's political agenda. The party's manifesto for the municipal elections ("A Better Life: Let's Make It Happen Where We Live") highlighted that democratic local councils make crucial decisions on delivering basic services, including "where new electricity supplies...will be put in" (ANC 1995). In contrast, the Democratic Party (DP) and NP campaigns emphasized crime, which was perceived as "principally a 'white' preoccupation" (Lodge 1999: 44). The IFP had other priorities still, in particular the future role of traditional authorities, which provided an important power base in its rural strongholds in KwaZulu-Natal (Beall et al. 2005). Given the ANC's high-level policy commitment to electrification, we expect the party to have strong incentives to influence the connection of households to the grid. We explore different mechanisms further below.

Electrification. While the ANC adopted an overall electrification target, the 1996 constitution left responsibility for electricity reticulation with local authorities, which had provided electricity to households prior to democracy. It is important to note that electricity generation did not impose a constraint during this period. Eskom had a 55 per cent reserve margin in 1990 (Bekker et al. 2008: 3126). Yet, not all municipalities had the capacity to connect households to the grid. Hence, the NEP split responsibility between established municipal distributors, which were responsible for one-third of the new connections, while Eskom took on the task of electrification and direct distribution to households elsewhere, especially in rural areas (Gaunt 2005: 1310). In assessing the determinants of electrification during this period, it is essential to understand the motivations of the different actors.

In the mid-1990s, the regulatory authorities issued all 362 existing municipal distributors with temporary licenses (National Electricity Regulator 2000: 93). These municipal distributors were under the direct control of local governments and subject to their political direction. Councils played an important role in identifying priorities and needs (Qase et al. 2001). Municipal electricity departments had to seek funding from the council, clarify technical requirements, and plan implementation (Gaunt et al. 2012: 9-11). This ensured direct local accountability for electrification, since “communities can use their vote to choose their representatives and influence policies”, as highlighted in an official evaluation of the NEP (Department of Minerals and Energy 2001: 15).

For Eskom, strategic considerations were important. The company initially had to rely on internal financing for electrification, including cross-subsidies as well as debt (Bekker et al. 2008: 3131). It also faced pressure to meet the demanding electrification targets. Hence, the company had incentives to prioritize areas that could be electrified easily and at the lowest possible cost per household, which depends on factors such as an area's distance from the grid, settlement density, as well as topographical conditions such as land gradients (Dinkelman 2011: 3084). Many new household connections were not financially viable (Davis 1997). Without direct government subsidies during the initial electrification period, Eskom could contain operating losses by focusing on areas where potential consumption was comparatively high.

Gaunt (2003: 196) argues that "it would be contradictory to expect no political interference in essentially social programmes." The politicization of electrification by Eskom dates back to the apartheid period. According to Davis and Steyn (1998: 68), "Eskom has always been used to serve the interests of key constituencies behind the political party in power." In the 1970s and 1980s, the company (then known as Escom) connected white-owned commercial farms to the electricity grid, even though many of the connections were not financially viable. Under the ANC-led government, pressures to deliver to the party's "key constituencies" were also likely. In addition, Eskom managers had incentives to please the ruling party in order to delay a restructuring of the electricity industry (Dinkelman 2011: 3083). While the process by which Eskom internally prioritized areas for electrification is not formally documented, we can examine its outcomes.

This discussion leads us to anticipate different effects on electrification depending on whether distribution was in the hands of Eskom or local authorities. In the latter case, the link to the political composition of the local council and its administration is through representation. If parties at the local level converge on the median voter in a Downsian fashion, then service delivery to the newly enfranchised should not be affected by the partisan composition of a council. On the other hand, if the ANC was more committed to electrification than other parties, then it may matter whether the party gained control of a council. Hence, we examine such representation effects in the municipal distribution subsample. However, partisan effects in Eskom distribution areas could be rather different. Since Eskom was at least indirectly subject to pressure from the dominant party at the central government level, its selection of electrification projects would always be political and could potentially be used to target specific voters. Our empirical analysis investigates such party political influences.

3.3 Variables and data

The empirical analysis is based on a dataset of 799 municipalities in South Africa. As shown in Figure 4, these municipalities are located within nine provinces. All variables are structured in a municipality-province-year format. The census data are from Statistics South Africa's 1996 and 2001 Community Profiles and GIS databases. Shortly before the 2001 census, a new local government structure with 262 municipalities replaced the one that existed at the time of the 1996 census. This process involved the redrawing of boundaries as well as the renaming of areas, which helps to explain why past analyses of service delivery over this period have been

limited to comparisons at the provincial level, the lowest geographical hierarchy that has stayed constant over time (Statistics South Africa 2005). In order to make 1996 and 2001 local government data comparable, we took advantage of their spatial attributes and aggregated information on 21,243 so-called ‘subplaces’ from the 2001 census up to the level of 1996 municipal boundaries. This generated a unique dataset with 1996 and 2001 level data using the municipal boundaries that existed in 1996.¹⁹ One contribution of this chapter is precisely to offer to the research community a new longitudinal dataset of South African electoral and socio-demographic variables at the municipal level. Studies of service delivery in South Africa often resort to coarse data on just nine provinces (Statistics South Africa 2005). Our dataset provides a much finer representation of South African society at the time of transition to democracy and we hope other researchers will use it to explore questions beyond the central theme of this chapter. Appendix B provides detailed information about the process of generating this new dataset and definitions of all variables.

¹⁹ Although our 2001 municipal boundaries are “virtual boundaries” this is not a problem for our empirical analysis because we use them only in order to measure the outcome of interest (electrification) in 2001. Political decisions taken by councils elected in 1995/6 directly affected areas belonging to a municipality defined according to 1996 boundaries. Hence, reporting 2001 outcomes at 1996 boundaries is the correct procedure if we want to evaluate the impact of the 1995/6 elections on outcomes during the first period of South Africa’s democratic local government.

Figure 4: Overview of South Africa's census geography in 1996

Map 1: South Africa's nine provinces



Map 2: Africa reference map



Map 3: Municipal boundaries in Gauteng

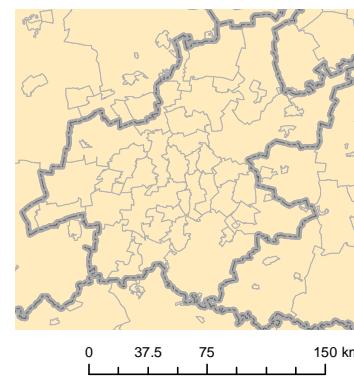
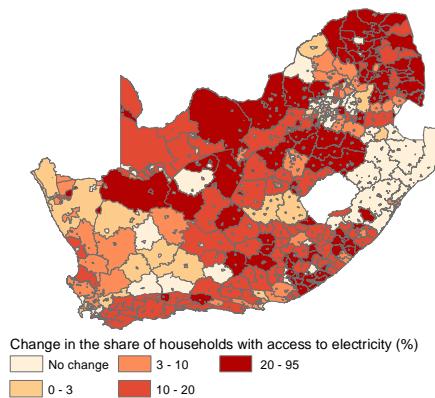
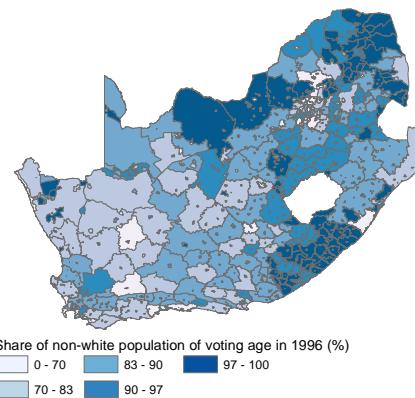


Figure 5: Geographic distribution of key variables

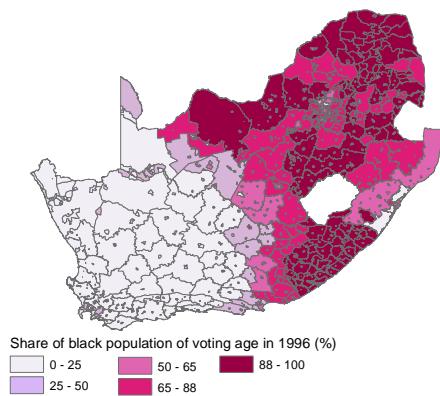
Map 1: Change in household access to electricity (1996-2001)



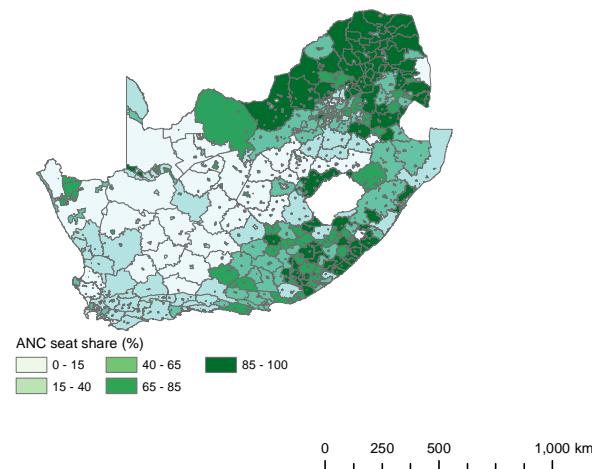
Map 2: Share of non-white enfranchisement (1996)



Map 3: Share of black enfranchisement (1996)



Map 4: ANC seat share on local councils (1996)



Dependent variables. Our main dependent variable is $Electricity_{it}$ measured as the percentage share of households with access to electricity, where i indexes a given municipality and t corresponds to time. This variable is based on the answer “electricity” to the census question: “What type of energy/fuel does this household MAINLY use for lighting?”²⁰ As non-technical losses such as theft or illegal connections are widespread, it is important to note that our dependent variable captures actual electricity use, not just connections and official users (Statistics South Africa 2005: 144). Map 1 in Figure 5 offers a visual impression of the change in Electricity between 1996 and 2001.²¹ Former “homeland” areas experienced some of the greatest changes in electrification. These areas, reserved for the black population under the previous regime’s policy of racial segregation, were synonymous with poverty and underdevelopment (Christopher 1994). At the same time, however, it is important to note that almost all municipalities across the country had

²⁰ We also have data on the type of energy used for cooking and heating. Both are highly correlated with lighting, but the latter is the most relevant indicator of whether or not a household has access to electricity (Statistics South Africa 2005: 144): “It should be noted that the census captures use of electricity, not access. In the case of heating and cooking, use may be limited for reasons of expense, both of the electricity itself and of appliances, so no conclusions can be drawn about access to electricity. The use of electricity for lighting, however, can be taken as a proxy for access to at least some level of electricity.”

²¹ Illegal connections might make it difficult to attribute electrification to Eskom or municipalities. Only the 1996 census distinguishes between municipal and Eskom connections from other connections. However, less than 2 percent of households indicated they get electricity from another source, which is likely to be the reason why this category was dropped from the 2001 census. Due to the presence of illegal connections we cannot assume that there were no connections other than those provided by a municipality or Eskom, but the findings in the 1996 census suggest that the electricity variables in both censuses reflect whether or not the main source of electricity is provided by a municipality or Eskom and thus whether or not households are connected to the grid. We can therefore reasonably assume that changes in electricity between 1996 and 2001 are due to Eskom and municipalities, either through new connections or newly used old connections.

electrification backlogs, and most of them had substantial numbers of households that were not connected to the grid. According to our calculations with the 1996 census data, merely seven municipalities had universal household access to electricity. The average (median) share of municipal households with access to electricity was 63 (70) per cent. In 1996, half of the municipalities had household electrification rates of 70 per cent or lower, and merely 132 (17 per cent of the total) had rates of 90 per cent or above. Hence, sizable improvements in household electrification were possible in the vast majority of South Africa's municipalities. Nonetheless, our analysis needs to account for possible ceiling effects. As we do not expect large changes in municipalities that already have high rates of household electrification, one of our robustness checks entails dropping those municipalities from our analysis. We try different thresholds to demonstrate that our findings are robust.

As an alternative dependent variable, we use $Nightlight_{it}$, which is calculated using a set of satellite images of the earth at night. The images were collected with the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) and released by the National Geophysical Data Center (NGDC) for the years 1992-2012. We use the annual composites of cloud-free, visible, and stable lights that are cleaned to filter out fires and background noise to calculate Nightlight for the years 1992, 1996, and 2001.²² Figure 6 visualizes how we transform this

²² Gas flares are visible in nighttime light images, which could be a concern. However, no gas flares have been identified on land in South Africa (email correspondence with Chris Elvidge, June 29, 2012).

data: we first use the raw nighttime lights images (Map 1) to classify each of the over one million pixels as either “lit” or “unlit” (Map 2), and then calculate the percentage share of lit pixels at the municipal level within the 1996 boundaries to yield information on local electrification levels (Map 3) as well as changes between 1996 and 2001 (Map 4). The data appendix contains full details.

An alternative method is to combine nighttime lights with population rasters to calculate the share of population in lit pixels (see Elvidge et al. 2010). While this variable is conceptually closer to our census-based measure, it relies heavily on the accuracy of population rasters.²³ In our case, daytime population counts at the local level differ substantially from census data due the legacy of the Group Areas Act, which under apartheid forced non-whites to live in designated areas often requiring long commutes to work (Christopher 1994). Moreover, these data are not available for the time period examined here. We prefer a methodology that only uses the nighttime images.

What is the relationship between the two dependent variables? Only the census measure directly captures our outcome of interest, household access to electricity, while the satellite images provide an aerial view of pixels that appear lit at night. On the municipal level the average

²³ Main sources for population rasters are the US Department of Energy’s Landscan data and gridded population datasets from the Center for International Earth Science Information Network (CIESIN). The former are based on daytime population counts using daily averages that capture populations on roads and other public areas. They are available to us only from 2002 onwards. The latter has a resolution of 2.5' and is thus less fine-grained than the nighttime lights data.

correlation between Nightlight and Electricity (in levels) for the years 1996 and 2001 is 0.47.²⁴ This is broadly consistent with other comparisons of nighttime lights to national electricity data (e.g., Elvidge et al. 2001, Min et al. 2013).²⁵ In the absence of stable outdoor lighting, pixels appear unlit even where households have access to electricity. This matters in our context, since street lighting was “often not provided” when households were connected to the grid (Department of Minerals and Energy 2001: viii). Even where outdoor lighting is present, it may be too weak to be detected.²⁶ Moreover, Nightlight captures light from uninhabited industrial sites and highways. South Africa’s mines and main highways are clearly visible from outer space. In areas around Johannesburg and Cape Town we also observe the overglow effect, due to the tendency of light “to travel to pixels outside of those in which it originates, and light tends to be magnified over certain terrain types such as water and snow cover” (Pinkovskiy 2011: 9). The satellite data might still be preferable were the census of poor quality. However, the quality of South Africa’s official statistics is recognized as uniquely high on the continent (Jerven 2013: 101). For our purposes the census is the ideal data source, but we exploit the satellite data to assess the robustness of our results and to analyze years for which census data are not available.

²⁴ This correlation increases for smaller subnational units, suggesting that the nighttime lights measure is more precise in smaller geographic areas.

²⁵ Recent contributions in the economics literature include the use of nighttime lights as a proxy for economic activity (Henderson et al. 2011, 2012; Storeygard 2012; Pinkovskiy 2011), regional development (Michalopoulos and Papaioannou 2012), and spatial inequality (Alesina et al. 2012).

²⁶ In his analysis of nighttime lights in India, Min (2014) points out the lowest population density at which lights are observed are 60 people per square kilometer.

Independent variable. The main independent variable is the percentage share of enfranchised citizens in the municipal electorate. The variable $Enfranchised_i$ refers to all non-white citizens of voting age in municipality i at the time of the first democratic local government elections. It is derived from a cross-tabulation of population group and age from the 1996 census. The census also allows us to break down this variable further into separate groups of newly enfranchised voters that are categorized as black, coloured, and Indian (or Asian). Maps 2 and 3 in Figure 6 show, respectively, the distribution of all newly enfranchised and black voters across municipalities.

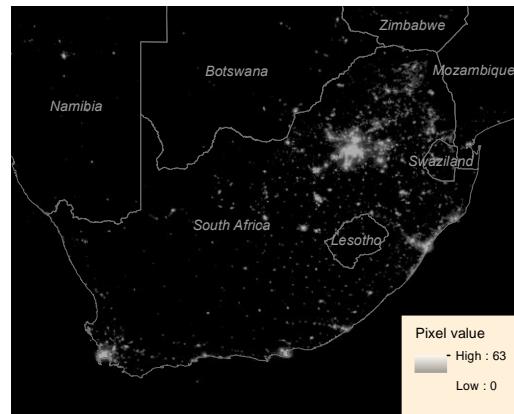
Control variables. In order to address concerns that Enfranchised is correlated with other variables that are in turn correlated with trends in Electricity, we include a number of control variables. We control for geographic covariates that were important in determining the supply of electricity: the distance to the main power grid in 1996, the distance to the closest main road, gradient/slope, and elevation. We control for the demand for electricity by including a number of socioeconomic variables, such as employment and education, as well as demographic variables, such as population density and the municipal share of the national population.²⁷ Finally, our regressions also include province fixed effects. Given that the dependent variable is expressed in differences, they capture province-specific shocks and absorb fixed spatial autocorrelation.

²⁷In order to avoid “post-treatment bias” (King 2010), we prefer specifications without “bad controls” (Angrist and Pischke 2009: 64), that is, control variables that are themselves potential outcomes affected by Enfranchised, such as changes in other public services that occurred after democratization.

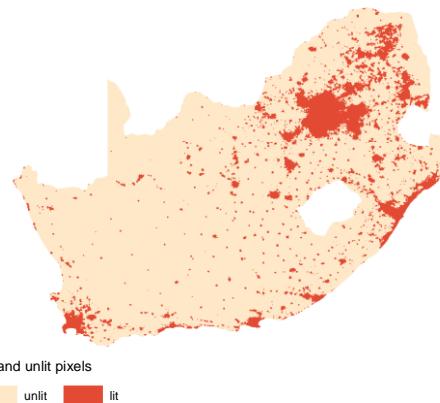
Appendix Table B1 reports summary statistics. As shown in panel (a), the average change in the share of electrified households is approximately 9 percentage points according to the census, whereas our data based on nighttime lights show an average change of only 2 percentage points in the share of lit pixels. Comparing these means in panels (b) and (c), which show summary statistics for our variables in municipalities with enfranchisement below and above the median respectively, suggests that there is a systematic discrepancy between the two measures, which is greater in high enfranchisement areas. At the same time, it is clear that high and low enfranchisement municipalities differ according to most other indicators. Our empirical strategy removes fixed municipal characteristics but we still need to worry about possible differential trends.

Figure 6: Nighttime lights as a proxy for municipal household electrification rates

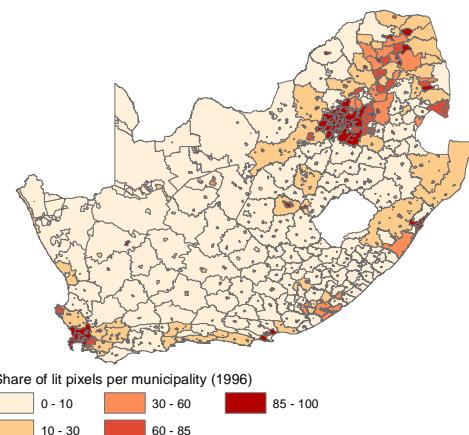
Map 1: Raw nighttime lights raster image (1996)



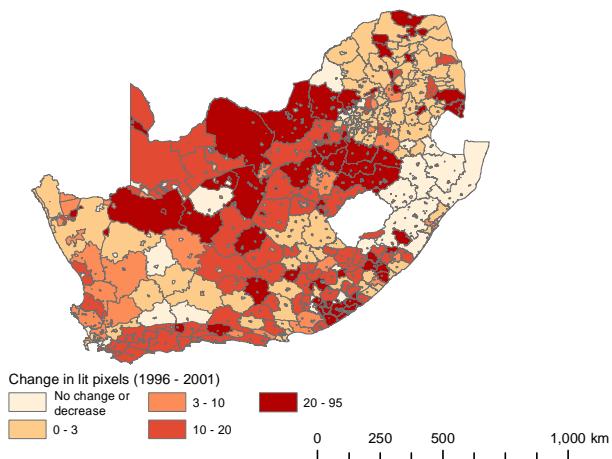
Map 2: Reclassified nighttime lights raster image (1996)



Map 3: Share of lit pixels per municipality (1996)



Map 4: Change in the share of lit pixels (1996-2001)



3.4 Empirical strategy

In an ideal setting, we would like to compare our data with a counterfactual scenario in which democratization did not occur. This comparison would allow us to determine precisely what difference democracy made in terms of delivering electricity to households. Since we cannot observe this counterfactual scenario we exploit the heterogeneity in enfranchisement in the first democratic elections across South African municipalities. Under apartheid, non-white citizens had no meaningful right to vote and the first fully democratic local elections took place in 1995/6. Municipalities with a high share of newly enfranchised voters can therefore be viewed as being part of the treatment group. Conversely we can think of municipalities with low levels of enfranchisement as being part of the control group. While this categorization helps to conceptualize the empirical strategy, Enfranchised is a continuous rather than binary treatment. This allows us to determine the effect of enfranchisement on electrification by using a difference-in-differences specification. The model can be expressed as follows, with i indexing a given municipality:

$$\Delta Electricity_{i1996-2001} = \alpha + \beta Enfranchised_{i1996} + \gamma Controls_{i1996} + \Delta \varepsilon_{i1996-2001} \quad (2)$$

Electricity corresponds to the proportion of households with access to electricity, Enfranchised corresponds to the share of non-white citizens of voting age in the first democratic local government elections and ε is the error term. There are a number of potential confounders that affect both the level of enfranchisement and electrification. We control for a wide range of such variables that we discussed above.

By using a specification in changes rather than levels, we purge our regressions of omitted unobservables that are fixed over time. However, we still remain concerned about changing characteristics of the municipalities that could be correlated with changes in electrification via Enfranchised. Our main identifying assumption is that, conditional on the observables we control for in our regressions, Enfranchised is uncorrelated with the error term. We run placebo regressions to rule out as far as possible that our results are driven by a violation of the parallel trends assumption. A paramount concern is that electrification in areas with high enfranchisement could simply constitute a catching-up effect, due to the fact that municipalities with high electrification rates in 1996 have little margins for further electrification. These municipalities are also those with higher white populations and therefore lower enfranchisement rates. Placebo regressions should help to rule out this possibility, since it is not obvious that a catching-up effect, if not due to democratization, should only occur after the first democratic elections. We document several additional robustness checks below.

3.5 Main results

Baseline estimates. Table 5 reports our baseline estimates, using census-based measure as the dependent variable and enfranchisement rates as the main explanatory variable. We report the results from several specifications, starting with a simple regression (column 1) and progressively including province-specific shocks (column 2), geographic controls (column 3), population and socioeconomic controls at their 1996 levels (column 4), and non-electrified households in 1996 (column 5). We

then include 1996-2001 differences in population and socioeconomic controls (column 6). This introduces the possibility of post-treatment bias (Angrist and Pischke 2009: 64, King 2010), but failing to control for these differences may also result in bias. The coefficient of interest is positive and statistically significant throughout. We can safely conclude that the share of newly enfranchised non-white voters has a positive and statistically significant effect on household electrification rates. The magnitude of the coefficients ranges from 0.19 to 0.34. The lower bound suggests that one standard deviation in enfranchisement led to an increase in electrification of at least 3.5 percentage points, more than one third of the sample mean: a substantively large effect.

In column 7 we distinguish between newly enfranchised black, coloured and Indian voters. The results show that the effect is particularly strong for black voters, positive but less strong for coloured voters, and statistically insignificant for newly enfranchised Indian voters. The coefficients indicate that an increase of one standard deviation in the share of black voters led to an increase of 12 percentage points in the household electrification rate, and about half that size for coloured voters. We address several empirical concerns in the remainder of this section.

Placebo regressions. The most important identification concern is that our estimates could be capturing changes that would have happened anyway in the absence of democratization. Our identification strategy relies on the “parallel trends” assumption (Angrist and Pischke 2009: 230-233): the idea that, once other covariates are taken into account, the various municipalities had similar pre-existing electrification trends,

irrespectively of their level of enfranchisement. Hence, we would like to rule out that our results could be driven by pre-existing trends associated with Enfranchised. To address this concern we follow the standard procedure to run placebo regressions that check whether enfranchisement rates can explain electrification before democratization (when they should not).

Comparable information for the period prior to 1996 does not exist, as apartheid-era census data from the early 1990s are unreliable and incomplete. For this reason we use Nightlight as an alternative dependent variable. This has two important advantages: first, it is based on an entirely independent data source with complete records of the distribution of electricity across South Africa. As such, the images function as an additional source of data that we use to corroborate our findings. Second, the satellite images are publicly available as annual composites from 1992, which means that they can provide information about electrification in the years prior to democratization.

Table 5: The impact of enfranchisement on electrification (census data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Enfranchised	0.343*** (0.032)	0.280*** (0.032)	0.293*** (0.033)	0.232*** (0.041)	0.191*** (0.040)	0.277*** (0.040)	
Enfranchised black							0.327*** (0.041)
Enfranchised coloured							0.204*** (0.047)
Enfranchised Indian							-0.022 (0.059)
Constant	-19.058*** (2.534)	-11.056*** (3.001)	-6.992* (3.619)	-5.269 (4.399)	-2.204 (4.328)	-12.815*** (4.757)	-13.817*** (4.565)
Province fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes	Yes	Yes	Yes
Population and socioeconomic controls (1996)	No	No	No	Yes	Yes	Yes	Yes
Households without electricity (1996)	No	No	No	No	Yes	Yes	Yes
Population and socioeconomic controls (1996-2001 diff.)	No	No	No	No	No	Yes	Yes
R-squared	0.111	0.161	0.173	0.192	0.245	0.307	0.325

Note: The dependent variable is the percentage share of households with electricity for lighting (difference 1996-2001) calculated from census data. Geographic controls are: (1) Distance from electricity grid; (2) Distance from main road; (3) Elevation; (4) Slope. Population controls are: (1) Population density; (2) Number of households. Socioeconomic controls are: (1) Share of population with no schooling; (2) Median income; (3) Share of labor force with low income (due to differences in the underlying variables in the 2001 census, this variable is only included as a 1996 level control and not as a 1996-2001 difference). Refer to the data appendix for full details. N = 799. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We begin by replicating the models reported in Table 6, this time using the change in the share of lit pixels over the 1996-2001 period as our dependent variable. The results appear in Table 6, panel (a). The coefficients of interest are statistically significant and very stable across the different specifications, but magnitudes are smaller than in our regressions with census data: a one standard deviation in the share of non-white new voters leads to a maximum 1.6 percentage point increase in the share of lit pixels. It is reassuring that the estimated effects of enfranchisement relative to mean changes in electrification are similar across models using census and satellite data.

In panel (b) of Table 6 we report the results of our placebo regressions. Here, we replace the dependent variable with changes in the share of lit pixels during 1992-1996. In this earlier period, Enfranchised should not matter if the “parallel trends” assumption holds. Reassuringly, all coefficients on Enfranchised are far from any acceptable significance level. We also experimented with a range of alternative nightlight-based variables used in the literature, notably the population-adjusted measures discussed earlier. The overall pattern is highly robust with significant enfranchisement effects for the period 1996-2001 and insignificant enfranchisement effects in the placebo regressions for the period 1992-1996. Despite the limitations of the Nightlight-based measure, these results strongly suggest that what we capture with our estimates is unlikely to be due to pre-existing trends in electrification.

Examining contiguous census tracts. The main identification concern is that the need for electrification at the municipal level is highly correlated

with enfranchisement (both being correlated with the share of non-white population at the time of democratization). It is therefore difficult to disentangle the effect of empowering the non-white population from a “catching up” effect. Placebo regressions help to rule out this second possibility. In this section we further corroborate our results by using an alternative empirical strategy based on spatial discontinuity. We use more fine-grained data at the census tract (henceforth CT) level and match adjacent CT pairs that lie on different sides of a municipal boundary, thus restricting the sample to CTs that lie on municipality borders only.²⁸ By including a fixed effect for each pair of neighbouring CTs, identification is obtained by matching CTs that belong to different municipalities (hence treated with differential levels of enfranchisement) but that are adjacent (hence generally similar in other respects).²⁹ By using this strategy we should be able to take into account several confounding factors. First, it is unlikely that economic and social conditions vary discontinuously along municipality borders, which makes comparisons more reliable. Second, electricity needs and socioeconomic conditions may be specific to a CT while the relevant decision-making unit is the municipality: municipal-level enfranchisement should now better capture the effect of democratization rather than a “catching up” effect. We estimate the following equation (omitting years from subscripts):

²⁸ This empirical strategy is described in detail in Holmes (1998), and has been used and extended by Dube et al. (2010) and Duranton et al. (2011).

²⁹ A CT bordering more than one CT of a bordering municipality enters multiple times into the sample, each time with a separate pair fixed effect. To correct for the resulting correlations across pairs on the same municipality boundary, we use two-way clustering (Cameron et al 2008), by municipality and by each border between municipalities.

$$\Delta Electricity_{cpi} = \alpha_p + \beta Enfranchised_i + \gamma Controls_c + \varepsilon_{cpi} \quad (3)$$

As before, i indicates a municipality, c indicates a CT and p indicates a pair of adjacent CTs that lie across a municipality boundary. Note that the controls are now at CT level but Enfranchised remains at the municipal level. Results are reported in Table 6 and are similar to those obtained with our difference-in-differences specification. The coefficient of Enfranchised is always statistically significant at 1% level and ranges from 0.17 to 0.28, slightly lower than what we obtained previously. When we distinguish between ethnic groups, the coefficient of black enfranchised remains positive, statistically significant and with a magnitude almost identical to that of Table 5. The coefficient of enfranchised coloured becomes instead statistically insignificant while that of enfranchised Indians becomes significant: these are the only relevant differences compared to our main specification. Overall, our results are robust to this quite demanding specification, both in terms of magnitude and statistical significance of the coefficients of interest.

Other robustness checks. As discussed in section 3.3, another concern is that our dependent variable has an upper bound of 100. Highly electrified areas in the pre-treatment period have lower margins to improve, and since pre-treatment electrification levels are certainly not orthogonal to enfranchisement, this may bias the coefficients of interest. Therefore we repeated our exercise with restricted samples, excluding municipalities where more than 90 per cent of the population had access to electricity in 1996, and progressively bringing this threshold down to 40 per cent (in 10-point intervals). Using the same specification as in

column 5 of Table 6, the coefficient on Enfranchised is about twice as large as in the full sample. While electrification gains were lower in areas with very high levels of household access to electricity in 1996, the basic pattern is very robust. These results appear in Appendix Table B2, panel (a).

Further robustness checks reported in Appendix Table B3, panel (a), show that the pattern of results does not change if we exclude, in turn, the municipalities that fall into any one of the nine provinces: the relationship we document is not limited to any particular region of the country.

Table 6: The impact of enfranchisement on electrification (satellite data, with placebo regressions)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>a. Dependent variable: ΔNightlight 1996-2001</i>							
Enfranchised	0.064*** (0.023)	0.063*** (0.024)	0.066*** (0.025)	0.092*** (0.034)	0.088** (0.035)	0.080** (0.037)	
Enfranchised black							0.109*** (0.038)
Enfranchised coloured							0.041 (0.044)
Enfranchised Indian							0.015 (0.049)
<i>b. Dependent variable: ΔNightlight 1992-1996</i>							
Enfranchised	0.068 (0.052)	0.029 (0.059)	0.037 (0.056)	0.047 (0.063)	0.059 (0.064)	-0.006 (0.070)	
Enfranchised black							0.007 (0.072)
Enfranchised coloured							-0.026 (0.078)
Enfranchised Indian							-0.078 (0.082)
Province fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes	Yes	Yes	Yes
Pop. and socioeconomic controls (1996)	No	No	No	Yes	Yes	Yes	Yes
Households without electricity (1996)	No	No	No	No	Yes	Yes	Yes
Pop. and socioeconomic controls (1996-2001 diff.)	No	No	No	No	No	Yes	Yes

Note: The dependent variable is the percentage share of lit pixels (difference 1996-2001 and 1992-1996, respectively) calculated from satellite data. All regressions also include a constant. Refer to Table 5 for a description of control variables, and the data appendix for full details. N = 799. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: The impact of enfranchisement on electrification (contiguous census tract pairs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Enfranchised	0.276*** (0.052)	0.261*** (0.056)	0.257*** (0.055)	0.188*** (0.055)	0.169*** (0.054)	0.192*** (0.055)	
Enfranchised black							0.260*** (0.063)
Enfranchised coloured							-0.020 (0.081)
Enfranchised Indian							0.188*** (0.074)
Contiguous CTs Pair fixed effects	Yes						
Province fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes	Yes	Yes	Yes
Population and socioeconomic controls (1996)	No	No	No	Yes	Yes	Yes	Yes
Households without electricity (1996)	No	No	No	No	Yes	Yes	Yes
Population and socioeconomic controls (1996-2001 diff.)	No	No	No	No	No	Yes	Yes
R-squared (overall)	0.084	0.062	0.038	0.074	0.083	0.089	0.103
R-squared (within)	0.017	0.019	0.027	0.069	0.082	0.088	0.095

Note: The dependent variable is the percentage share of households with electricity for lighting (difference 1996-2001) calculated from census data. Refer to Table 5 for a description of control variables, and the data appendix for full details. All variables are calculated at the CT level, except Enfranchised and Enfranchised black/coloured/Indian, which are calculated at the municipality level. N = 7528. Standard errors are double clustered (see Cameron et al. 2011) at the municipality and the border level. There are 687 clusters for municipalities and 1172 clusters for borders. *** p<0.01, ** p<0.05, * p<0.1.

3.6 The mediating role of political parties

This section focuses on how party politics affected electrification during this period. As noted earlier, partisan effects could differ depending on whether Eskom or a local distributor was in charge of electrification. In municipalities served by Eskom, electricity distribution could be part of a national strategy of rewarding core supporters or swing voters. On the other hand, local distributors could be more responsive to changes in the median voter (irrespective of party representation at the local level). We use 1996 membership data from the Association of Municipal Electricity Undertakings (AMEU) to distinguish the two groups. Appendix B contains a detailed description.

Our statistical analysis draws on the variable $ANC\ seat\ share_i$, which is the share of total seats on local council i won by the ANC in the 1995/6 elections (Elections Task Group 1996). This was the first time following its unbanning that the ANC was able to fully contest municipal elections. Hence, the share of seats obtained in those elections at the same time represents the change in the share of ANC seats from zero prior to democracy. The variable is highly heterogeneous and ranges from zero to 100 per cent with a mean of 55 and a standard deviation of 31. Map 4 in Figure 5 shows the geographic distribution, with spatial clusters of ANC seats between 76 and 100 per cent in the Eastern Cape, North West, and Limpopo (the latter was initially called Northern Transvaal and then Northern Province). These provinces contain the ANC's core constituencies and coincide with the areas that experienced the greatest changes in electrification as depicted in Map 1.

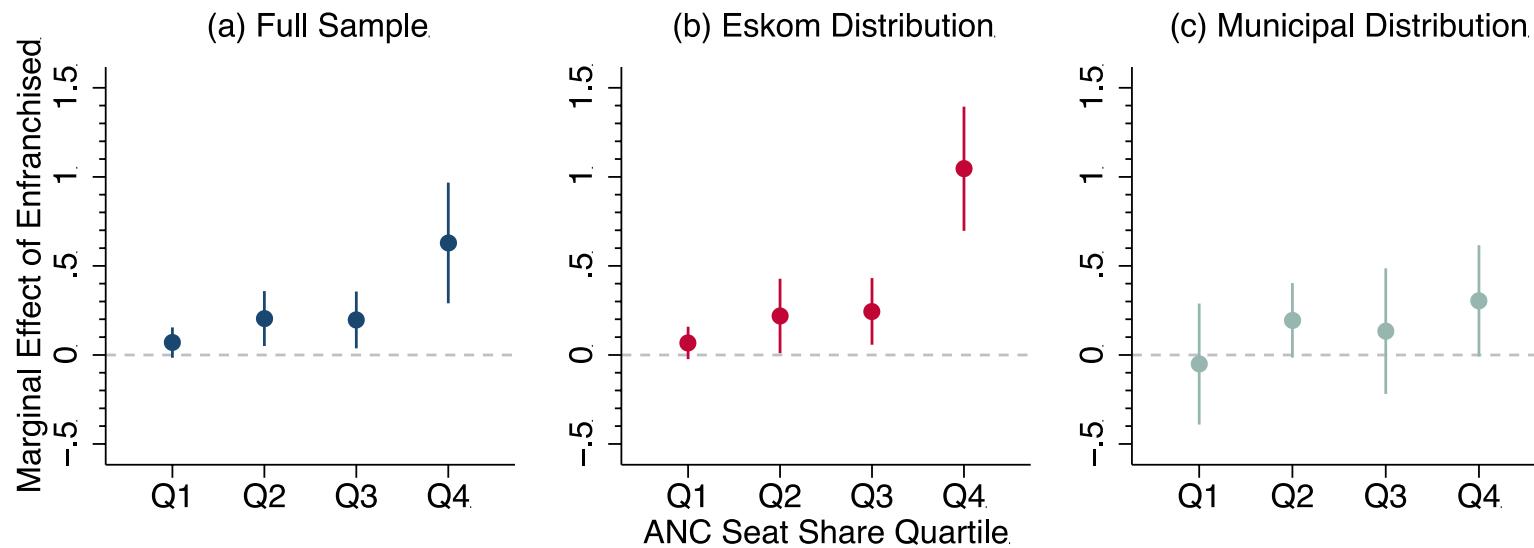
We first test whether enfranchisement is merely a proxy for the changed landscape in municipal representation, or put differently, if the impact of democratization on electrification is channelled via municipal representation of the ANC. Hence, we start by augmenting our model with the seat share variable. Column (1) of Table 6 shows that the ANC's seat share had no direct effect on electrification outcomes and that enfranchisement, which remains positive and statically significant, is therefore not merely a proxy for ANC representation. Only when we exclude enfranchisement, the effect of ANC representation on electrification is not statistically distinguishable from zero (column 2). We also find no evidence of a direct effect of ANC seat share when we distinguish Eskom (column 3) and municipal distribution (column 4) areas. Moreover, the coefficient on Enfranchised is the same in both subsamples.

If not via municipal representation, another possibility is that the impact of enfranchisement itself differed as a function of ANC strength. To test this, we augment our model with an interaction between Enfranchised and ANC seat share. We report results for the full sample (column 5) and separately for Eskom (column 6) and municipal distributors (column 7).³⁰ The coefficient on the interaction term is positive and significant only in the full sample and the Eskom subsample. The magnitude of the coefficient is larger in the latter case. To probe the precise nature of the partisan effect, we construct separate dummies for quartiles of ANC seat

³⁰ When we replaced the ANC's seat share with NP or IFP seat shares, the results convey a similar story in different ways: the effect of enfranchisement is statistically significant only when these parties had low levels of representation on local councils, and it is decreasing in their representation. See Appendix Figure B3.

share. We then use these dummies and their interactions with enfranchisement in our regressions (the first quartile is the reference category), replacing the continuous seat share measure and its interaction. The results in columns (8) to (10) and in Figure 7 reveal important nuances. First, the 50 percent threshold does not matter; the conditional coefficients for the second and third quartiles are not statistically distinguishable. Second, F-tests indicate that in Eskom-served municipalities the marginal effect of franchise extension when the party controls 75 percent of seats or more is statistically different from the other conditional coefficients. If it simply were the case that the coordination between a non-ANC council and Eskom was more difficult, then it should matter whether the ANC had a majority or not, irrespective of its size, but this is not the pattern we detect. For councils with municipal distributors, the effect of enfranchisement is still larger in the three highest quartiles of ANC representation, compared to the baseline category, but the differences are smaller and statistically insignificant.

Figure 7: Enfranchisement conditional on ANC seat share quartile



Note: Graphs (a), (b), and (c) are based on the results in columns (8), (9), and (10) of Table 8, respectively.

We also conduct a formal test of whether the corresponding coefficients depicted in panels (b) and (c) of Figure 7 are statistically different. Using the full sample, we estimate a model with a three-way interaction of enfranchisement, the seat share quartile dummies, and an indicator of whether a municipality is supplied directly by Eskom.³¹ The coefficient on the interaction between Enfranchised, ANC seat share Q4, and the Eskom indicator is .655 (the difference between .953 and .297 in columns 9 and 10 of Table 8), with a standard error of .258 ($p = .011$). For all other seat share quartiles, the corresponding differences between the two subsamples are not statistically distinguishable. In other words, Eskom delivered a significant top-up to core constituencies of the ANC, in both a substantive as well as a statistical sense.

These patterns are robust to different combinations of controls included in Tables 2 and 3; we do not report these results. Supplementary results in appendix Table B2, panel (b) show that pattern is also evident in the restricted samples introduced earlier to account for a possible ceiling effect. Table B3, panel (b), further shows that the interactive pattern is robust to excluding municipalities in any one of the nine provinces.

In sum, we detect two distinct patterns of service delivery, depending on the assignment of institutional responsibility for electrification. In municipal distribution areas, the partisan composition of local councils

³¹ To recover the precise coefficients reported in columns 9 and 10 of Table 8, we also allowed the coefficients on all controls to vary across the two subsamples. The results reported here are substantively the same whether we do this or not.

does not condition the effect of enfranchisement. This implies that, when decision-making happened at the local level, party convergence was more likely in line with standard models of electoral competition (Downs 1957). The results for Eskom distribution areas, on the other hand, point to a strong role for the partisan composition of municipal assemblies and appear compatible with core-voter models in the distributive politics literature (e.g., Cox and McCubbins 1986). Here, the pattern of results suggests that the dominant party in the national government, via Eskom, rewarded its “key constituencies”, as predicted by Davis and Steyn (1998: 68). Since Eskom was responsible for two-thirds of the electrification target under the NEP, this effect dominates in the full sample.³²

³² Without the subsample analysis, the average pattern of results could have been interpreted as evidence against convergence on the median voter while our results are compatible with this hypothesis. This highlights the central importance of understanding the precise delivery mechanisms for the proper interpretation of the spatial patterns we document.

Table 8: The role of the ANC's seat share on local councils

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enfranchised	0.194*** (0.043)		0.216*** (0.050)	0.115 (0.086)	0.053 (0.044)	0.065 (0.048)	-0.005 (0.143)	0.080* (0.044)	0.082* (0.047)	-0.082 (0.168)
ANC seat share	-0.005 (0.026)	0.050** (0.025)	-0.011 (0.032)	-0.005 (0.051)	-0.370*** (0.086)	-0.491*** (0.114)	-0.187 (0.153)			
Enfranchised × ANC seat share					0.004*** (0.001)	0.006*** (0.001)	0.002 (0.002)			
Enfranchised × ANC seat share Q2							0.125 (0.083)	0.174 (0.107)	0.224 (0.187)	
Enfranchised × ANC seat share Q3							0.165* (0.087)	0.211** (0.103)	0.167 (0.199)	
Enfranchised × ANC seat share Q4							0.511*** (0.159)	0.953*** (0.173)	0.297 (0.194)	
ANC seat share Q2							-12.401** (6.211)	-17.632** (8.236)	-15.247 (13.649)	
ANC seat share Q3							-12.898* (6.836)	-16.789** (8.373)	-9.201 (14.527)	
ANC seat share Q4							-44.830*** (14.614)	-86.916*** (15.934)	-20.830 (14.458)	
Observations	799	799	551	248	799	551	248	799	551	248
Distributors in the sample	All	All	Eskom	Municipal	All	Eskom	Municipal	All	Eskom	Municipal
R-squared	0.245	0.227	0.240	0.311	0.262	0.264	0.314	0.266	0.277	0.318

Note: The dependent variable is the percentage share of households with electricity for lighting (difference 1996-2001) calculated from census data. All regressions include a constant, province fixed effects, geographic controls, population and socioeconomic controls (1996), and households without electricity (1996). Refer to Table 5 for a description of control variables, and the data appendix for full details. The pattern of results is not affected when we vary the combination of controls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.7 Conclusion

We find that enfranchisement increased household electrification rates during the first period of the democratic local government in South Africa. For the period 1996-2001, we estimate an average increase in the share of households with electricity access across municipalities of between 3 and 6 percentage points per standard deviation of enfranchisement. When distinguishing different groups of newly enfranchised voters, we find that the effect of enfranchisement on electrification is largest in municipalities with higher shares of black voters. Our placebo regressions, in which we proxy for electrification by using nighttime lights satellite imagery, provide evidence that this finding is unlikely to be due to pre-existing trends. We conduct further robustness checks that return the same pattern of results. With regard to the question we raised at the outset of this chapter, our evidence suggests that in the case of South Africa, democracy did indeed affect the delivery of electricity, as enfranchisement shifted the median voter.

Our analysis also shows that the liberalization of political organization mediated the effect of enfranchisement. While enfranchisement had a positive effect on electrification across municipalities, ANC core constituencies supplied by Eskom saw even larger gains in this initial period of the post-apartheid electrification campaign. In other words, Eskom delivered an ANC top-up in areas where it distributed electricity. Conversely, the partisan composition of the local council made no difference in areas with municipal distributors controlled by local councils, suggesting convergence on the median voter. Hence, it appears that all political parties tried to deliver electrification when in power at

the local level, but the ANC had its hands on one crucial additional lever: Eskom. These findings highlight the importance of pinpointing how political parties can affect service provision in order to understand observed spatial patterns of service delivery.

Our results contribute to the literature on democracy and public services in several ways. Much of the literature is cross-national and uses comparative indices of democracy that make it difficult to pin down what precisely affects public service provision. Our analysis is sub-national and establishes an analytically separate contribution of franchise extension and changes in partisan representation; not only does it matter that more people receive the right to vote, but under certain conditions it also matters which party they vote for. This is a valuable first step in addressing the “compound treatment” problem in cross-national work on this topic.

Moreover, by directly measuring service delivery outcomes our analysis focuses on the ultimate outcome of interest: whether people’s lives were actually affected. This is essential for assessing the implications of democracy on the poor when resource allocation is a limited indicator for actual service delivery, as is the case in many developing countries. Our empirical analysis uses two independent data sources to assess electrification outcomes. This is an advance over studies based on data collected with involvement by the very governmental units that are under examination. By using both census data and satellite images, our work is also among the first in documenting their relationship. We demonstrate potential for cross-validation but also limitations. More

research is needed to analyse how precisely these alternative measures relate to one another in different contexts.

With respect to the South African context of this chapter, we recognize that our scope is narrowly defined: we focus on the initial period of electrification after the demise of apartheid. Various institutional adjustments, including the redrawing of municipal boundaries, and changes to the electricity sector, make it difficult to trace effects for a more extended period. While, on balance, electricity access continued to improve, disconnections due to non-payment became a widespread phenomenon that undermined access (Fjeldstad 2004). More recently, electricity blackouts occurred as a sustained lack of investment in generation capacity became evident. It is important to point out that these problems were not an inevitable consequence of the initial electrification campaign that we examine, but rather the result of inadequate planning (Johnson 2009: 473-481).

How generalizable are our results? On the one hand, the South African context was unusual in that the conditions for a rapid rollout of electricity existed at the time of the transition to democracy. Due to a massive reserve margin, electricity generation was not an obstacle to expansion. Post-democratization electrification gains might be less impressive in countries with lower reserve margins. Related work on how democracy affects health outcomes in Sub-Saharan Africa suggests that the basic pattern that we document for electricity applies more widely to other services (Kudamatsu 2012). We plan detailed follow-up work on access to drinking water and housing, which have their own complexities.

Nonetheless, we have reasons to believe that the patterns we document apply not only to other services, but also widely across countries. Recent work on enfranchisement in very different geographic and historical contexts (e.g., Miller 2008, Vernby 2013) gives external validation to our finding that enfranchisement matters for service delivery. We contribute to this literature by drawing attention to the potentially crucial mediating role of political parties and how they affect spatial patterns of service delivery in response to franchise extension.

Chapter IV

How can we study socioeconomic outcomes from outer space? Opportunities and pitfalls of using nighttime lights data in quantitative social science research

Abstract

How can we study socioeconomic outcomes from outer space? This chapter makes a methodological contribution to answering this question by providing the first review of nighttime lights applications from a social science perspective. It highlights the possibilities for using these data to proxy for a range of socioeconomic outcomes, from electrification, to economic activity, and educational attainment. More importantly, I demonstrate that there are significant limitations using the data and that

a strong positive correlation between nighttime lights and a given socioeconomic outcome on the national level does not warrant their application on the subnational level. Both Type I and Type II errors are likely to arise. The chapter also draws attention to important differences between the ways in which we can turn the information stored in the satellite images into proxies for socioeconomic outcomes. Crucially, the results we get can differ greatly, depending on which method is chosen.

4.1 Introduction

In 1960, the National Aeronautics and Space Administration (NASA) of the United States launched its first experimental weather satellite. Initially, the aim was to obtain global meteorological data, but it soon became clear that the images obtained by the satellite provided a unique perspective of human activity from outer space. Today, the images present an opportunity to obtain independent, globally consistent and complete proxies for a range of socioeconomic outcomes such as electricity provision, economic growth, greenhouse gas emissions, or urbanisation. This value proposition is appealing to social scientists engaging in quantitative research as they depend on consistent, high-quality data, which are often difficult to obtain. While data collection efforts by governments and other organizations have increased, availability remains a problem, particularly for those hoping to study socioeconomic outcomes in developing countries. But even in the developed world, high local precision is difficult to achieve in studies that transcend national borders. For example, with access limited to traditional data sources, it would hardly be possible to compare socioeconomic outcomes in an area of precisely 20 square kilometres on

either side of a national border. Similarly, we would not be able to answer questions that require the measurement of rural electrification rates across countries in sub-Saharan Africa, energy consumption levels in China, or urbanisation trends in Indian villages.

It is thus not surprising that the use of the satellite images is starting to spill over from weather forecasting to the social sciences. Taking advantage of the revolution in remote sensing technology and advances in geospatial analysis, a few social scientists with an empirical research agenda have started using the images to proxy for a range of socioeconomic outcomes in various country contexts, such as measuring electrification in India (Min 2010), economic development in Africa (Michalopoulos and Papaioannou 2013), or human well-being across the globe (Ghosh et al. 2013). The global breadth and at the same time local accuracy of these data opens up avenues for exploring new research questions. The investigation of contemporary differences in well-being across ethnic groups within 17 Sub-Saharan countries in Alesina et al. (2012) is another fitting example. While social science applications of the images are still in their infancy, the remote sensing literature is well established and offers a wealth of insights into the many possible applications of the data. This chapter provides the first review of the existing literature from a social science perspective, thereby highlighting the potential of this technological advance to study human development.

More importantly, however, the aim of this chapter is to draw attention to three main risks associated with any application of the images in quantitative social science research: first, even though several scholars

have shown high correlations between satellite-based and census-based measures of socioeconomic outcomes on the global level, there are large outliers in each of these studies. We cannot therefore assume that the images work equally well in all countries of the world. Second, even if the correlation coefficient between satellite-based and census-based measures is high at the national level, this does not mean that the correlation is equally high on the subnational level. Third, there are important differences between the ways in which we can turn the information stored in the satellite images into proxies for socioeconomic outcomes. The results we get can differ greatly, depending on which method is chosen. This risk is amplified when researchers refer to the variables with vague concepts such ‘luminosity’ or ‘light density,’ even though this can easily lead to interpretations that are removed from what the variable captures in reality.

Unfortunately, these risks are easily overlooked given the substantial advantages of using the data. I want to highlight at the outset of this chapter that data extracted from the images can never directly capture socioeconomic outcomes. They can only serve as proxies, which by definition makes them imperfect. Access to electricity is a useful example to consider: in an ideal world we would be able to directly capture whether someone has access to electricity or not. In the absence of such a perfect variable, there are at least two types of data sources available. On the one hand most countries collect censuses or administrative data, for example on household electricity usage typically recorded or electricity sales per customer. On the other hand, nighttime lights images are available online and capture the geographic location of manmade

lighting that is persistent enough to be picked up by a satellite in outer space. Both data sources are discussed in more detail below, the point I want to make here is that we can expect there to be great differences between what these two data sources can tell us about the true outcome of interest.

To do so, the discussion focuses predominantly on electricity because this is arguably the most direct way of interpreting the images: after all, each pixel reflects the presence or absence of anthropogenic light. At the same time, electricity access is often considered a “marker” of development (Dinkelman 2011: 3078), making it a relevant outcome of concern, be it from a public administration, policy, or economics angle. To demonstrate how the discussion extends to other social science applications, I also look at income and educational attainment, given their importance for studying human development.

In order to highlight the aforementioned risks, I reproduce the results presented in Elvidge et al. (2010), to demonstrate variation in cross-national correlations between census-based and satellite-based electrification measures. I also draw attention to large outliers in their study that are easily overlooked. To demonstrate that cross-country results are not necessarily portable to the subnational level, I take advantage of South African census data to cross-validate the satellite-based measures.

There are two main reasons why South Africa is a particularly useful setting for this exercise: first, the massive scale of South Africa’s

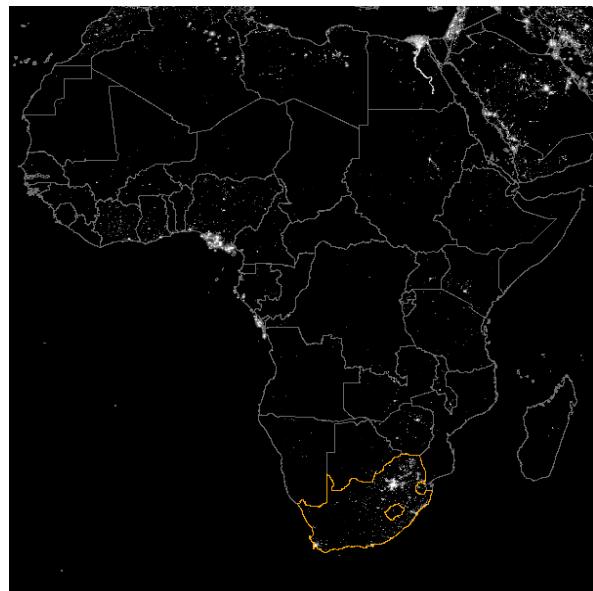
electrification programme in the 1990s makes for a fertile testing ground for assessing the extent to which both different levels of and changes in socioeconomic outcomes are observable from outer space. Following South Africa's transition from apartheid to democracy, over 10 million people are estimated to have gained access to electricity. These changes are well documented in traditional data sources and should be clearly visible from outer space. Second, the availability of high quality census data offers a rare opportunity for cross-validating the satellite images. During the two decades for which nighttime lights data are currently available (1992-2012), South Africa has had three censuses (1996, 2001 and 2011). Since even most industrialised countries only carry out one census per decade, the frequency of the data presents a strong comparative advantage of South Africa. More importantly, coupled with its exceptionally high quality compared to most other countries in Sub-Saharan Africa (Jerven 2013: 101).

The remainder of this chapter is structured as follows: first, I provide a social science perspective on nighttime light images by reviewing the existing literature, including both remote sensing and social science contributions. I then offer an overview of underlying data sources and detailed definitions of variables in section 4.3. On this basis, I cross-validate satellite-based measures of socioeconomic indicators with census-based measures on the cross-country level as well as with data from South African municipalities. Section 4.5 categorizes the main pitfalls to explain why such discrepancies might arise and demonstrates ways to navigate them. Section 4.6 concludes.

4.2 A social science perspective on nighttime light satellite imagery

When looking at a satellite image of Africa by night, at least three distinct hubs of brightness stand out (see Map 1, Figure 8): the densely populated Nile Delta in Egypt stretching south of Cairo, the Niger Delta Basin in Nigeria, where most of the country's oil fields are located and the area around Johannesburg in South Africa, which generates approximately one tenth of the GDP of the entire continent. The rest of the continent appears almost completely dark, particularly in comparison to the shining lights of Western Europe by night (see Map 2, Figure 8). Densely populated countries such as the Netherlands or Germany display high levels of brightness, consistently and nearly ubiquitously. Zooming in on the geographic extent of South Africa reveals a much darker image, which likely reflects not only a lower population density, but also lower access to electricity (see Map 3, Figure 8). Nonetheless, major urban areas, such as the Cape Town Peninsula in the south-west and Port Elizabeth, Pietermaritzburg, and Durban along the southern coastline are clearly visible.

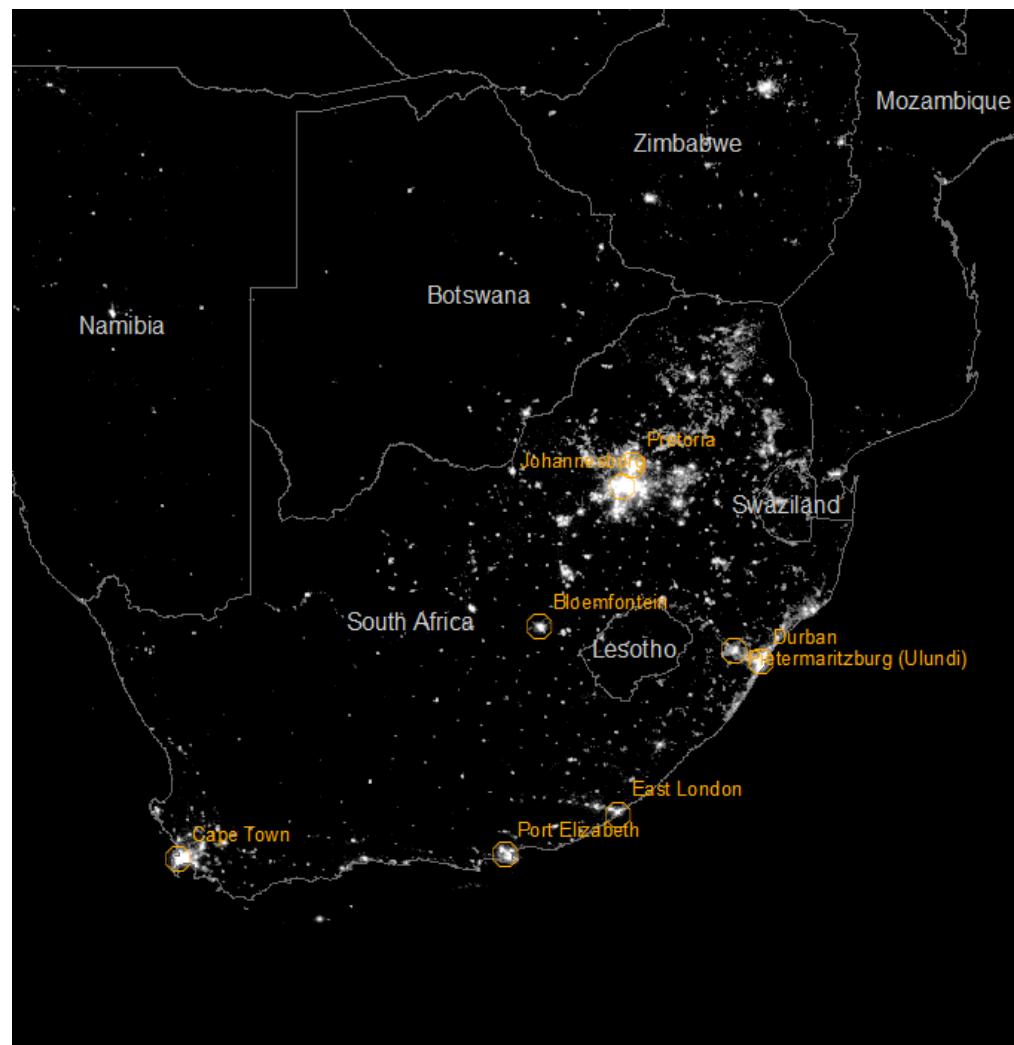
Figure 8: Views of the African continent, Western Europe and South Africa by night (2001)



Map 1: Darkness on the African continent



Map 2: Western Europe shining bright



Map 3: South Africa's eight biggest cities by night

Note: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by US Air Force Weather Agency. Above maps use images acquired with satellite F14 in 2001.

It seems intuitive that at the most basic level these striking images would be able to inform us about electrification: where we see lights, there must be electricity. Indeed, the images directly capture the existence of stable, electric lights as seen from outer space. But this was not the original intention. When NASA launched satellite TIROS-I in 1960 the mission was to provide global, operational meteorological data to meet military commitments (Air Weather Service 1974: ii). When the images became publicly available in the early 1970s – by then recorded with the succeeding Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS)³³ – the group of scientists using the data extended from US military scientists to the global community of meteorologists. While the early literature thus focused predominantly on the use of the images for weather forecasting purposes, it soon became clear that the images provided a unique perspective of human activity as observed from outer space (see Croft 1978: 1979).

When the National Oceanic and Atmospheric Administration (NOAA) in Boulder, Colorado established a digital archive in 1992, a burgeoning literature focussing on empirical applications of the images emerged. Today researchers working with nighttime light images generally fall into one of two groups: scientists in the fields of remote sensing on the one hand and quantitative social scientists on the other hand. While the two groups do not have a common research agenda, they share an interest in making use of the images. Research by the latter group draws heavily on the insights and findings produced by the former, particularly

³³ This is now a US Department of Defense programme run by the Air Force Space and Missile Systems Center.

the wealth of papers produced by Christopher Elvidge and his team at NOAA's Earth Observation Group (EOG), who process and release the data. This makes the remote sensing literature a natural starting point for gaining a social science perspective on the data.

Insights from the remote sensing literature. Much of the early work by the scientists at the EOG is descriptive in nature, clarifying the functionality of the satellite sensors and basic applications (e.g., see Elvidge et al. 1997, Imhoff et al. 1997, Sutton et al. 1997). More recent work has focused on the use of nighttime lights as a proxy for socioeconomic indicators, thereby establishing the relevance of the images for social science applications. Elvidge et al. (2010) document a strong positive relationship between nighttime lights and electrification rates across 229 countries and more than 2,000 subnational units. Their work presents the first systematic global assessment of electrification rates, which paved the way for subsequent social science applications. By overlaying lights and population rasters – a method that is further discussed below – they estimate that 1.62 billion people worldwide lacked access to electricity in 2006, compared to 1.58 billion estimated by the International Energy Agency (IEA). The resulting correlation coefficient between national level nighttime lights and IEA data is 0.90, suggesting a strong positive relationship between the two. The paper concludes that “while there are some known sources of error in the current product, the method does provide electrification rates using a standardized definition and standardized data sources, with complete global coverage.” As I demonstrate below, large outliers are hidden behind the high correlation

reported in this paper. I will therefore revisit this pioneering study throughout this chapter.

Other scientists have confirmed the high correlation between satellite-based and census-based measures of electrification. For example, Townsend and Bruce (2010) reinforce the correlation between satellite-based and census-based measures of electricity in their analysis of the spatial distribution of electricity consumption in Australia. For the period 1997 to 2002, they show that there was a very high correlation between state electricity consumption and nighttime lights with a correlation coefficient of 0.93 at the state and territory spatial resolution. Chand et al. (2009) show that in India between 1993 and 2002 changes in electricity consumption can be traced using nighttime light satellite imagery. Over this period, they find that a population increase of 170 million was accompanied by an increase in power consumption by 261,393 billion kWh. Nighttime light imagery mirrored this change with an increase in the number of nighttime lights of up to 26 per cent in all states. From these studies we may infer that there is a high correlation between electricity and nighttime lights – as per the definitions used in each of these studies and taking outliers into account – but as I demonstrate in this chapter the relationship on the subnational level is a lot less robust.

Besides electrification, the remote sensing literature has applied nighttime light images to study economic activity: Doll et al. (2006) present a global map of GDP at a one-degree resolution using nighttime light imagery. They also compare satellite-based measures of income growth with GDP statistics for 46 countries. Building on this research

Ghosh et al. (2010a) construct a map of total economic activity in which they include both formal and informal economic activity. They argue that the map they generate provides an alternative means for measuring global economic activity, which for the first time includes the informal economy. In more recent work Ghosh et al. (2013) provide a review of efforts to use the nighttime lights images, not only to measure GDP, but also human well-being more generally. The possibility to proxy for economic growth has attracted particular interest from the economics strand of the social science literature. The pioneering study by Henderson et al. (2011) is further discussed below.

In the spirit of capturing human development more generally, Elvidge et al. (2012) develop a Night Light Development Index, measuring human development across the globe with satellite images. Their index correlates relatively strongly with the Human Development Index, with an R^2 of 0.71. This offers an interesting possibility for social scientists to enrich their datasets with an index that is a “simple, objective, spatially explicit and globally available empirical measurement of human development derived solely from nighttime satellite imagery and population density” (Elvidge et al. 2012: 24). However, any such application will need to be carefully tailored to a given country and temporal context.

Doll’s (2008) guide to the use of nighttime lights provides a useful overview of other interdisciplinary applications, such as urbanisation. He makes reference to earlier research in which he correlates areas that appear lit in the images with census based population figures to trace urban extents in the US. Building on this work, Sutton (2003) determines

the level of urban sprawl in US cities by comparing their radiance thresholds. More recent work compares regional and global urban growth in India, China, Japan, and the US, measures the dynamics of urbanization in India (Pandey et al. 2013), and contrasts patterns of urban expansion in Colombia, Ecuador, Peru, and Bolivia between 1992 and 2009 (Alvarez-Berrios et al. 2013). Agnew et al. (2008) show how the images can be used in a more qualitative way: they evaluate the US military surge in Iraq by assessing the extent to which the distribution of nighttime lights in Baghdad changed between March 2006 and December 2007.

There have also been a number of studies in the area of disaster management, for example analysing the extent of light disruption in Mississippi and Louisiana following hurricane Katrina in September 2005 (Doll 2008: 29). Another interesting application of nighttime lights data is the approximation of greenhouse gas emissions. On a global level there is a strong positive correlation between nighttime lights and carbon dioxide (CO₂) emissions. This relationship was first identified by Elvidge et al. (1997) and then built on by Doll et al. (2000), who map global CO₂ emissions at a resolution of one square kilometre. Ghosh et al. (2010b) make the distribution of carbon dioxide emissions visible on a global map. The possibility to map and monitor emissions and identify specific sources, such as gas flaring, make them an indispensable resource for climate change researchers and policy makers. In fact, the World Bank is leading a global gas flaring reduction initiative in partnership with the NOAA, who provide estimates of global volumes of gas flaring using nighttime light satellite imagery. The unbiased and independent nature

of the images render them particularly useful for research on countries, such as Russia or China, where the objectiveness of administrative data may be called into question.

The aforementioned papers point to the breadth of possible social science applications of nighttime light imagery: from proxying for electrification, to economic activity, urbanisation to disaster management, the images provide a distinct view of the world from outer space. It is important to note, however, that most of the remote sensing literature has established robust correlations between the lights and socioeconomic indicators on a global level. Indeed, one of the main advantages of the remotely sensed images is to obtain information that is universally consistent, which makes it possible to establish proxies or indices that help us better understand socioeconomic development across the globe. To a certain extent, this is at odds with recent developments in the social science literature. Due to the high risk of endogeneity in cross-section data, an increasing number of empirical studies analyses socio-economic outcomes on the subnational level. Each application of nighttime lights in a subnational context requires careful consideration.

Social science applications. The economics literature was possibly the first social science branch to make use of the images to proxy for different variations of economic development. Henderson et al. (2012) use the images to measure economic growth from outer space. Published in the American Economic Review, this paper is generally considered the pioneering study in this literature, making it a common reference in subsequent studies. The authors develop a composite measure with

roughly equal weights on census-based and satellite-based growth to augment official income statistics in poor countries. They find that their composite measure differs from official data by up to three percentage points annually. They also take advantage of the high resolution of the images by calculating economic growth measures for subnational units in Sub-Saharan Africa.

Michalopoulos and Papaioannou (2013) and Alesina et al. (2012) follow the study in using nighttime lights data as a proxy. The former paper uses the satellite-based measure as a proxy for regional development. The authors argue that the use of what they refer to as 'luminosity' builds directly on contributions such as Elvidge et al. (1997) and Henderson et al. (2012) in the literature showing that "light density at night is a robust proxy of economic activity" (Michalopoulos and Papaioannou 2013: 120). They perform cross-validation between what they term 'light density' or 'luminosity' – by which they refer to the variable they calculate using the digital number that is contained in each pixel – and a census-based wealth index in Nigeria, Tanzania, the Democratic Republic of Congo, and Zimbabwe. They find a correlation coefficient of around 0.75 per cent, which is robust to the exclusion of the top 1 per cent of lit areas. Their use of an online appendix to provide additional cross-validation sets a good example for how studies that do not focus primarily on the images can still provide justification to support their use of the images. Notably, the cross-validation they provide is not at the same geographical hierarchy as the nighttime lights data they use in their empirical analysis: they validate the data across countries and regions within countries. Their

main units of analysis are, however, 10 square kilometre areas and the pixel level.

Alesina et al. (2012) focus on spatial inequality in their application of the lights. Specifically, the authors proxy for the level of development in ethnic homelands with what they term ‘average luminosity per capita’ – again, the term luminosity refers to the digital number contained in each pixel. To generate the per capita measures, they draw on population data stemming from the Gridded Population of the World dataset. There are important drawbacks when combining nighttime lights with this dataset; these are addressed below. Nonetheless, they use this measure to generate a spatial Gini coefficient with which they measure spatial inequality in ethnic homelands. Acknowledging likely problems of measurement error, they construct an alternative spatial Gini coefficient based on Thiessen polygons and include both in their empirical specifications. Since nighttime lights data is only one of multiple data sources in their study, they devote less than one page to explaining the data, with a justification that is symptomatic of other social science applications: “since comparable data on income per capita at the ethnicity level across all countries in the world do not exist, following Henderson, Storeygard, and Weil (2012) ... we use satellite image data on light density at night as a proxy” (Alesina et al. 2012).

As one of the first political scientists, Min (2014) provides further evidence of the validity of nightlights data as a measure of electrification. He uses the images to study the effect of electoral accountability on public goods provision and the distribution of power to the poor in India.

Recognizing the lack of robust evidence for the relationship between nighttime lights and electricity for smaller subnational units, Min makes an effort to validate the data in the Indian context by correlating village level electricity consumption, as well as the percentage of households using electricity for lighting, with what he terms nighttime light output – finding relatively strong coefficients of 0.82 and 0.70 respectively.

Unlike the remote sensing literature, social science papers have a tendency to use the images as one of many data sources rather than a primary subject.³⁴ As mentioned above, they also tend to use the data on the subnational level, whereas the remote sensing literature predominantly focuses on global applications of the images. Appendix Table C1 provides an overview of selected empirical applications in both literatures in descending order. While the list of papers in panel (b) are only a very small selection of the available remote sensing literature, panel (a) includes the main social science applications in the fields of economics and political science that have been produced to date.

Chen and Nordhaus (2010: 1) claim that “to date, virtually all studies have used the nighttime luminosity data without comparing them with other measures.” I would argue that this is an exaggeration, but their claim does reflect a general lack of sufficient cross-validation when using the data in social science applications. That is precisely why the use of nighttime lights alongside South African census data in chapter 3 makes an important contribution to the growing body of social science applications. However, it is important to recognize that oftentimes the

³⁴ Henderson et al. (2011) is an exception.

reason for using the images in the first place is precisely the fact that no alternative data sources are available. However, at closer inspection the images produce indicators of socioeconomic outcomes that are very noisy. In this context, Chen and Nordhaus (2010: 3) argue that “luminosity data has little value added in countries where sophisticated statistical information is available because compared to nighttime lights, measurement errors in standard economic data are relatively small.” The value added of the images is higher in countries where no data on socioeconomic indicators are available, but at the same time, this does not warrant their use without any attempts for cross-validation.

4.3 Turning nighttime lights into proxies for socioeconomic outcomes

In order to make the best use of the information in the images and extract proxies for socioeconomic outcomes, we have to understand the underlying data sources. For those unfamiliar with remote sensing data, Appendix C lays the groundwork for using and interpreting them appropriately by providing definitions not only nighttime lights as a data source, but also gridded population data, which are often used in combination with the images. The main point is that these are very different data sources compared to census data, which is why we can expect differences in empirical applications. It is important to keep the distinctions between nighttime lights, gridded population and census data in mind, when turning the data into proxies for socioeconomic outcomes. This section shows how.

While there are a number of ways to turn the digital number included in each pixel into a proxy for socioeconomic outcomes, existing social science applications have tended not to justify why they perform this step in a particular way. Instead, they simply refer to the data as ‘luminosity’ or ‘light density’ leaving the reader to wonder what precisely the measure actually means. In order to shed some light on the way in which we can use the data, the following shows that there are at least three distinct ways in which we can convert the information stored in each pixel into proxies for socioeconomic outcomes. This is an important exercise because the variable choice we make can affect the results we get.

Digital number. Arguably the most straightforward way of using the information contained in the nighttime light images is to determine the pixel value at the centre, the geographical centroid of a settlement and use that information to draw conclusions about the availability of electricity there. Alternatively, it is possible to add up pixel values for a desired level of geographic aggregation. This calculation uses the boundaries of a shape file to define the zones for which the digital numbers of each pixel are summed up.³⁵ In the literature, the resulting value is often termed luminosity (e.g., Michalopoulos and Papaioannou 2013). Based on this method, the average digital number can be calculated by dividing the sum of all pixels by the number of pixels in a given area. This continuous measure is in fact the average of an average because the digital numbers in the images are averages of daily, cloud-free observations over a given year. Michalopoulos and Papaioannou (2013) use this method to calculate the average digital number at various different levels of aggregation as a proxy for ethnic inequality.

However, one issue to be aware of when using this variable is that there are clusters of very brightly-lit pixels in main cities of developed countries. This problem is referred to as top-coding in the literature and arises when satellite sensors become saturated (Chen and Nordhaus 2010: 7). This is less of a problem in poorer countries, and especially in sub-Saharan Africa, where almost no pixels are top-coded (Storeygard 2012: 9). In South Africa, only 0.16 per cent of pixels had a value of 63 in 1992

³⁵ In ArcMap, this calculation can be performed with the ‘zonal statistics as a table’ tool, using the nighttime lights image as the input value raster and the shape file with boundaries as the feature zone to define the boundaries for which the data are to be calculated.

and while this figure has increased drastically in recent years, the number of top-coded pixels was still less than 0.5 per cent in 2012. If a researcher studies an area with a high share of top-coded pixels, such as New York City, the *digital number* is unlikely to be the appropriate variable.

Share of lit pixels. By reclassifying pixels it is possible to reduce the containing information into a binary code, indicating whether a given pixel is lit or not. To do so, pixel values between 1 and 63 are reclassified to take the value 1, indicating the presence of light, and pixels with value 0 retain the same value, indicating the absence of light. Pixels with value 1 are termed lit pixels, whereas pixels with value 0 are referred to as unlit pixels. This binary measure captures the presence or absence of light, rather than the intensity. For example, if a settlement grows brighter over the years, due to the introduction of more and brighter outdoor lighting, the share of lit pixels remains unchanged if this is not also associated with a geographic expansion. The share of lit pixels does not take the luminosity information provided in the nighttime lights into account. Crucially, this makes the variable less sensitive to variations in the satellite's recording instruments or changes in atmospheric conditions over time (Min 2010: 16). The importance of this feature is further discussed below.

The share of lit pixels in a given area is calculated as the sum of all lit pixels, divided by the total number of pixels in that area. Alternatively, this measure too can be determined for the centroid of a settlement, thus determining whether a given place is electrified at its centre or not. Indeed, this is the primary dependent variable used in Min's study of

electrification in rural India (2010). As above, the appropriateness of the centroid approach crucially depends on the size of the settlement and how straightforward it is to determine its centre.

Share of population living in lit pixels. The above measures electrification in terms of the presence of light in a pixel. Social scientists are, however, likely to be more interested in whether or not people in a given square kilometre have access to electricity, rather than whether or not a given square kilometre appears lit from outer space. One way of adjusting the share of lit pixels is to combine it with information about population. The share of populated lit pixels can be calculated as the sum of pixels that are both lit and populated, divided by the total number of populated pixels. Alternatively, rather than using binary information about the presence of people in a given pixel, it is possible to use the population counts rasters to determine how many people live in pixels that are lit. Dividing this figure by the total number of pixels equals the share of population living in lit pixels. This is the method used in Elvidge et al. (2010). Based on the definition of this measure, we may expect that this variable is closest to the census.

Share of households with electricity. The census based electrification rate for a given geographic boundary is defined as the share of households with electricity. Based on this definition, it is calculated as the share of households in a given area answering that electricity is the type of energy mainly used for lighting. Subsequent references to census-based electrification measures are based on this definition.

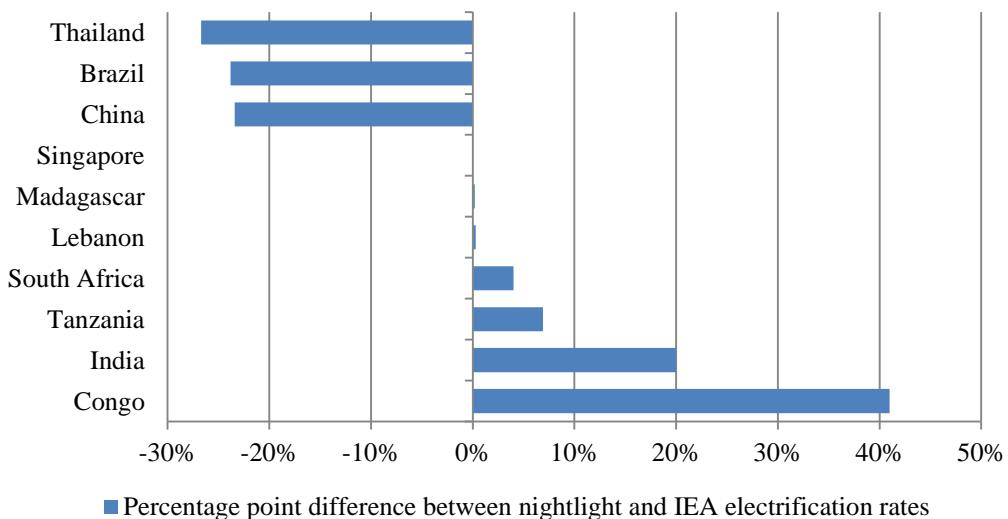
Each of the four aforementioned variables could potentially serve as a proxy for a socioeconomic outcome. However, given their distinct definitions, they are unlikely to give us the same answers. The following section puts each of these variables to the test in order to determine how they relate to each other and specifically how the satellite-based measures can help us study socioeconomic outcomes from outer space.

4.4 Cross-validation with cross-country and South African census data

Even though several scholars have shown high correlations between satellite-based and census-based measures of socioeconomic outcomes on the global level, we cannot assume that the images can serve as a reliable proxy for socioeconomic indicators in all contexts. Upon closer inspection, the findings presented in Elvidge et al. (2010) provide an excellent example: their analysis of global electrification rates overestimates that of the International Energy Agency by 2.5 per cent. This figure seems insubstantial on a global level. However, when comparing their estimated national electrification rates with data from the IEA, there are large differences in many cases. As illustrated in Figure 9, the satellite-based estimates are 41 and 23 percentage points larger than the IEA estimates in the Congo and India respectively. Conversely, the IEA estimates are 24 and 27 percentage points larger than the satellite-based estimates in Brazil and Thailand. This means that when Elvidge et al. (2010: 7) report that the total number of people found to be without electricity is only slightly larger than the figure estimated by the IEA, it is

important to understand that their estimate includes both under and overestimations of the IEA figures.

Figure 9: Showcasing differences between satellite and census based electrification rates



Source: Author's calculations with data from Elvidge et al. (2010).

In fact, the absolute difference between the national estimates is much greater than the 40 million that a naive interpretation of Elvidge et al.'s result suggests. In China alone the nighttime lights estimate identifies 320 million more people without electricity than the IEA (Elvidge et al. 2010: 8).³⁶ Then again, in India the IEA statistic exceeds the nighttime lights estimate by 220 million.³⁷ In absolute terms, the estimates for these two countries amount to a discrepancy of nearly half a billion people compared to the IEA data. While it is unclear whether IEA data are the

³⁶ Doll and Pachauri (2010: 5663) find a similar discrepancy in China.

³⁷ Author's calculation with data provided in Elvidge et al. (2010).

appropriate yardstick against which to measure the accuracy of nighttime lights data, this example demonstrates that there is a risk in using nighttime lights as a proxy for electrification. This observation draws attention to the presence of outliers, warranting careful cross-validation in each application of the images.

In addition to the problem of outliers, correlations between census-based and satellite-based measures of socioeconomic indicators can differ widely depending on how the information in the images is used. To illustrate this idea, I draw on the data used in Elvidge et al. (2010) once more: panel (a) in Table 9 reports correlation coefficients between satellite-based and census-based electrification measures for 81 countries in the year 2006. The total digital number is only weakly correlated with census data (0.22). This is not surprising given that this measure simply adds the value of every single pixel in a given country, regardless of size and population. When using the images to analyse electrification in an entire country, relative to others, normalization of the digital number is clearly important. That explains why the correlation coefficient grows slightly stronger when using the average digital number instead (0.39).

Instead of using the average, the digital number can be normalized by a fixed area as demonstrated in Michalopoulos and Papaioannou's (2013) paper on pre-colonial ethnic institutions and contemporary African development. They calculate the average digital number, what they call light density at night, for the 10 kilometre radius from the centroid of several enumeration areas in each country of interest. They show a robust and highly significant correlation between the average digital number

and electrification in Tanzania, Zimbabwe, Nigeria, and the Democratic Republic of Congo (Michalopoulos and Papaioannou 2013: Appendix Table 1).³⁸ Their method ensures that comparisons of the average digital number correspond to precisely the same area for each unit (i.e. a 10 kilometre radius). Since this is not the case when the *digital number* is calculated for the entire geographic extent of different countries, this variable is unlikely to be suitable for comparing geographical units that differ greatly in size.

Compared to the *total digital number* the correlation between the *share of lit pixels* and the IEA based measure is more than twice as strong (0.55). This indicates that the binary reclassification of the images and the normalization of the measure, improves the accuracy considerably. With a correlation coefficient of 0.66, the *share of lit pixels* also displays a strong correlation with the *share of population in lit pixels*. The latter in turn displays by far the highest correlation coefficient compared to the census-based variable.

As reported in Elvidge et al. (2010) the correlation coefficient of 0.90 suggests a strong positive relationship between the two measures on the global level. When it comes to cross-sectional analyses, the *share of population in lit pixels* appears to be the most appropriate out of the three aforementioned proxies for electrification.

³⁸ Interestingly, the correlation coefficient decreases when they condition the correlation on log population density. It is not clear, however, how accurate the population data are at such a low level of geographical disaggregation.

When extending this analysis to census-based measures of income and education, a high variation in correlation coefficients can be observed: as reported in panel (b) in Table 9, the correlation between the *share of population in lit pixels* and GDP per capita is only 0.47, far weaker than the correlation between the former and electrification. Regarding the share of the labour force without primary education, panel (c) in Table 9 reveals that the correlation coefficients between this indicator and satellite-based measures have the expected negative sign. The higher the share of labour force without primary education, the lower the satellite-based proxies. However, coefficients are relatively weak, varying between -0.07 and -0.23.

This exercise demonstrates that there are large differences between the results we get – with correlations ranging between 0.23 to 0.90 in the case of electrification – depending on how we choose to turn nighttime lights data into proxies for socioeconomic outcomes. The picture looks different again on the subnational level or when we want to trace socioeconomic outcomes over time.

Table 9: Correlating different measures of socioeconomic outcomes in 81 countries

Census-based indicator	Digital number (DN, total)	Digital number (DN, mean)	Share of population in lit pixels (%)	Share of lit pixels (%)
(1)	(2)	(3)	(4)	(5)
<i>a. Electrification</i>				
Census/IEA: share of hhs with access to electricity	1.00			
Digital number (total)	0.22	1.00		
Digital number (mean)	0.39	-0.06	1.00	
Share of population in lit pixels	0.90	0.17	0.45	1.00
Share of lit pixels	0.55	0.00	0.80	0.66
				1.00
<i>b. Income</i>				
Census/World Bank: GDP per capita (current US\$)	1.00			
Digital number (total)	-0.06	1.00		
Digital number (mean)	0.59	-0.06	1.00	
Share of population in lit pixels	0.47	0.17	0.45	1.00
Share of lit pixels	0.57	0.00	0.79	0.66
				1.00
<i>c. Education</i>				
Census/World Bank: share of labour force without primary education	1.00			
Digital number (total)	-0.07	1.00		
Digital number (mean)	-0.34	-0.22	1.00	
Share of population in lit pixels	-0.23	0.04	0.48	1.00
Share of lit pixels	-0.23	-0.33	0.80	0.62
				1.00

Note: Figures calculated for 81 countries in 2006. Census and share of population in lit pixels are from Elvidge et al. (2010). Share of lit pixels and digital number figures are based on author's own calculations using data acquired with satellite F16 in 2006.

Figure 10: Visualizing changes in electrification in South Africa between 1996 and 2011

Figure 10a: Changes in lit pixels between

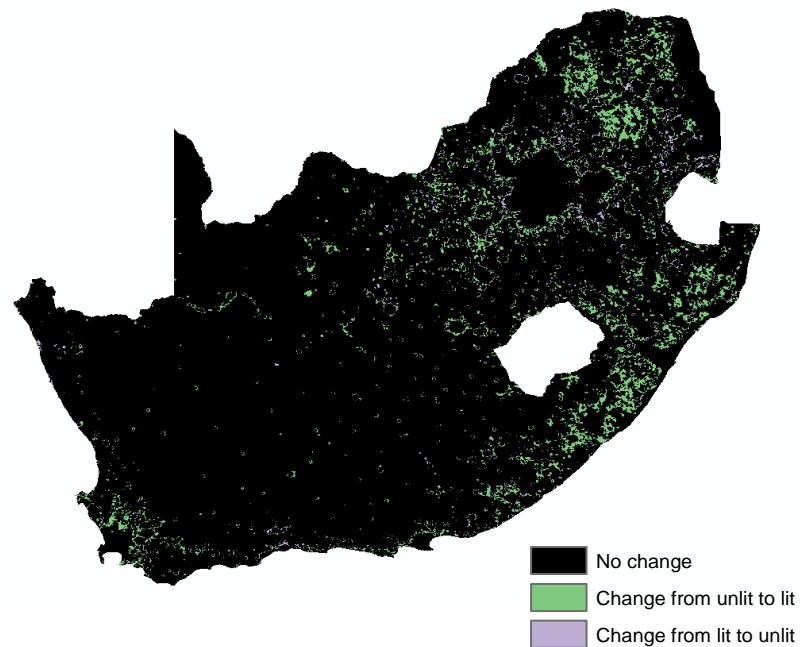
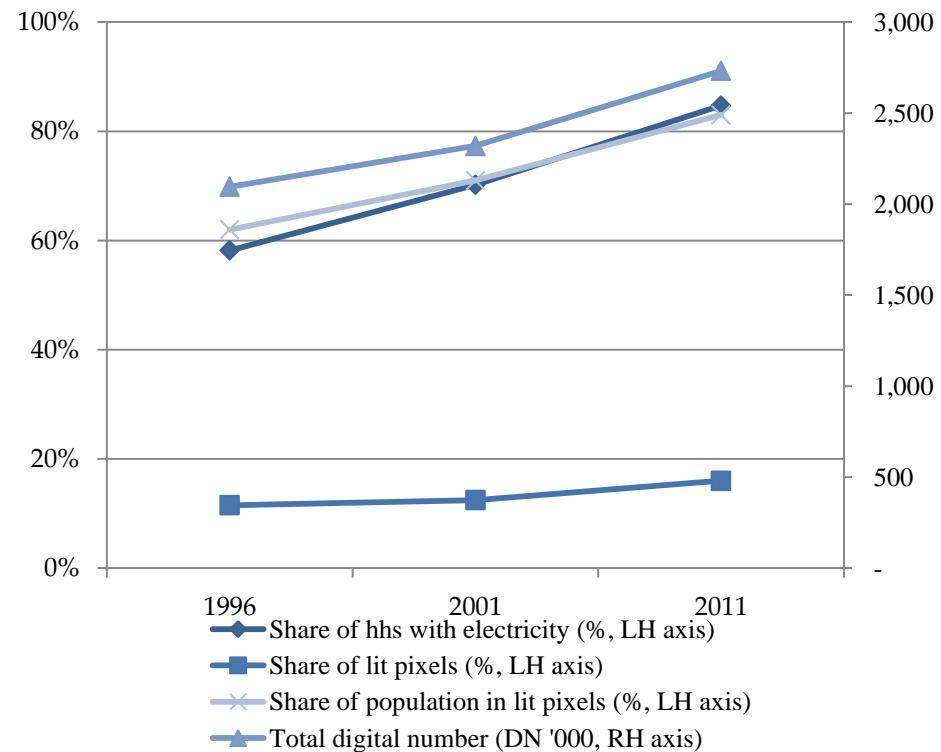


Figure 10b: Changes in census and satellite-based measures of electrification



Note: Images and data processing by NOAA's National Geophysical Data Center. DMSP data collected by US Air Force Weather Agency. Figure 10a was created by the author by subtracting binary versions of the image acquired with satellite F18 in 2011 from the image acquired with satellite F12 in 1996. Figure 10b uses South African census data from 1996, 2001 and 2011 as well as data acquired with satellites F12 in 1996 and 2001, and F18 in 2011.

Indeed, even if the satellite-based measure accurately reflects a socio-economic outcome on the national level, the same may not necessarily be true of the subnational level. South Africa is a case in point. According to the South African census, 85 per cent of households had access to electricity in 2011. This figure represents an increase of nearly 30 percentage points compared to the country's first democratic census in 1996. This period encompasses the beginning of South Africa's post-apartheid mass electrification campaign, which was inherently linked to its transition to democracy. As illustrated in Figure 10a, a comparison of lit and unlit pixels over the period 1996-2011 clearly indicates a positive change over time. Consistent with the census, the spatial concentration of these pixels is much higher in the northeastern half of the country, particularly in the former bantustan areas of Ciskei, Transkei or KwaZulu along the east coast. At closer inspection, the green pixels also validate the well-documented increase in urban sprawl around cities like Cape Town or Johannesburg.

In addition to helping us visualise electrification – which is a powerful research tool in its own right – the images can help us quantify both levels and changes of South Africa's electrification programme. Figure 10b shows a line graph of census data and the three satellite-based electrification rates. The *share of population in lit pixels* increased in step with the census based measure from 62 per cent in 1996 to 83 per cent in 2011. Over the same time period, the *total digital number* increased by 636,510, or 30 per cent. The increase in the *share of lit pixels* is minimal, from 12 per cent in 1996 to 16 per cent in 2011. This is not surprising: vast

areas of the country are uninhabited so even if electricity access were universal we would not expect universal coverage of lit pixels.

While the positive trend in both the *total digital number* and *share of lit pixels* are consistent with census data, it is not immediately clear how to interpret the figures presented by these two measures. What does it mean to observe a change in the *share of lit pixels* or in the *digital number*? As per the definition above, an increase in the share of lit pixels reflects an increase in the geographic expansion of average visible stable lights captured by the satellite sensors. This could reflect the expansion of the electricity grid to previously unelectrified settlements. However, the fundamental problem with this measure is that there is no way of knowing how this expansion relates to the people that live in those settlements.

An increase in the *total digital number* on the other hand could indicate a stronger intensity of existing lights with no geographic expansion. This too does not offer any conclusive evidence as to whether or not more people gained access to electricity. Moreover, as pointed out above, the comparability of digital numbers over time may be compromised due to the sensitivity of the sensors, as discussed above. Even if digital numbers are calibrated to ensure comparability over time, then an issue of interpretability remains. In a cross-sectional analysis, it is difficult to interpret what it means to have a total digital number of 2.7 million as is the case in South Africa in 2011. In the same year, a country like Singapore, by comparison, has a total digital number of just over 40,000,

while the US's total digital number in the same year is over 70 million. Yet, both countries have near universal access to electricity.

The missing piece of information seems to be population or population density. But there is a trade-off: on the one hand we would like to use pixel level population data to determine whether a given pixel is populated. On the other hand population data are not available for most years for which nighttime light images exist and in available years well-documented measurement problems apply. These are particularly pertinent at low levels of aggregation and in rural areas (Doll and Pachauri 2010: 5665). While the *share of population in lit pixels* is most closely related to the census based-electrification rate, both in a cross-section and over time, the reliance on population data can be problematic.

Moreover, the high correlation on the national level does not entail a high correlation on the subnational level. Indeed, a subnational breakdown reveals that the difference between the different measures varies greatly across South Africa's municipalities. As shown in panel (a) of Table 10 below, the correlation between the *share of population in lit pixels* and the census-based electrification measure drops from 0.90 to 0.51 compared to the results presented in Table 9. The correlation between the latter and the *share of lit pixels* is also lower but only decreases slightly from 0.51 to 0.45. The *total digital number* is now hardly correlated with the census, while the latter's correlation with the *average digital number* remains relatively constant, with only a slight drop from 0.37 to 0.30. Regarding the *share of population in lit pixels* the drop in the correlation coefficient is

likely related to the large differences in the Landscan based population figures.

The same is true of income and education indicators, measured here as the average household income in a given municipality and the share of the population without any schooling respectively. Regarding the former, the strongest correlation is observed in association with the *share of population in lit pixels*, at 0.48. The *share of lit pixels* is similarly associated with this measure of economic activity, with a correlation coefficient of 0.43. The weakest association is with the total digital number and average digital number (0.11 and 0.37 respectively). As for education, the variation across variables is similarly large: the *total digital number* shows the weakest correlation with -0.06.

These estimates look at the level of a given socioeconomic outcome, but one of the advantages of the images is that they can help us trace changes over time. It is therefore useful to reproduce the above table for the changes in socioeconomic outcomes, both using the census and satellite-based variables. The results, presented in Appendix Table C2, reveal that correlation coefficients are significantly weakened by this exercise. Regarding the change in electrification presented in panel (a), the highest correlation is between the change in households with access to electricity and share of lit pixels at 0.18 per cent. Satellite-based measures calculated using the digital number are marked grey because digital numbers are not directly comparable over time due to sensitivity of the satellite sensors. They are reported for consistency, but their use is not recommended when looking at changes over time.

The main takeaway from this discussion is that correlation coefficients between satellite-based and census-based socioeconomic outcomes are unstable and outliers may be significant. Therefore, we cannot blindly trust the satellite images to deliver the correct results. As I have shown with South African census data, even if the correlation coefficient between satellite-based and census-based measures is high at the national level, this does not mean that the correlation is equally high on the subnational level. The choice of the satellite-based variable can deliver very different results – social scientists therefore need to investigate which variable is the most appropriate depending on what variable generates the highest correlation coefficient.

Table 10: Correlating different measures of socioeconomic outcomes in South African municipalities

Census-based indicator	Digital number (total)	Digital number (mean)	Share of population in lit pixels	Share of lit pixels
(1)	(2)	(3)	(4)	(5)
<i>a. Electrification</i>				
Census: share of hhs with access to electricity	1.00			
Digital number (total)	-0.07	1.00		
Digital number (mean)	0.30	0.28	1.00	
Share of population in lit pixels	0.51	0.01	0.51	1.00
Share of lit pixels	0.45	-0.01	0.63	0.88
				1.00
<i>b. Income</i>				
Census: share of hhs with average income	1.00			
Digital number (total)	0.11	1.00		
Digital number (mean)	0.37	0.28	1.00	
Share of population in lit pixels	0.48	0.01	0.51	1.00
Share of lit pixels	0.43	0.01	0.62	0.90
				1.00
<i>c. Education</i>				
Census: share of population without any schooling	1.00			
Digital number (total)	-0.06	1.00		
Digital number (mean)	-0.43	0.28	1.00	
Share of population in lit pixels	-0.45	0.01	0.51	1.00
Share of lit pixels	-0.47	-0.01	0.62	0.90
				1.00

Note: Figures calculated for South Africa's municipalities in 2001, using South African census data and nighttime light images acquired with satellite F15.

4.5 Categorizing pitfalls and navigating them

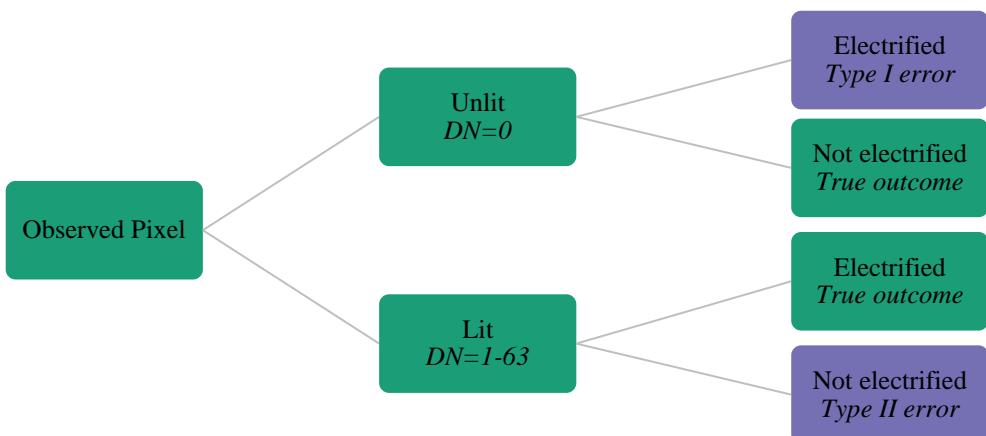
The previous section emphasized large variation between different satellite-based measures of socioeconomic outcomes. This raises two important questions: first, why might the images not capture the true level of change in a socioeconomic outcome? Second, how can we navigate the pitfalls? Answers to both questions follow in this section.

There are a number of pitfalls that may undermine the usefulness of the images for measuring socioeconomic outcomes. Small et al. (2005) and Doll (2008) provide useful summaries of the main technical issues with the satellite sensors. Their summaries are however quite technical and implications for social science research are difficult to interpret. Yet, it is these implications that are most relevant to those who are more concerned with the practical question of how to use the final data product than understanding the technical intricacies of the satellite sensors. In order to explain why the images might not capture socioeconomic outcomes, I categorize the main pitfalls in terms of whether they are likely to induce Type I or Type II errors in social science applications.

To illustrate the logic behind this categorization, it is useful to consider residential access to electricity as an example. As demonstrated in the decision tree in Figure 11 below, there are two possible outcomes for each observed pixel: a pixel is either unlit, with a digital number of 0, or the pixel is lit, with a digital number between 1 and 63. If a pixel is unlit, then no stable, anthropogenic lights are visible in this pixel. From this observation we may infer that people living in this pixel do not have

access to electricity. If this is correct, then we are directly inferring the true outcome from the images. However, it should be clear that there is a gap between our observation and the conclusion we infer. It is of course entirely feasible that people living in the unlit pixel do have electricity, even if we cannot infer this from the images. In this case we incur a Type I error: we are wrongly inferring that fewer people have electricity access than is actually the case. In this case, we are underestimating the actual electrification rate.

Figure 11: Possible outcomes of observed pixels in nighttime light imagery



If the observed pixel is lit, we can conclude that average stable lights are present in this pixel. From this observation we may infer that people living in the lit pixel have access to electricity. If this is the case, then we are inferring a true outcome. However, it is possible that some or even none of the people in this pixel may in fact not have access to electricity, in which case we incur a Type II error. This in turn leads to an overestimation of the actual residential electrification rate since we are assuming that more people have electricity than is the case.

Which type of error is more likely to occur depends on the area we are interested in studying. Since social scientists are generally interested in areas covered by more than one pixel, they may incur a combination of Type I and Type II errors in their samples.

Type I errors. Since nighttime light images offer an aerial view of the world, they can only capture light that is visible from above. When we observe an unlit pixel, there are at least three reasons why this may lead to Type I error: first, the pixel may appear unlit because there is no stable, visible, outdoor lighting. Instead, there may only be indoor lighting, which simply cannot be captured from outer space. Second, even if outdoor lighting is present, it may be too weak to be detected. There is a minimum threshold at which the nighttime light images are able to detect lighting. If an area is very sparsely populated, with only 15 to 20 street lamps within a one-kilometre radius for example, then the satellite sensor may not be able to pick them up. Min (2014) points out this issue in his analysis of nighttime lights in India, noting the lowest population density at which lights are observed are 60 people per square kilometre. All pixels with fewer than 60 inhabitants were unlit regardless of the presence of outdoor lighting. Nighttime lights will therefore literally not be able to shed light on the most sparsely populated areas and associated research questions. This point is further reinforced by Doll and Pachauri (2010) who focus on rural populations without access to electricity. Their study points to large discrepancies between nighttime light measures of electrification and IEA data. They demonstrate that nighttime light data may not be able to proxy for electrification if population density is not

high enough or if electricity usage is not dense enough or not used in outdoor lighting (Doll and Pachauri 2010: 5665).

Third, lit pixels capture stable lights averaged over a one-year period. In many countries, however, within-year disconnections or power outages are a common feature, which depending on the severity may cause pixels to appear unlit even though inhabitants would be considered to have access to electricity under traditional definitions of the term. Unfortunately the images provide “little ability to discern periods of blackouts, power outages, or changes in infrastructure and customer access” (Min et al. 2013: 8130).

In all three cases, the evidence from the images would lead us to reject the hypothesis that residents living in a given area have access to electricity, even if the hypothesis is true. This would lead us to underestimate the true level of electrification. Social scientists need to be mindful about both the presence of outdoor lighting, population densities, and the role of disconnections pertaining to the geographic area under review.

Type II errors. One problem that can lead to Type II error is the inability to identify the source of visible lighting. Using nighttime lights as a proxy for residential access to electricity is particularly risky in areas with persistent public or industrial light sources. In their study of nighttime lights in Israel, Levin and Duke (2012: 4) show that outside of towns, the main source of lights is from highways and other lit roads that connect urban areas. The inspection of a nighttime light image of an oil producing

country such as Nigeria also reveals that industrial sites such as mines and gas processing plants produce a lot of light output. Gas flares too are visible in the imagery explaining off shore lighting, for example in the North Sea. Such light sources are not necessarily indicative of residential access. Moreover, in many developing countries public street lighting may be available in areas long before individual dwellings are connected to the grid (or vice versa). In such cases the inference that residents have electricity based on the observation of a lit pixel, is a Type II error. Using the variable *share of lit pixels*, for example, the aggregation of lit pixels coupled with the absence of residential access will produce an overestimation of the true level of residential access to electricity.

While this source of Type II error is likely to affect most studies that seek to distinguish between public and private entities, it has remained largely unaddressed in existing social science applications. One possible way to address this issue is to mask large industrial areas prior to carrying out the calculation. Elvidge et al. (2010: 6) point out the importance of this step and explain how this process works in practice. When applying the images to a given country context, it is important to be aware of the presence of gas flaring or other large industrial sights. Masking out such areas can significantly reduce the likelihood of Type II error.

Another problem that increases the likelihood of Type II error is the overglow effect observed in nighttime light images. This effect is observed when lights from one pixel spill over to another. This is particularly pronounced over water or snow cover (Pinkovskiy, 2011: 9), but also in pixels adjacent to large urban centres. As above, the overglow

effect creates observations of lit pixels that suggest the presence of stable manmade lights where they may not in fact exist. This can create considerable inaccuracies when using nighttime lights as a measure of electricity consumption (Townsend and Bruce 2010: 4461). Townsend and Bruce (2010) have developed an overglow removal model, which can help social scientists navigate the problem. This reinforces my earlier point about understanding the context to which the images are applied.

Comparability in time-series. One of the main advantages of the images is that they provide a unique, high-resolution view of human activity on an annual basis for two decades, an ideal foundation for panel data analysis. Yet, comparisons over time pose a challenge: due to the use of different satellites and thus sensors to obtain nighttime light images, there are concerns about their comparability over time. The problem is that “satellites differ in their optical quality and may degrade over time” (Chen and Nordhaus 2010: 12) and there is no in-flight calibration of the visible band on the OLS (Elvidge et al. 2013: 3). Indeed, according to NOAA’s website, sensor problems led to the discontinuation of data collection from sensors F10 and F11 in 1995 (NOAA 2013a). The text file accompanying each download from NOAA makes this explicit: “while the time-series of annual cloud-free composites were produced using the same algorithms and stringent data selection criteria, the digital number (DN) values are not strictly comparable from one year to the next.” (NOAA 2013b). Doll (2008: 15) shows that there can be considerable differences in brightness between the sensors flying on different satellites by comparing the sum of digital numbers in India for the years 1992-2003 using overlapping years from different sensors. Consequently, the value

of a pixel in one year is not necessarily comparable with the value of that same pixel in another year. This is clearly a significant limitation.

Ongoing research is developing calibration methods to help interpret brightness changes over time. Elvidge et al. (2013) make a first attempt to quantitatively analyse an intercalibrated time series. They conclude that intercalibrated nighttime lights can successfully reveal electrification patterns, but they also highlight the importance of verifying the data in a given context before drawing conclusions. Without calibration, direct comparisons between the digital numbers over time can be misleading. The argument against using digital numbers to make comparisons over time is an argument in favour of using a binary classification to analyse the images. Instead of interpreting the brightness of a pixel as light intensity, a binary variable focuses on the presence or absence of light. The comparability concern can therefore be mitigated by reclassifying pixel values between 1 and 63 as lit pixels and the remainder as unlit pixels. As noted by Michalopoulos and Papaioannou (2013: 17) the binary nature of this measure implies that the non-linearity of luminosity is no longer a concern.

Given these pitfalls it should be clear that the assumption that the images can serve as a reliable proxy for socioeconomic outcomes in all contexts is flawed. To help researchers decide what satellite-based measure is most appropriate for their application of the images, Table 12 provides an overview of the satellite-based variables and how they compare to each other and census data. The main advantage of census data is that they capture household electrification and other socioeconomic outcomes

more closely than the satellite-based measures. In terms of the unit of analysis, households or persons are usually of more interest, which means that census-based or other administrative data are the more obvious choice. The main drawback is that researchers are limited to administrative boundaries and years that may not match the researchers' data requirements. Administrative boundaries may change over time or may be too coarse. The main advantage of the images: the flexibility of zooming in on any geographic area with a resolution of one square kilometre undoubtedly generates great value for social science research. The availability of annual data over two decades is certainly also advantageous. The respective advantages and pitfalls of the two data sources highlight their complementarity.

As outlined in section 4.4 above, there are different ways of turning the images into proxies of socioeconomic outcomes. Each variable has pros and cons and it is important to understand these when deciding which variable to use. The variable most closely related to the census is the *share of population living in lit pixels*. By taking the location of local populations into account, the unit of analysis of this variable are people, therefore allowing for an interpretation that is most closely related to the census data. The comparability in time-series is, however, limited to years with available gridded population data, which as explained above is not available free of charge and the publisher recommends the use of the latest available year. In fact, “previous versions of the LandScan data are made unavailable as new datasets are released because the makers of the LandScan data ... caution against using the data as a change detection or migration tool” (Elvidge et al. 2013: 4690/1). This may or may not be

problematic for social scientists depending on whether it is realistic to assume constant population growth over a given period. There may, however, be additional concerns regarding the reliability of the population data at low levels of geography. Since Landscan data capture daytime population counts, we need to critically assess the applicability of the data, not just in the context of a given country, but also at the desired level of geographical disaggregation. This concern also applies to the *share of populated pixels*. Even though this variable is less reliant on the accuracy of population counts in a given pixel, it still relies on the accuracy of the location of populated pixels.

The *share of populated pixels* is the least restrictive satellite-based variable, when it comes to both data availability and use in time-series. Unlike the population-based variables, this variable only relies on the images and is therefore not reliant on the accuracy of population counts. It also mitigates the comparability concern by reclassifying pixels into a binary classification of lit or unlit pixels rather than using the full range of digital numbers. These characteristics render this variable the most useful for panel data analysis. However, both the *share of lit pixels* and the *digital number* use units of analysis that are more difficult to interpret. It is not immediately clear how the *share of lit pixels* or *digital number* in a given area relates to the number of electrified households or the level of income. That is why these variables are likely to be most useful in contexts where they are used to trace changes over time rather than determining the level of human activity at one point in time.

Table 11: Overview matrix of electrification measures

	Census-based electricity measure		Satellite-based electricity measures		
	Share of hhs with electricity	Share of population living in lit pixels	Share of populated lit pixels	Share of lit pixels	Digital number
<i>Unit of analysis</i>	Households	Persons	Pixels	Pixels	Digital number
<i>Data sources required</i>	1. Census data	1. Nightlights 2. Landscan	1. Nightlights 2. Landscan	1. Nightlights	1. Nightlights
<i>Years Available</i>	1. 1996, 2001, 2011	1. 1992-2012; 2. 2011*	1. 1992-2012; 2. 2011*	1. 1992-2012	1. 1992-2012
<i>Lowest available resolution</i>	Varies with changes in census boundaries; e.g. census wards 2011	~1km ²	~1km ²	~1km ²	~1km ²
<i>Comparability in time-series</i>	Yes, but limited to census years.	Limited to years with available gridded population.	Limited to years with available gridded population.	Yes.	No, unless digital number calibrated.
<i>Main advantage</i>	Most closely captures access to electricity.	Takes population location and counts into account; most closely related to census data.	Takes population location into account, less reliant on accuracy of population counts than the share of population living in lit pixels.	No reliance on second data source; can be calculated for two decades with 1km ² resolution.	Takes radiance characteristic of nightlights into account.
<i>Main pitfall</i>	Limited to administrative boundaries and census years.	Limitations of population data apply: reliance on accuracy of pixel level population counts which represent ambient population; limited availability.	Limitations of population data apply: accuracy of location of populated pixels.	Unit of analysis is pixel; does not take into account where people live.	Digital numbers are difficult to interpret. Data not normalized, making comparisons limited to areas of equal size.

Note: *Landscan data can be requested online. They cannot be downloaded immediately and are not free of charge. 2011 is the latest year for which they are currently available.

To summarise this discussion, the following four steps may help social scientists navigate the potential pitfalls presented by nighttime lights data:

1. *Geography.* Examine the desired geographical hierarchy in the country/countries of interest. Identify large industrial sights, particularly gas processing plants as they produce a lot of light. Identify national parks, large lakes, forests, etc, i.e. areas where you do not expect any lighting to be present. Mask them if possible. Identify areas with population counts lower than the minimum population density threshold.
2. *Variable choice.* Choose the appropriate satellite-based proxy variables based on geography, data availability, and requirements outlined in Table 12. E.g., if the share of top-coded pixels is likely to be very high in the area under review, the *digital number* is unlikely to be the appropriate variable. Do not compare digital numbers over time. Use binary classification of the images or intercalibrate.
3. *Cross-validation.* Produce simple correlations between satellite-based variables and alternative data source(s). Choose the satellite-based proxy with the highest correlation. Given the high risks of using the data, it is not recommended to do so for a level of geography for which no suitable data for cross-validation is available.

4. *Interpretation.* Interpret the satellite-based measure precisely as per the definitions outlined in sections 4.3 and 4.4 above. Avoid using vague concepts such as luminosity or light density.

4.6 Conclusion

The accuracy of empirical analyses crucially depends on the quality of the data used to generate the results. Yet, high quality data on socioeconomic outcomes are inherently difficult to obtain, particularly in the developing world where inconsistent data collection methods often pervade official estimates. Many developing countries lack reliable population censuses. And even if data exists, quality is often poor and there is no way to verify the data released by governments. A potentially powerful substitute is presented by data collected with remote sensing technology. Nighttime light images allow researchers to zoom in on any area of interest with a local accuracy of approximately one square kilometre. The scale and scope is unparalleled by administrative data, as even the smallest administrative units usually encompass several square kilometres and data collection methodologies often vary substantially across countries.

As I have argued in this chapter, there are important problems with this data product that tend to be insufficiently addressed in existing applications of the data to social science contexts. In the long term, further improvements in technology will hopefully address many of these issues. Some progress is already being made: in 2011 a new instrument, the Visible Imaging Radiometer Suite (VIIRS), was launched. According to Elvidge et al. (2013: 1) “the VIRRS offers substantial

improvements in spatial resolution, radiometric calibration and usable dynamic range when compared to the DMSP low light imaging data". The usefulness of nighttime lights for social science applications will increase with improvements to the sensor and corresponding data products.

However, for the time being, social scientists cannot simply trust that remote sensing technology delivers datasets that we can use as perfect substitutes for socioeconomic outcomes. It is important to use these data critically, with an understanding of their limits and what can go wrong. Given the risks of Type I and Type II errors, it is not sufficient to rely on the relatively robust positive correlation between satellite-based and administrative data in cross-country studies to justify the application of the images to individual country contexts and for different geographical hierarchies. Based on the analysis of potential pitfalls as well as cross-validating the data with South African census data, I generate lessons learned that will hopefully guide social scientists in applying nighttime lights data in the future.

The findings presented in this chapter have two important implications: first, from a practical point of view, it will require more time and effort from social scientists to use the data. Due to the formal constraints of journal articles, an extensive justification of the use of the images may not seem realistic at first. Luckily, online appendices are offering new possibilities for doing so. Michalopoulos and Papaioannou's work (2013) provides useful guidance for how to combine the use of the data with rigorous cross-validation using online appendices. Second, having to

cross-check the data before using it limits the usefulness of the satellite data because it is likely to be most valuable in cases where alternative data does not exist. This is unfortunate, but given the significant potential for error further research is needed to leverage the impact of remote sensing technologies in the social sciences. This research will require the integration of the remote sensing literature with the social science literature: only by advancing our understanding of the science behind the technology can we optimise the use of the data to further our knowledge about human development. In the end, empirical results generated through social scientific inquiry will only ever be as good as the data used to generate them.

Chapter V

Discussion and implications

This thesis provides insights into the important question of how government affects the distribution of financial resources and services. As stated in the introduction, each paper was authored as a self-contained piece of research with separate conclusions and implications. In general, each paper makes a contribution to answering the question of whether public goods and services improve as a result of democratization. The first two papers do so by analysing a specific outcome influenced by government policy on the provincial and municipal level respectively. The third paper adds to this analysis by enhancing the utility of nighttime light satellite imagery for quantitative social scientists. One of the centrepieces of this research is the incorporation of geographic analysis to study spatial patterns of the distribution of government resources. In addition to the conclusions

already provided in sections 2.7, 3.7 and 4.6, this final chapter consolidates the findings, summarises the contributions, and provides a discussion on areas for further research.

Substantive contributions. Existing empirical work often leaves unresolved what precisely it is about democracy that accounts for the effect on the outcome of interest. The main substantive contribution of this thesis is to shed some light on this issue: chapter 3 shows that enfranchisement increased household electrification rates during the first period of democratic local government in South Africa. This finding reinforces the existing literature and adds a first study in the context of South Africa as a contemporary case of democratization. It also addresses a main limitation in the literature, which often examines the impact of enfranchisement on resource allocation rather than services delivered, for example, by focusing on spending on local public health (Miller 2008) or education and social services (Vernby 2013). However, if funds are not spent as intended – as is often the case in developing countries – such an analysis cannot inform us about how the lives of the people for whom the allocations are intended are affected, even though ultimately that is what we care about. Another key contribution of chapter 3 is therefore that it adds to our understanding of how people's lives were affected in post-apartheid South Africa: in the period 1996-2001, households in municipalities with higher shares of newly enfranchised voters, experienced higher gains in access to electricity (of between 3 and 6 percentage points per standard deviation of enfranchisement).

These improvements are observed across the country, but the chapter draws attention to a racial pattern, too: households in municipalities with higher shares of black voters benefitted the most in terms of receiving access to electricity (approximately 9 percentage points per standard deviation of enfranchisement). The effect is also positive but less strong for coloured voters, and statistically insignificant for newly enfranchised Indian voters. Previous “homeland” areas experienced some of the greatest changes in electrification. These areas, reserved for the black population under the policy of racial segregation, became synonymous with poverty and underdevelopment, reflecting the geography of apartheid (Christopher 1994). Recognizing that South Africa continues to be one of the most unequal societies in the world, the finding suggests that enfranchisement did have a positive effect on spatial inequality with respect to electricity provision.

Unlike many other studies on the effects of democratization, this thesis takes one step further by carefully separating the effect of enfranchisement and partisan representation to clarify the mechanism by which democratization affects the provision of essential basic services. The findings presented in chapter 3 demonstrate that in South Africa the state-owned electricity company supplied the incumbent party’s core constituencies with an additional top-up in this initial period of the post-apartheid electrification campaign. In the South African context, it therefore did not only matter that people got the right to vote, but also who they voted for. Interestingly, this effect is only observed in municipalities where the state-owned electricity company controls electricity reticulation, not in those with municipal distributors even if

they were controlled by the incumbent party. In other words, the partisan composition of the local council made no difference to the distribution of electricity to households. The contribution of this finding to the existing literature is that it highlights the importance of pinpointing who is responsible for service provisions in order to understand observed spatial patterns of service delivery.

The thesis also adds to our understanding of the role of subnational electoral competition as an incentive to manipulate the distribution of public funds. The detection of political budget cycles in the equitable share of South Africa's intergovernmental transfer system suggests that even in the absence of electoral competition on the national level, electorally motivated distortions in financial resources may occur on the subnational level. The direction of the distortion points in the same direction as the Eskom-effect uncovered in chapter 3, only in this case the ANC's core provinces rather than municipalities benefit from the distributive patterns. To stop the distortion of the equitable share for political gain, the government's ability to manipulate the formula would have to be removed. In order to do so, more research is required to identify the precise mechanism through which past manipulations have been implemented. Regardless of the fiscal instrument, however, chapter 2 demonstrates that both the timing of elections and electoral competition can play an important role when it comes to the distribution of financial resources to subnational units in a new democracy.

An important limitation is that this thesis does not touch on the vast problems of corruption that burden South Africa's financial management

system: unauthorized payments and misallocations, contracts without competitive bidding, and manipulation of tenders are just a few examples that are commonly featured in the country's news stream. According to the civil society organization Corruption Watch (2013), the South African government loses approximately USD 2.7 billion to procurement corruption each year. However, an important implication of the results presented in chapter 2 is that in addition to off-the-book malpractices, official channels may be subject to manipulation, too.

One of the main obstacles in this line of research is that high-quality data on the delivery of goods and services, and socioeconomic outcomes in general, are inherently difficult to obtain. This is particularly true of Africa, where inconsistent data collection methods often pervade official estimates (Jerven 2013). Chapter 4 makes a contribution to the social science toolkit by demonstrating how nighttime light satellite imagery can be used to proxy for socioeconomic outcomes. The chapter provides the first literature review of interdisciplinary applications of the images tailored to the social sciences. By synthesising the pitfalls that arise from such applications, I emphasize the risk of Type I and Type II errors, adding to our substantive knowledge about the images. The paper also makes important methodological contributions, which are further discussed below.

Methodological contributions. This thesis demonstrates how distributive outcomes can be studied on the subnational level, not only by adding two examples for such studies to the still small inventory of subnational analyses on the subject, but also by offering new datasets to the research

community. In the introduction to this thesis, I argued that the availability of high quality data speaks in favour of using South Africa as a testing ground. Here I would like to add to this point that availability does not necessarily imply user friendliness. The budget information obtained for chapter 2, for example, is readily available from the South African Treasury's website, but only in portable document format such that budget information cannot be directly converted into tables. Instead, putting together a dataset from the available information required time intensive data entry. The result is a new longitudinal data set, consisting of annual observations of all nine provinces for the period 1995 to 2010. Researchers can build on this by adding additional years of budget information in the coming years. This year's twentieth anniversary of South Africa's first democratic election presents a particularly exciting opportunity to extend existing datasets with a fifth election and further develop our understanding of South Africa's political economy. Fortunately, this will be a much easier task as the South African Treasury now makes budgetary and financial information available in number format.

The data behind the results in chapter 3 also offer a unique longitudinal dataset to the research community: since administrative boundaries below the provincial level were redrawn between South Africa's first census in 1996 and the second census in 2001, it is not possible to directly compare socioeconomic outcomes below the level of the provinces over this period. Chapter 3 develops a methodology for carrying out precisely such analyses: using the spatial attributes of 21,243 subplaces from the 2001 census, we aggregate their information up to the level of 799

municipal boundaries from the 1996 census. We chose 1996 municipal boundaries as our benchmark in order to further combine the data with 1995/6 local election results, but the methodology could equally be applied to other boundaries. Since changing administrative boundaries are a common occurrence, particularly in developing countries, the method applied to this research (documented in detail in Appendix B) will hopefully be helpful for other researchers in the future.

The empirical analysis in chapter 3 uses two independent data sources to assess electrification outcomes, providing an example for how nighttime lights data can be used. This is an advance over studies solely based on data collected with involvement by the governmental units that are under examination. However, as demonstrated in chapter 4, it is important to be mindful of the risks of using these data. By carrying out the first literature review of social science applications of nighttime lights, chapter 4 highlights the opportunities presented by the images to social scientists engaged in quantitative research. The value proposition is clearly attractive: the images offer free and easy access to global, high-resolution data that can proxy for different measures of human development over the past two decades.

At the same time, chapter 4 demonstrates that the importance of cross-validating the data cannot be underestimated. By using South African census data, I show how this can be done in practice. In the future, the usefulness of nighttime lights for social science applications is likely to increase in line with improvements to the sensor, already ongoing with the aforementioned Visible Imaging Radiometer Suite, and

corresponding data products. This will open up new opportunities for combining the images with traditional datasets and further integrating them into social science research.

Overall, the three papers included in this thesis emphasize the importance of considering spatial patterns when it comes to the distribution of government resources. Be it the allocation of financial resources to regions with different electoral profiles or the distribution of a basic public service such as electricity: geography matters. By using maps as a visualization tool for key variables, this thesis offers examples for how to add a geographical dimension to the way in which we can study socioeconomic outcomes. This is a powerful method for detecting spatial patterns and uncovering correlations for further investigation. The use of geographic information systems is not currently receiving a lot of attention in political economy, but hopefully this will change with time.

External validity and directions for future research. For the most part, the empirical findings presented in this thesis are based on a single country and specific outcomes. As argued above, the advantage of focusing on subnational units addresses some of the problems observed in cross-country studies. The disadvantage is that interpretations of single cases are often unrepresentative of larger patterns (Golden and Min 2013: 83). An important challenge for this type of research is therefore to manage the opposing forces between data precision on the one hand and external validity on the other hand. Admittedly, this thesis has been more concerned with the former than the latter, with a preference for

generating precisely defined results rather than ones that are easily generalizable.

However, a comparison of the findings with other research papers uncovers interesting parallels: the findings presented in chapter 2 are consistent with the studies at the national level that have found stronger and more consistent evidence of political budget cycles in developing countries (Golden and Min 2013: 83). A comparison with Banful's (2011) study on intergovernmental transfers in Ghana further substantiates this observation on the subnational level: Banful also finds that there is scope for politically motivated targeting in a formula-based system. Regarding the magnitude of the distortion, comparisons with similar studies suggest the effect of vote margin on the equitable share in pre-election years is large³⁹ and point in the same direction (e.g., see Case 2001, Miguel and Zaidi 2003). The results also confirm Brender and Drazen's (2005) finding that PBCs weaken over time. A natural extension of this research is to determine whether the size of the effect is linked to the support incumbents receive, i.e. whether votes can in fact be retained with such allocations. Empirical work is divided over whether incumbents are rewarded or punished by voters for such distributive allocations (Golden and Min 2013: 84). This would be an interesting area for future research in the South African context.

The finding that enfranchisement improves the delivery of public services, as suggested in chapter 3, conforms with the theoretical model

³⁹ See section 2.5 for examples.

established by Meltzer and Richard (1981). This establishes an interesting link between the effect of enfranchisement in South Africa and Sweden: Vernby (2013) finds that enfranchising non-citizens in Sweden caused substantial shifts in budget priorities in municipalities where non-citizens made up a non-negligible share of voters. While Vernby studies spending on education and social services as an outcome, not electrification, it is remarkable that these two studies on enfranchisement in completely different settings point in the same direction. Similarly, Miller (2008) finds that the enfranchisement of American women led to increases in local public health spending and traces these to subsequent improvements in child mortality. These comparisons give external validation to the finding that enfranchisement matters for service delivery.

This suggests that in the South African context, public services other than electricity may have improved with democracy, too. Indeed, Kudamatsu's (2012) work on how democracy affects health outcomes in Sub-Saharan Africa demonstrates that the basic pattern documented for electricity may apply more widely, e.g., regarding the provision of public housing. Exploring this further presents an interesting avenue for future research, which should also address the important issue of the quality and sustainability of services delivered.

While the examples above are useful to draw parallels to other work in the existing literature, it is important not to overgeneralise the results. Kramon and Posner (2013: 469) reinforce this by showing that patterns of distribution can and do vary across outcomes, even within countries. By

the same token, measurement of outcomes and the way we choose to interpret the data matter greatly for the way we understand results. Only by recognising the fact that our answers depend on both the outcomes and the data we use, can we continue building on our knowledge of the social consequences of democracy.

References

- Ablo, E., and R. Reinikka (1998). "Do Budgets Really Matter? Evidence from Public Spending on Education and Health in Uganda." World Bank Policy Research Working Paper Series 1926. Washington DC.
- Acemoglu, D., and J. A. Robinson (2000a). "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4): 1167-1199.
- Acemoglu, D., and J. A. Robinson (2000b). "Democratization or repression?" *European Economic Review* 44(4): 683-693.
- African National Congress (1994). The Reconstruction and Development Programme: A Policy Framework. Johannesburg, African National Congress by Umanyano Publications.
- African National Congress (1995). "1995 Local Government Elections Manifesto. A Better Life: Let's Make It Happen Where We Live." August 24, 1995. Downloaded from <http://www.anc.org.za/show.php?id=2541> (last accessed February 20, 2013).
- Agnew, J., T. Gillespie, J. Gonzalez and B. Min (2008). "Baghdad nights: evaluating the US military surge using nighttime light signatures." *Environment and Planning* 40: 2285-2295.
- Air Weather Service (1974). "Defense Meteorological Satellite Program (DMSP) User's Guide." Illinois, Air Weather Service.

Akhmedov, A., and E. Zhuravskaya (2004). "Opportunistic Political Cycles: Test in a Young Democracy Setting." *The Quarterly Journal of Economics* 119(4): 1301-1338.

Alence, R. (2004). "South Africa After Apartheid: The First Decade." *Journal of Democracy* 15(3): 78-92.

Alesina, A. (1987). "Macroeconomic Policy in a Two-Party System as a Repeated Game." *Quarterly Journal of Economics* 102: 651-678.

Alesina, A., N. Roubini with G. Cohen (1997). "Political Cycles and the Macroeconomy." Cambridge, MIT Press.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2012). "Ethnic Inequality." NBER WP 18512, CEPR Discussion Paper 9225.

Alt, J., and D. Lassen (2005). "The Political Budget Cycle is Where You Can't See It: Transparency and Fiscal Manipulation." Economic Policy Research Unit Working Paper Series.

Alt, J., and D. Lassen (2006a). "Transparency, Political Polarization, and Political Budget Cycles in OECD Countries." *American Journal of Political Science* 50: 530-50.

Alt, J., and D. Lassen (2006b). "Fiscal Transparency, Political Parties, and Debt in OECD Countries." *European Economic Review* 50: 1403-39.

Alt, J., and S. Rose (2007). "Context-Conditional Political Budget Cycles" in Boix, C. and S. Stokes (Ed.) *Oxford Handbook of Comparative Politics*,

New York, Oxford University Press.

Álvarez-Berrios, N. L., I.K. Parés-Ramos, and T.M. Aide (2013). "Contrasting patterns of urban expansion in Colombia, Ecuador, Peru, and Bolivia between 1992 and 2009." *Ambio* 42(1): 29-40.

Angrist, J. D., and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, Princeton University Press.

Arulampalam, W., S. Dasgupta, A. Dhillon, and B. Dutta (2009) "Electoral goals and center-state transfers: A theoretical model and empirical evidence from India." *Journal of Development Economics* 88(1): 103-119.

Baleiras, R., and J. da Silva Costa (2004). "To be or not to be in Office again: an Empirical Test of a Local Political Business Cycle Rationale." *European Journal of Political Economy* 20: 655-671.

Banful, A. (2011). "Do Formula-Based Intergovernmental Transfer Mechanisms eliminate Politically Motivated Targeting? Evidence from Ghana." *Journal of Development Economics* 96(2): 380-390.

Beall, J., S. Mkhize, and S. Vawda (2005). "Emergent Democracy and 'Resurgent' Tradition: Institutions, Chieftaincy and Transition in KwaZulu-Natal." *Journal of Southern African Studies* 31(4): 755-771.

Bekker, B., A. Eberhard, T. Gaunt, and A. Marquard (2008). "South Africa's Rapid Electrification Programme: Policy, Institutional, Planning, Financing and Technical Innovations." *Energy Policy* 36(8): 3125-3137.

Benito, B., Bastida, F., and C. Vicente (2013). "Creating Room for Manoeuvre: a Strategy to Generate Political Budget Cycles under Fiscal Rules." *Kyklos* 66(4): 467-496.

Besley, T., Persson, T. and D. Sturm (2010). Political Competition, Policy and Growth: Theory and Evidence from the United States. *Review of Economic Studies* 77(4): 1329-1352.

Blais, A., and R. Nadeau (1991). "The Electoral Budget Cycle." *Public Choice* 74: 389-403.

Blaydes, L., and M.A. Kayser (2011). "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55(4): 887-908.

Block, S. (2002a). "Elections, Electoral Competitiveness, and Political Budget Cycles in Developing Countries." Center for International Development at Harvard University Working Paper 78.

Block, S. (2002b). Political Business Cycles, Democratization, and Economic Reform: The Case of Africa. *Journal of Development Economics* 67: 205-228.

Block, S., S. Singh and K. Ferree (2003). "Multiparty Competition, Founding Elections and Political Business Cycles in Africa." *Journal of African Economies* 12(3): 444-468.

Boix, C. (2003). Democracy and Redistribution. Cambridge, Cambridge University Press.

Bratton, M., and N. van de Walle (1997). *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective*. Cambridge, Cambridge University Press.

Brollo, F. and T. Nannicini (2012). "Tying your enemy's hands in close races: The politics of federal transfers in Brazil." *American Political Science Review* 106(4): 742-761.

Brender, A. (2003). "The effect of fiscal performance on local government election results in Israel: 1989–1998." *Journal of Public Economics* 87(9): 2187-2205.

Brender, A., and A. Drazen (2005). "Political Budget Cycles in New versus Established Democracies." NBER Working Paper 10539.

Brown, D. S., and A. M. Mobarak (2009). "The Transforming Power of Democracy: Regime Type and the Distribution of Electricity." *American Political Science Review* 103(2): 193-213.

Bueno de Mesquita, B., A. Smith, R. M. Siverson, and J. D. Morrow (2003). *The Logic of Political Survival*. Cambridge, MA, MIT Press.

Cameron, R. (1996). "The Democratisation of South African Local Government." *Local Government Studies* 22(1): 19-39.

Cameron, A. C., J. B. Gelbach and D. L. Miller (2011). "Robust inference with multiway clustering." *Journal of Business and Economic Statistics* 29(2): 238-249.

Carlsen F. (1998). "Rational partisan theory: empirical evidence for the US." Southern Economic Journal 65: 64-82.

Case, A. (2001). "Election goals and income redistribution: Recent evidence from Albania." European Economic Review 45(3): 405-423.

Casey, K. E. (2013). "Crossing Party Lines: The Effects of Information on Redistributive Politics." Stanford GSB Research Paper 2099R.

Cerda, R., and R. Vergara (2007). Business Cycle and Political Election Outcomes: Evidence from the Chilean Democracy. Public Choice 132: 125-136.

Chand, T. K., K. V. Badarinath, C. D. Elvidge and B. T. Tuttle (2009). "Spatial characterization of electrical power consumption patterns over India using temporal DMSP-OLS night-time satellite data." International Journal of Remote Sensing 30(3): 647-661.

Chauvet, L., and P. Collier (2009). "Elections and Economic Policy in Developing Countries." Economic Policy 24(59): 509-550.

Chen, X., and W.D. Nordhaus (2010). "The Value of Luminosity Data as a Proxy for Economic Statistics." Cowles Foundation Discussion Papers 1766, Cowles Foundation for Research in Economics, Yale University.

Christopher, A. J. (1994). The Atlas of Apartheid. Johannesburg, Witwatersrand University Press.

Cingermayer, J., and B. Wood (1995). "Disentangling Patterns of State Debt Financing." *American Political Science Review* 89(1): 108-120.

Conradie, S. R., and L. J. M. Messerschmidt (2000). *A Symphony of Power: The Eskom Story*. Johannesburg, Chris van Rensburg Publications.

Converse, N., and E. Kapstein (2006). "The Economics of Young Democracies: Policies and Performance." MPRA Paper 553.

Corruption Watch (2013). "Understanding Tender Corruption." October 15, 2013. Downloaded from <http://www.corruptionwatch.org.za/print/959> (last accessed February 1, 2014).

Cox G. (2010). "Swing voters, core voters and distributive politics." In Shapiro, I., S. C. Stokes, E. J. Wood, and A. S. Kirshner (Ed.) in *Political Representation*, Cambridge, Cambridge University Press: 342-357.

Cox, G. W., and M. D. McCubbins (1986). "Electoral Politics as a Redistributive Game." *Journal of Politics* 48(2): 370-389.

Croft, T.A. (1978). "Nighttime images of the earth from space." *Scientific American* 239: 68-79.

Croft, T.A. (1979). "The brightness of lights on earth at night, digitally recorded by DMSP satellite." Stanford Research Institute final report prepared for the U.S. Geological Survey, SRI, Stanford, CA.

Dahlberg, M., and E. Johansson (2002). "On the Vote-Purchasing Behavior of Incumbent Governments." *American Political Science Review* 96(1): 27-40.

Davis, M. (1997). "Financial Analysis of the Electrification Programme." *Development Southern Africa* 14(1): 97-120.

Davis, M., and G. Steyn (1998). *Electricity in South Africa*. London, Financial Times Energy.

De Haan, J., and J. Klomp (2009). *Political Budget Cycles and Election Outcomes*. CES Institute for Economic Research, Munich, Germany.

De Souza Filho, C.R., J. Zullo, and C. Elvidge (2001). "Brazil's 2001 energy crisis monitored from space." *International Journal of Remote Sensing* 25 (12): 2475-2482.

Department of Housing (1997). White Paper - A New Housing Policy and Strategy for South Africa. Downloaded from <http://www.housing.gov.za/content/Doc/Housing%20White%20Paper.pdf> (last accessed June 19, 2012).

Department of Minerals and Energy (2001). *National Electrification Programme (NEP) 1994-1999: Summary Evaluation Report*. Pretoria, Department of Minerals and Energy.

Diaz-Cayeros, A., F. Estevez, and B. Magaloni (2013). *Strategies of Vote-Buying: Poverty, Democracy, and Social Transfers in Mexico*. Cambridge, Cambridge University Press.

Dinkelman, T. (2011). "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101(7): 3078-3108.

Dixit, A., and J. Londregan (1996). "The Determinants of Success of Special Interests in Redistributive Politics." *Journal of Politics* 58(4): 1132-1155.

Dixit, A., and J. Londregan (1998). "Ideology, Tactics, and Efficiency in Redistributive Politics." *Quarterly Journal of Economics* 113(2): 497-529.

Doll, C.N. (2008). "CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications." Center for International Earth Science Information Network of Columbia University, Palisades, NY.

Doll, C.N. (2010). "Population detection profiles of DMSP-OLS night-time imagery by regions of the world." *Proceedings of the 30th Asia-Pacific Advanced Network Meeting*, 191-207.

Doll, C.N., and S. Pachauri, (2010). "Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery." *Energy Policy* 38: 5661-5670.

Doll, C.N., J-P. Muller, and C.D. Elvidge (2000). "Nighttime imagery as a tool for global mapping of socioeconomic parameters and greenhouse gas emissions." *Ambio* 29: 157-162.

Doll, C.N., J. Muller and J.G. Morley (2006). "Mapping regional economic activity from night-time light satellite imagery." *Ecological economics* 57: 75-92.

Dombroski, K. (2003). "South Africa After Apartheid." *Journal of Democracy* 17(3): 43-57.

Downs, A. (1957). *An Economic Theory of Democracy*. New York, Harper and Row.

Drazen, A. (2000). *The Political Business Cycle after 25 Years*. NBER Macroeconomic Annual. Cambridge, MIT Press.

Dube, A., T. W. Lester, and M. Reich (2010). "Minimum wage effects across state borders: Estimates using contiguous counties." *The Review of Economics and Statistics* 92(4): 945-964.

Duranton, G., L. Gobillon, and H. G. Overman (2011). "Assessing the Effects of Local Taxation using Microgeographic Data." *The Economic Journal* 121(555): 1017-1046.

Eldridge, M., and J. Seekings (1996). "Mandela's Lost Province: The African National Congress and the Western Cape Electorate in the 1994 South African Elections." *Journal of Southern African Studies* 22(4): 517-540.

Elections Task Group (1996). *Local Government Elections in South Africa 1995/1996*.

Elvidge, C.D., F.C. Hsu, K.E. Baugh, and T. Ghosh, (2013). "National Trends in Satellite Observed Lighting: 1992-2012." in Global Urban Monitoring and Assessment Through Earth Observation, CRC Press.

Elvidge, C. D., K.E. Baugh, S.J. Anderson, P.C. Sutton, and T. Ghosh (2012). "The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data." Social Geography 7(1): 23-35.

Elvidge, C.D., K.E. Baugh, P.C. Sutton, B. Bhaduri, B.T. Tuttle, T. Ghosh, D. Ziskin, and E.H. Erwin (2010). "Who's In The Dark: Satellite Based Estimates Of Electrification Rates" in Xiaojun Yang (Ed.) Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment, Chichester, Wiley-Blackwell.

Elvidge, C. D., M. L. Imhoff, K. E. Baugh, V. R. Hobson, I. Nelson, J. Safran, J. B. Dietz, and B. T. Tuttle (2001). "Night-time lights of the world: 1994-1995." ISPRS Journal of Photogrammetry and Remote Sensing 56(2): 81-99.

Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis (1997). "Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption." International Journal of Remote Sensing 18(6): 1373-1379.

Faal, E. (2007). "Political Budget Cycles in Papua New Guinea." IMF Working Paper 219.

Fanoe, J. A. (2011). "Lessons from Census Taking in South Africa: Budgeting and Accounting Experiences." *Journal Statistique Africain* 13: 82-109.

Fedderke, J., J. Lutz, and P. Perkins (2005). "An Analysis of Economic Infrastructure Investment in South Africa." *South African Journal of Economics* 73(2): 211-228.

Ferree, K. (2004). "The Microfoundations of Ethnic Voting: Evidence from South Africa." *Afrobarometer Working Paper* 40.

Fjeldstad, O. (2004). "What's Trust Got to Do With It? Non-payment of Service Charges in Local Authorities in South Africa." *Journal of Modern African Studies* 42(4): 539-562.

Folscher, A., Krafchik, W., and I. Shapiro (1999). "Fiscal Transparency and Participation in the Budget Process – South Africa Country Report." Idasa Budget Information Service.

Franzese, R. (2002). "Electoral and Partisan Cycles in Economic Policies and Outcomes." *Annual Review of Political Science* 5(1): 369-422.

Friedman, S. (2009). "An Accidental Advance? South Africa's 2009 Elections." *Journal of Democracy* 20(4): 108-122.

Fujiwara, T. (2013). "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil." Unpublished paper. Princeton University.

Gaunt, C. T. (2003). Electrification Technology and Processes to Meet Economic and Social Objectives in Southern Africa. Thesis Presented for the Degree of Doctor of Philosophy, Department of Electrical Engineering, University of Cape Town.

Gaunt, C. T. (2005). "Meeting Electrification's Social Objectives in South Africa, and Implications For Developing Countries." *Energy Policy* 33(10): 1309-1317.

Gaunt, T., M. Salida, R. Macfarlane, S. Maboda, Y. Reddy, and M. Borchers (2012). Informal Electrification in South Africa: Experience, Opportunities and Challenges. Cape Town, Sustainable Energy Africa.

Ghosh, T., R. Powell, C. Elvidge, K. Baugh, P. Sutton, and S. Anderson (2010a). "Shedding Light on the Global Distribution of Economic Activity." *The Open Geography Journal* 3: 148-161.

Ghosh, T., C. Elvidge, P. Sutton, K. Baugh, D. Ziskin, and B. Tuttle (2010b). "Creating a global grid of distributed fossil fuel CO₂ emissions from nighttime satellite imagery." *Energies* 3(12): 1895-1913.

Ghosh, T., S. Anderson, C. Elvidge, and P. Sutton (2013). "Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being." *Sustainability* 5(12): 4988-5019.

Golden, M., and B. Min (2013). Distributive politics around the world. *Annual Review of Political Science* 16: 73-99.

Gonzales, M. (2002). "Do Changes in Democracy Affect the Political Budget Cycle? Evidence from Mexico." *Review of Development Economics* 6(2): 204-224.

Grier, R., and K. Grier (2000). Political Cycles in Nontraditional Settings: Theory and Evidence from the Case of Mexico. *Journal of Law and Economics* 43: 239-263.

Harding, R., and D. Stasavage (2014). "What Democracy Does (and Doesn't) do for Basic Services: School Fees, School Inputs, and African Elections." *The Journal of Politics* 76(1): 229-245.

Henderson, J. V., Storeygard A., and D. N. Weil (2011). "A bright idea for measuring economic growth." *American Economic Review Papers and Proceedings* 101(3): 194-199.

Henderson, J. V., Storeygard A., and D. N. Weil (2012). "Measuring economic growth from outer space, American Economic Review." 102(2): 994-1028.

Hibbs, D. (1977). Political Parties and Macroeconomic Policy. *American Political Science Review* 71(4): 1467-1487.

Hollyer, J. R., B. P. Rosendorff, and J. R. Vreeland (2011). "Democracy and Transparency." *Journal of Politics* 73(4): 1191-1205.

Holmes, T. J. (1998). "The effect of state policies on the location of manufacturing: Evidence from state borders." *Journal of Political Economy* 106(4): 667-705.

Humphreys, M. and R. Bates (2005). "Political Institutions and Economic Policies: Lessons from Africa." *British Journal of Political Science* 35(3): 403-428.

Husted, T. A., and L. W. Kenny (1997) "The Effect of the Expansion of the Voting Franchise on the Size of Government." *Journal of Political Economy* 105(1): 54-82.

IDASA (2004). "The Local Government Grant System in South Africa. Budget Information Service." *Budget Brief* 138.

Imhoff, M. L., W. T. Lawrence, C. D. Elvidge, T. Paul, E. Levine, M. V. Privalsky, and V. Brown (1997). "Using nighttime DMSP/OLS images of city lights to estimate the impact of urban land use on soil resources in the United States." *Remote Sensing of Environment* 59(1): 105-117.

Independent Electoral Commission (2012). Results of Past Elections. Downloaded from <http://www.elections.org.za/Elections94.asp>. (last accessed March 10, 2013).

Ishiyami, J. and J. Quinn (2006). "African Phoenix? Explaining the Electoral Performance of the Formerly Dominant Parties in Africa." *Party Politics* 12(3): 317-340.

Jacobs, D., N. Schoeman, and J. van Heerden (2002). "Alternative Definitions of the Budget Deficit and its Impact on the Sustainability of Fiscal Policy in South Africa." *South African Journal of Economics* 70(3): 544-559.

Jerven, M. (2013). Poor numbers: how we are misled by African development statistics and what to do about it. Ithaca, Cornell University Press.

Johansson, E. (2003). Intergovernmental Grants as a Tactical Instrument: Empirical Evidence from Swedish Municipalities. *Journal of Public Economics* 87: 883-915.

Johnson, R. (1996). The 1994 Election: Outcome and Analysis. In Launching Democracy in South Africa: The First Open Election, April 1994, Johnson, R. and Schlemmer, L. eds. New Haven, Yale University Press.

Johnson, R. W. (2009). South Africa's Brave New World: The Beloved Country Since the End of Apartheid. New York, Allen Lane.

Khemani, S. (2004). Political Cycles in a Developing Economy: Effect of Elections in the Indian States. *Journal of Development Economics* 73: 125-154.

Khemani, S. (2007). "Does delegation of fiscal policy to an independent agency make a difference? Evidence from intergovernmental transfers in India." *Journal of Development Economics* 82(2): 464-484.

Khumalo, B. and G. Rao (2004). Sharing the Cake: A Review of Provincial Equitable Share Formula in South Africa. Financial and Fiscal Commission Conference Papers, South Africa.

King, G. (2010). "A Hard Unsolved Problem? Post-Treatment Bias in Big Social Science Questions." Presentation at the "Hard Problems in Social Science" Symposium, Harvard University, April 10.

Kneebone, R. and K. McKenzie. (2001). Electoral and Partisan Cycles in Fiscal Policy: An Examination of Canadian Provinces. *International Tax and Public Finance* 8: 753-774.

Kramon, E., and D. N. Posner (2013). "Who benefits from distributive politics? How the outcome one studies affects the answer one gets." *Perspectives on Politics* 11(2): 461-474.

Kudamatsu, M. (2012). "Has Democratization Reduced Infant Mortality in Sub-Saharan Africa? Evidence from Micro Data." *Journal of the European Economic Association* 10(6): 1294-1317.

Kwon, H. (2005). "Targeting Public Spending in a New Democracy: Evidence from South Korea." *British Journal of Political Science* 35(2): 321-341.

Lake, D. A., and M. A. Baum (2001). "The Invisible Hand of Democracy: Political Control and the Provision of Public Services." *Comparative Political Studies* 34(6): 587-621.

Oak Ridge National Laboratory. (2014). LandScan Data Availability. Downloaded from http://web.ornl.gov/sci/landscan/landscan_data_avail.shtml (last accessed 1 February, 2014)

Lanegran, K. (2001). South Africa's 1999 Election: Consolidating a Dominant Party System. *Africa Today* 48(2): 81-102.

Larcinese, V. (2014). Enfranchisement and representation: Italy 1909-1913. Unpublished paper.

Larcinese, V., J. Snyder, and C. Testa (2013). "Testing Models of Distributive Politics Using Exit Polls to Measure Voter Preferences and Partisanship." *British Journal of Political Science* 43(4): 845-875.

Larcinese, V., L. Rizzo and C. Testa (2006). "Allocating the US Federal Budget to the States: the Impact of the President." *Journal of Politics* 68 (May): 447-456.

Levin, N., and Y. Duke (2012). "High spatial resolution night-time light images for demographic and socio-economic studies." *Remote Sensing of Environment* 119: 1-10.

Lewis-Beck, M. (1988). *Economics and Elections*. Ann Arbor, University of Michigan Press.

Lindbeck, A., and J. Weibull (1987). "Balanced Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52: 237-297.

Lindbeck, A., and J. Weibull (1993). "A Model of Political Equilibrium in a Representative Democracy." *Journal of Public Economics* 51(2): 195-209.

Lodge, T. (1999). *South African Politics Since 1994*. Cape Town, David Philip Publishers.

Lodge, T. (2006). The Future of South Africa's Party System. *Journal of Democracy* 17(3): 153-166.

Mattes, R. B. (2002). "South Africa: Democracy Without the People?" *Journal of Democracy* 13(1): 22-36.

Mattes, R. B. (1995). *The Election Book: Judgement and Choice in South Africa's 1994 Election*. Cape Town, Idasa.

Meltzer, A. H., and S. F. Richard (1981). "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5): 914-927.

Michalopoulos, S., and E. Papaioannou (2013). "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1): 113-152.

Miguel, E., and F. Zaidi (2003). "Do Politicians Reward their Supporters? Regression Discontinuity Evidence from Ghana." Working Paper, University of California, Berkeley.

Miller, G. (2008). "Women's Suffrage, Political Responsiveness, and Child Survival in American History." *Quarterly Journal of Economics* 123(3): 1287-1327.

Min, B. (2014). "Electrifying the Poor: Power and Politics in India." Revise and resubmit. *American Journal of Political Science*.

Min, B., K. M. Gaba, O. F. Sarr, and A. Agalassou (2013). "Detection of Rural Electrification in Africa using DMSP-OLS Night Lights Imagery." International Journal of Remote Sensing 34(22): 8118-8141.

Mocanu D., A. Baronchelli, N. Perra, B. Gonçalves, Q. Zhang, (2013). "The Twitter of Babel: Mapping World Languages through Microblogging Platforms." PLoS ONE 8(4): e61981.

Momoniat, I (1999). Fiscal Decentralisation in South Africa: A Practitioner's Perspective. National Treasury Intergovernmental Relations, South Africa.

Munck, G. L., and J. Verkuilen (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." Comparative Political Studies 35(1): 5-34.

National Electricity Regulator (2000). Lighting Up South Africa: A Century of Electricity Serving Humankind. Fichardt Park, South Africa, Open Hand Press.

National Oceanic and Atmospheric Administration (2013a). Data Availability. <http://ngdc.noaa.gov/eog/availability.html> (last accessed April 01, 2013).

National Oceanic and Atmospheric Administration (2013b). Version 4 DMSP-OLS Nighttime Lights Time Series. <http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html> (last accessed April 01, 2013).

Nichols, A. and M. Schaffer. (2007). Clustered Standard Errors. In Stata Users Group, United Kingdom Stata Users' Group Meetings 2007(07).

Nordhaus, W. (1975). "The Political Business Cycle." *The Review of Economic Studies* 42(2): 169-190.

Open Budget Index (2012). Open Budget Index 2012 – Key Findings and Recommendations.

Downloaded from <http://internationalbudget.org/wp-content/uploads/OBI2012-Report-English.pdf> (last accessed February 10, 2014).

Paldam, M. (1997). "Is There an Electoral Cycle? A Comparative Study of National Accounts." *Scandinavian Journal of Economics* 81: 323-342.

Pandey, B., P.K. Joshi, and K.C. Seto (2013). "Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data." *International Journal of Applied Earth Observation and Geoinformation* 23: 49-61.

Perotti, R., and Y. Kontopoulos. 2002. Fragmented Fiscal Policy. *Journal of Public Economics* 86: 191-222.

Persson, T., and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*, Cambridge, MA: MIT Press.

Persson, T., and G. Tabellini (2003). Do Electoral Cycles differ across Political Systems? CEPR Working Paper Series 232.

Pinkovskiy, M.L. (2011). Economic Discontinuities across Borders: Evidence from Satellite Data on Lights at Night." Mimeo, Department of Economics, MIT.

Qase, N., T. Gaunt, H. Winkler, J. Mavhungu, M. Borchers, Y. Afrane-Okese, C. Thom (2001). National Electrification Programme Evaluation. Interim Outcomes Report: Programme Evaluations. Cape Town, Energy and Development Research Centre, University of Cape Town.

Rakner, L., and N. van de Walle (2009). Opposition Weakness in Africa. *Journal of Democracy* 20(3): 108-121.

Reinikka, R., and J. Svensson (2004). "Local Capture: Evidence From a Central Government Transfer Program in Uganda." *Quarterly Journal of Economics* 119(2): 679-705.

Reynolds, A. (1999). *Election 1999 – From Mandela to Mbeki*. New York, St Martin's Press.

Rogoff, K., and A. Sibert (1988). "Elections and Macroeconomic Policy Cycles." *Review of Economic Studies* 55(1): 1-16.

Rogoff, K. (1990). "Equilibrium Political Budget Cycles." *American Economic Review* 80: 21-36.

Rose, S. (2006). "Do fiscal rules dampen the political business cycle?" *Public choice* 128(3-4): 407-431.

Ross, M. (2006). "Is Democracy Good for the Poor?" *American Journal of Political Science* 50(4): 860-874.

SA Constitution (1996). *Constitution of the Republic of South Africa*. Downloaded from <http://www.info.gov.za/documents/constitution/1996/a108-96.pdf> (last accessed March 20, 2013).

SA Treasury (1999-2011). *Intergovernmental Fiscal Review*. Department of Finance, Pretoria.

SA Treasury (2005). *Intergovernmental System in South Africa*. Powerpoint presentation for PMIG Meeting. Johannesburg, South Africa.

Schneider, C. (2009). "Fighting with one Hand tied behind the Back: Political Budget Cycles in the West German States." *Public Choice* 142: 125-150.

Schuknecht, L. (1996). "Political Business Cycles and Fiscal Policies in Developing Countries." *Kyklos* 49: 155-170.

Schuknecht, L. (2000). "Fiscal Policy Cycles and Public Expenditure in Developing Countries." *Public Choice* 102: 115-130.

Schultz, K. (1995). "The Politics of the Political Business Cycle." *British Journal of Political Science* 25(1): 79-99.

Sen, A. (1999). "Development as Freedom." Oxford: Oxford University Press.

Shepsle, K., and B. Weingast (1981). "Political Preferences for the Pork Barrel: A Generalization." *American Journal of Political Science* 25(1): 96-111.

Shi, M., and J. Svensson (2002). "Political Budget Cycles in Developed and Developing Countries." Institute for International Economic Studies, Stockholm University.

Shi, M., and J. Svensson (2003). "Political Budget Cycles: A Review of Recent Developments." *Nordic Journal of Political Economy* 29: 67-76.

Shi, M., and J. Svensson (2006). "Political Business Cycles: Do they differ across Countries and why?" *Journal of Public Economics* 90(8-9): 1367-1389.

Small, C., and C. Elvidge (2013). "Night on Earth: Mapping decadal changes of anthropogenic night light in Asia." *International Journal of Applied Earth Observation and Geoinformation* 22: 40-52.

Small, C., F. Pozzi and C. Elvidge (2005). "Spatial analysis of global urban extent from DMSP-OLS night lights." *Remote Sensing of Environment* 96: 277-291.

Southall, R. (1994). "The South African Elections of 1994: The Remaking of a Dominant-Party State." *Journal of Modern African Studies* 32(4): 629-655.

Sparks, A. (1995). "Tomorrow is Another Country: The Inside Story of South Africa's Negotiated Revolution." London, Heinemann.

Stasavage, D. (2005). "Democracy and Education Spending in Africa." *American Journal of Political Science* 49(2): 343-358.

Stasavage, D., and D. Moyo (2000). "Are Cash Budgets a Cure for Excess Fiscal Deficits (and at What Cost)?" *World Development* 28(12): 2105-2122.

Statistics South Africa (2001). *Census 2001 Household Questionnaire*. Pretoria, Statistics South Africa.

Statistics South Africa (2005). *Achieving a Better Life For All: Progress Between Census '96 and Census 2001*. Pretoria, Statistics South Africa.

Stokes, S. C. (2005) "Perverse accountability: A formal model of machine politics with evidence from Argentina." *American Political Science Review* 99(03): 315-325.

Storeygard, A. (2012), "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa." Working Paper.

Sutton, P. (2003). "Estimation of human population parameters using night-time satellite imagery" in Mesev, V. (Ed.) *Remotely Sensed Cities*. London, Taylor and Francis: 301-334.

Sutton, P., D. Roberts, C. Elvidge, and H. Meij (1997). "A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States." *Photogrammetric Engineering and Remote Sensing* 63(11): 1303-1313.

Townsend, A. C., and D. A. Bruce (2010). "The use of night-time lights satellite imagery as a measure of Australia's regional electricity consumption and population distribution." *International Journal of Remote Sensing* 31(16): 4459-4480.

Tufte, E. (1978). *Political Control of the Economy*. Princeton, Princeton University Press.

UNECA (2011). The 2010 Round of Population and Housing Censuses in Africa. Downloaded from
http://www.uneca.org/sites/default/files/page_attachments/2010roundpopulationhousingcensusesafrica.pdf

Van Zyl, A. (1999). "A Review of Provincial Budgets, 1996/7 - 1999/000." IDASA Budget information Service, South Africa.

Van Zyl, A. (2003). "Financing the Provinces." IDASA Budget Information Service. Unpublished, South Africa.

Vernby, K. (2013). "Inclusion and Public Policy: Evidence from Sweden's Introduction of Noncitizen Suffrage." *American Journal of Political Science* 57(1): 15-29.

Wehner, J. (2000). "Fiscal Federalism in South Africa." *The Journal of Federalism* 30(3): 47-72.

Wright, G. (1974). "The Political Economy of New Deal Spending." *Review of Economics and Statistics* 56(1): 20-29.

Zhang, Q. and K. C. Seto (2011). "Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data." *Remote Sensing of Environment* 115(9): 2320-2329.

Appendix

Appendix A

Appendix Table A1: Components of the equitable share formula between 1995 and 2010

Component	Explanation
Education	Based on the average size of the provincial school-age population (ages 6-17) and the number of learners enrolled in public ordinary schools in each province.
Health	Based on the proportion of the population without access to medical aid funding in each province.
Social welfare*	Based on the estimated number of people in each province entitled to social security grants (i.e. the elderly, disabled and children) weighted using a poverty index.
Basic	Derived from each province's share of the total population of the country.
Backlog/poverty	Based on the distribution of the provincial capital needs as captured in the schools register of needs, the audit of hospital facilities and the share of the rural population.
Economic activity	Based on the provincial share of remuneration.
Institutional	Independent of data; this is divided equally among the provinces.

Note: This table is based on information from the SA Treasury Intergovernmental Fiscal Reviews 1999 - 2010. *As of 2006, the social welfare component was removed from the formula as responsibility for social security grants shifted from the provincial to the national sphere of government with the implementation of the Social Assistance Act of 2004.

Appendix Table A2: Description of variables

Variable name	Type	Unit	Construction	Source
Revenue variables:			Log of inflation adjusted rand of revenue variable divided by the number of habitants for each year between 1995 and 2010	
- Total Transfer - Equitable share - Conditional grants	Dependent	Log of Rand per capita		SA Treasury, Intergovernmental Fiscal Reviews, Statistics South Africa, Population Estimates
Pre-election	Independent	Provinces	Dummy = 1 one year before an election, 0 otherwise	Independent Electoral Commission Reports.
Vote margin (1994, 1999, 2004, 2009)	Independent	Provinces	Equals the absolute difference between the share of votes of the first and second party in the previous provincial election	Independent Electoral Commission Reports.
Vote margin (1994)	Independent	Provinces	Equals the absolute difference between the share of votes of the first and second party in the 1994 provincial elections	Independent Electoral Commission Reports.
GDP per capita	Control	Rand per capita	GDP divided by the number of habitants for each year between 1995	Statistics South Africa, 1995, 2000, 2005, 2010 Income and Expenditure Surveys

and 2010				
				IHS Global Insight supplied by Gauteng Provincial Government.
Total population	Control	Number of people	Sum of habitants	
Population < 14	Control	Number of people	Sum of habitants under the age of 14	IHS Global Insight supplied by Gauteng Provincial Government.
Population > 60	Control	Number of people	Sum of habitants over the age of 60	IHS Global Insight supplied by Gauteng Provincial Government.

Appendix B

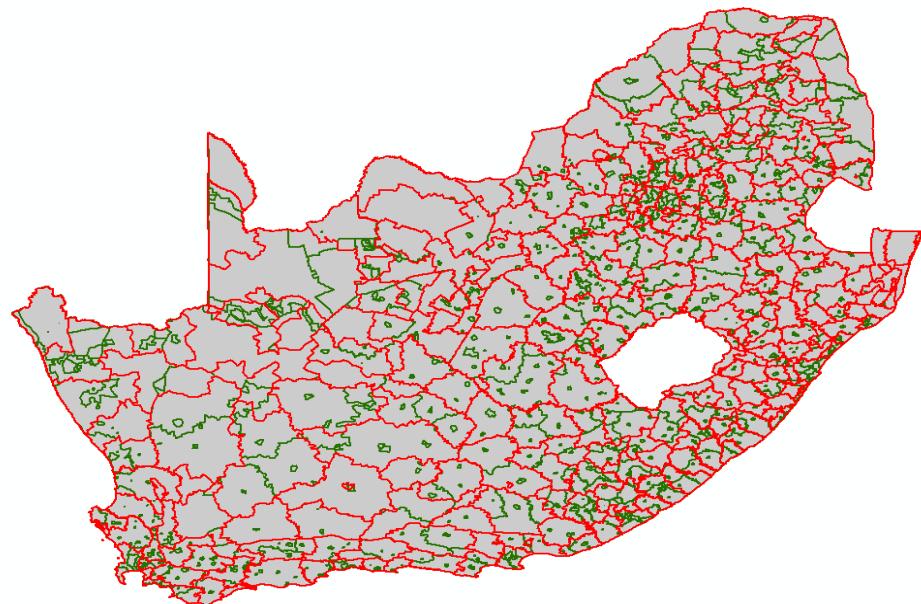
Generating a single dataset with 1995/6 election, 1996 and 2001 census data

Comparability problem. Although census data for 1996 and 2001 are readily available, comparisons over time have been rare. The main reason for this is that the overhaul of local government during the 1990s dramatically changed the geographical hierarchies relevant for the censuses. For example, the 12,852 so-called placenames (PNs) in the 1996 census were regrouped into 21,243 subplaces (SPs) in the 2001 census. Moreover, the 813 transitional authorities⁴⁰ that applied to the 1996 census were demarcated into 262 municipalities prior to the 2001 census. In this process, boundaries were redrawn and areas were renamed. The image below shows an overlay of 1996 and 2001 boundaries and clearly demonstrates how drastically the boundaries changed over time. As a consequence, the comparisons that do exist are mostly limited to the provincial level, which is the lowest geographical level that has stayed constant since 1994.⁴¹

⁴⁰ Under apartheid, municipal structures were referred to as local government. In the transition period subsequent to democratization, they were termed transitional local authorities. Since early 2000 they were referred to as municipalities. The three terms (local government, transitional authorities and municipalities) are used interchangeably to refer to local government entities.

⁴¹ Note that the number of magisterial districts stayed constant (354), however, some of the boundaries in the Western Cape were redrawn.

Appendix Figure B1: Municipal demarcation (green lines indicate 1996 and red lines 2001 boundaries)



Overview of census data. The census data stems from the 1996 and 2001 Community Profile and GIS databases purchased from Statistics South Africa. Census 1996 variables were obtained on the municipal level, i.e. the level used in the analysis. Census 2001 variables were obtained on the subplace level and then aggregated up to the 1996 municipal level boundaries as explained below. Table B1 provides details on the geographical hierarchies available for 1996 and 2001 census data.

Appendix Table B1: Geographical hierarchies in census 1996 and 2001

1996		2001	
Hierarchy	Entities	Hierarchy	Entities
Provinces	9	Provinces	9
District Councils	45	District Councils	60
Magisterial Districts	354	Magisterial Districts	354
Local Authority	799 ⁴²	Municipality	262
Place Name	12,851	Main Place	3,109
		Sub Place	21,243

Generating 2001 census data for 1996 boundaries. In order to obtain 2001 census data on the 1996 municipal level, we carried out the following steps. First, we extracted all census variables based on the smallest geographical hierarchy in the census 2001, i.e. 21,243 subplace units. Second, we merged this data file with the corresponding subplace shape file in ArcGIS based on the numerical code that uniquely identifies each subplace in both the shape file and the data file. Finally, we assigned each 2001 sub place (21,243 polygons and 267 variables) to a 1996 level municipality (799 polygons, also obtained from the StatsSA GIS database). In some cases a 2001 subplace was assigned to more than one 1996 level municipality. The following explains how this is accounted for.

⁴² The number of municipalities in the original shape file is 813 – to see why we use 799, please refer to the discussion on ‘Number of municipalities in 1996’ below.

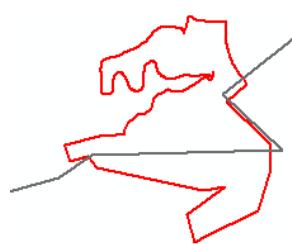
Adjusting for overlapping subplaces. One possible method to execute this step is to use ArcMap's spatial join analysis tool. The default method of this tool is called 'intersect' and assigns a polygon from the lower geographical hierarchy to the higher level, if there is an intersection between the two. This method works best when the lower level geographical units are perfectly contained within the higher-level geographical units. Most subplace polygons are smaller than the municipality polygons and are often fully contained within a higher-level polygon, as shown in Figure B2, Example 1. However, there are also several cases in which 2001 subplaces cross the border of a 1996 municipality, as shown in Figure B2, Example 2. This introduces a double counting problem because the spatial join / intersect method adds the attributes of the overlapping subplace to every municipality with which the subplace intersects. In order to take these cross-border subplaces into account, we calculated a weight based on the area size of the subplace that is covered by a municipality. For example, if a subplace has 100 individuals and 10 per cent of the subplace area falls within municipality A and 90 per cent falls within municipality B (i.e., it is cut by the municipal boundary), then 10 individuals will be allocated to municipality A by assigning a weight of 0.1 and 90 to municipality B by assigning a weight of 0.9. The key underlying assumption of this method is that the population is evenly spread throughout the subplace. Dinkelman (2011) makes the same assumption in her work on the effects of rural electrification on employment in KwaZulu-Natal.

Appendix Figure B2: Overlay of geographical hierarchies (green lines indicate 1996 municipalities and red lines 2001 subplaces)

Example 1: 2001 SPs (red) perfectly contained within 1996 municipality (grey)



Example 2: 2001 SP (red) overlaps with two 1996 municipalities (grey)



Election data. Municipal election results were obtained from the Elections Task Group (ETG), which administered the 1995/6 local government elections.⁴³ The ETG report is not available electronically and was digitized based on photocopies from an original version held by the University of Cape Town Library. The data was geocoded in ArcGIS by using the municipalities provided in the 1996 census shape file as a geo-locator. Roughly 80 per cent of municipalities were perfectly matched based on the names provided in the election report and those provided in the census shape file. The remaining 20 per cent were matched

⁴³ Subsequent elections have been managed by the Independent Electoral Commission, which was established as a permanent, independent institution by the Constitution.

manually.⁴⁴ The names in the shape file were adjusted in correspondence with the names in the electoral report in order to avoid duplication during the merger. Finally, the electoral results were assigned to the 799 municipality polygons in the 1996 census shape file.

Number of municipalities in 1996. The number of municipalities in the analysis is 799. It is important to note that the Local Government Transition Act (LGTA) in 1993 created 843 interim local government structures. Each of these 843 structures belonged to one of three municipal categories, i.e. metropolitan, local or district council. However, in terms of both geography and local government representation these structures were not uniquely identified. For example, a district council generally included more than one municipality. Our analysis focuses on municipalities that are uniquely identified in terms of both geography and representation, which is not the same as the number of local government structures created through the LGTA 1993. Notably, our analysis also includes 14 municipalities less than in the original shape file obtained from StatsSA, which shows 813 municipalities. This small discrepancy has the following reasons: one municipality (Doringberg TRC) appears twice in the census shape file, i.e. it has mistakenly been

⁴⁴ The two main reasons for the absence of a perfect match between the names provided in the electoral report and the shape file are 1) spelling discrepancies and 2) duplications. For example, the electoral report provides abbreviated versions of long names, which do not match perfectly with the complete spelling embedded in the list of municipalities in the shape file. Another common inconsistency between the two lists was that sometimes the word 'rural' forms part of the name of a transitional rural council and sometimes it did not. Duplicates included Middelberg TRC and Middelberg TLC, which are municipalities in Mpumalanga and the Eastern Cape. Richmond TLC is the name of a municipality in the Northern Cape as well as in KwaZulu-Natal.

assigned to the same polygon twice. The thirteen other municipalities⁴⁵ are shown as separate geographic entities, but are not represented by a separate local government. For example, Elukwatini appears as a local and as a rural council, but according to the election report, both areas were governed by the Elukwatini local council (see 1995/6 Election results).

Description of variables

Dependent variables:

Electricity. This variable is defined as the percentage share of households with access to electricity in a given municipality. This variable is based on the answer “electricity” to the census question: “What type of energy/fuel does this household MAINLY use for lighting?” See section B, 2.1 and H-28 in the census household questionnaires in 1996 and 2001 respectively.

Source: Statistics South Africa, 1996 and 2001 Community Profiles and GIS databases.

Nightlight. We define nightlight as the percentage share of lit pixels in a municipality. The calculations include the following steps, which we apply to the raw nighttime lights rasters for the years 1992, 1996 and

⁴⁵ 1) Badklass TRC, 2) Blaauberg MLC, 3) Eastern Johannesburg MLC, 4) Elukwatini TRC, 5) iLembe Regional Council, 6) Indlovu Regional Council, 7) Mzinyathi Regional Council, 8) Nqamakwe TLC, 9) Oostenberg MLC, 10) Renovaal TLC, 11) Vierfontein TLC, 12) Waterkloof TLC, and 13) Zululand Regional Council.

2001⁴⁶: First, we convert the stable lights image into a binary grid, where 1 represents pixels in which light is detected and 0 where no light is detected. We do so by using the ArcGIS reclassify tool. Note that there were no pixels with missing values in any of the years in the South African geographic extent. The resulting raster image is depicted in Map 2 in Figure 5. Second, we calculate zonal statistics using the municipal boundaries shape file. This produces a table in which the variable “count” represents the number of pixels in a given municipality and the variable “sum” represents the sum of lit pixels in a given municipality. The ratio of the two is therefore the share of lit pixels in a municipality, which yields our variable nightlight (see Map 3 in Figure 5).

Source: http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html

Independent variables:

Enfranchised. This variable is defined as the percentage share of enfranchised citizens in the municipal electorate and thus refers to all non-white citizens of voting age in a given municipality at the time of the first democratic local government elections. It is derived from a cross-tabulation of population group and age from the 1996 census and thus calculating the share of non-white citizens aged 18 or above. The variables *Enfranchised black*, *Enfranchised coloured*, and *Enfranchised Indian* are subsets of Enfranchised that calculate the percentage share of enfranchised citizens for the black, coloured and Indian or Asian population groups respectively. For the corresponding census questions

⁴⁶ In 2001 two satellites recorded nighttime lights (F14 and F15), in this case we use the average between the two images.

see section A, 5 and P-06 in the census household questionnaires in 1996 and 2001 respectively.

Source: Statistics South Africa, 1996 and 2001 Community Profiles and GIS databases.

ANC seat share. This variable is defined as the percentage share of total seats on a given local council won by the ANC in the 1995/6 elections. The related variables *NP seat share* and *IFP seat share* are defined as the percentage share of total seats won by the National Party and the Inkatha Freedom Party respectively.

Elections Task Group (1996). Local Government Elections in South Africa 1995/1996.

Geographic controls:

Elevation. We calculate various descriptive statistics (mean, median, max, min, standard deviation, range) to determine the elevation and slope of a given municipality. We do so by using the zonal statistics tool in ArcGIS and obtain elevation indicators in meters above sea level. To obtain these measures, we use the following projected coordinate system: "WGS 1984 UTM Zone 35S". In the paper, the variable elevation is the average municipal elevation in meters above sea level.

Source: USGS (2004). Shuttle Radar Topography Mission, 1 ArcSecond scene SRTM_u03_n008e004, Unfilled Unfinished 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000.

Slope. In order to calculate the slope of a given municipality, we perform a surface analysis in ArcGIS and generate a new raster layer. We apply

the zonal statistics tool to this raster, using the municipality shape file and thus calculate various descriptive statistics (mean, median, max, min, standard deviation, range) of slope in percent rise. These calculations are performed using the projected coordinate system: "WGS 1984 UTM Zone 35S"

Source: USGS (2004). Shuttle Radar Topography Mission, 1 ArcSecond scene SRTM_u03_n008e004, Unfilled Unfinished 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000.

Distance from electricity grid. This variable measures the shortest distance in kilometers between the centroid of a municipality and the South African high-voltage power grid in 1996. This measure is calculated using the proximity tool in ArcGIS 10 using the municipality shape file and power grid shape file, containing the geo-referenced high-voltage Eskom power lines in South Africa in 1996.

Source: Eskom (1997). Annual Report 1996. Johannesburg: Eskom.

Distance from main road. This variable measures the shortest distance in kilometers between the centroid of a municipality and the closest main road in 1996. This measure is calculated using the proximity tool in ArcGIS 10 using the municipality shape file and the shape file of main roads in 1996 included in the 1996 GIS database.

Source: Statistics South Africa, 1996 GIS databases.

Population controls:

Population density. We calculate this variable by dividing the total population in a given municipality, using 1996 and 2001 census data, by

the area of that municipality measured in square kilometers. The area is calculated using the area tool in ArcGIS 10 using the 1996 municipality shape file.

Source: Statistics South Africa, 1996 and 2001 Community Profiles and GIS databases.

Population share. We define this variable as the percentage share of the total South African population living in a given municipality.

Source: Statistics South Africa, 1996 and 2001 Community Profiles.

Socioeconomic controls:

Employment rate. This variable is calculated as the percentage share of employed people out of the total labor force in a given municipality. The number of employed people is a census variable corresponding to the questions in section A, 17 and P-19 in the census household questionnaires in 1996 and 2001 respectively. The total size of the labor force is calculated by summing the relevant labor force sub-categories.

Source: Statistics South Africa, 1996 and 2001 Community Profiles.

Low income. This variable is the percentage of individuals in the labor force earning less than 500 Rand in 1996. Since data is grouped in earning ranges and since the 1996 and 2001 census rounds use different groupings, it was not possible to match 1996 and 2001 figures. Hence, we can use the 1996 level variable in our regressions, but not the 1996-2001 difference. For the census questions corresponding to income see section A, 20 and P-22 in the census household questionnaires in 1996 and 2001 respectively.

Source: Statistics South Africa, 1996 Community Profiles.

Median income. This is the median household income in South African Rand (ZAR) in 1996 (nominal) and 2001 (real). Figures for the two years were made comparable by inflation-adjusting the 2001 data using the Consumer Price Index (CPI). CPI base 2008 (i.e., equal to 100 in 2008) is 48.6 in 1996 and 66.1 in 2001, so we deflate 2001 income by dividing by 1.36. For the census questions corresponding to income see section A, 20 and P-22 in the census household questionnaires in 1996 and 2001 respectively.

Source: Statistics South Africa, 1996 and 2001 Community Profiles.

Share of population with no schooling. This variable is the number of people with no schooling as a percentage of the total population in a given municipality. For the census questions corresponding to schooling see section A, 16.1 and P-17 in the census household questionnaires in 1996 and 2001 respectively.

Source: Statistics South Africa, 1996 and 2001 Community Profiles.

Distinguishing Eskom and municipal electricity distribution areas

The National Energy Regulator of South Africa (NERSA) was unable to supply the original list of 362 municipalities that had been granted electricity distribution licenses by the previous National Electricity Regulator (NER), which no longer exists. Instead, we obtained a 1996

membership list of the Association of Municipal Electricity Undertakings (AMEU), which identified 248 municipalities that had electricity departments at that time. We use this information to approximate a split into municipal and Eskom distribution areas. Since the AMEU data exclude some municipalities with electricity undertakings, we have about a hundred wrongly classified municipalities in the Eskom distribution subsample. Moreover, in some areas with municipal distributors, Eskom was nonetheless involved in electrification. For example, the company took on the electrification of Kayelitsha, located on the Cape Flats and now part of the City of Cape Town, through a joint venture (Qase et al. 2001). These imprecisions make it less likely that we detect differences across the subsamples.

Appendix Table B2: Descriptive statistics

Variable	a. All observations (N = 799)				b. Enfranchised non-white < 86.65 (N = 399)				c. Enfranchised non-white ≥ 86.65 (N = 400)			
	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max
ΔElectricity	9.1	19.1	-54.1	95.5	3.7	14.7	-54.1	60.2	14.6	21.3	-42.8	95.5
ΔNightlight	2.4	12.1	-36.7	100.0	1.2	9.7	-30.0	100.0	3.6	14.0	-36.7	100.0
Electricity (1996)	63.1	27.0	0.0	100.0	74.0	19.8	6.2	100.0	52.3	28.8	0.0	100.0
Nightlight (1996)	62.9	39.2	0.0	100.0	71.5	38.3	0.0	100.0	54.4	38.4	0.0	100.0
Electricity (2001)	72.3	20.0	0.5	100.0	77.7	16.5	21.2	100.0	66.9	21.7	0.5	100.0
Nightlight (2001)	65.4	38.2	0.0	100.0	72.7	37.7	0.0	100.0	58.1	37.3	0.0	100.0
Enfranchised	82.2	18.5	0.0	100.0	69.8	18.8	0.0	86.6	94.6	4.8	86.7	100.0
Enfranchised black	57.2	36.3	0.0	100.0	40.5	30.7	0.0	86.0	73.9	33.6	0.0	100.0
Enfranchised coloured	22.9	31.5	0.0	99.2	26.8	30.1	0.0	84.2	19.0	32.4	0.0	99.2
Enfranchised Indian	1.4	6.5	0.0	87.6	1.6	5.8	0.0	53.9	1.1	7.1	0.0	87.6
ANC seat share	54.7	31.0	0.0	100.0	40.7	26.6	0.0	100.0	68.7	28.8	0.0	100.0
NP seat share	17.9	24.0	0.0	100.0	25.2	26.1	0.0	100.0	10.7	19.3	0.0	100.0
IFP seat share	1.5	8.6	0.0	100.0	1.5	7.5	0.0	71.4	1.5	9.7	0.0	100.0
Distance from electricity grid (1996)	49.2	46.5	0.1	266.7	39.7	33.5	0.1	211.5	58.8	54.9	0.4	266.7
Distance from main road (1996)	19.5	23.1	0.0	201.4	18.0	23.2	0.0	134.0	20.9	22.9	0.0	201.4
Slope	1.4	0.9	0.2	4.9	1.3	0.9	0.2	4.9	1.4	0.9	0.2	4.1
Elevation	970.8	530.4	5.2	2061.6	891.9	583.0	5.2	1899.0	1049.5	459.6	24.1	2061.6
Number of households (1996)	10367.2	26375.7	4.0	265111.0	10446.7	29086.4	4.0	239313.0	10287.8	23397.5	52.0	265111.0
Population density (1996)	3797.1	17951.4	0.3	233598.8	4730.9	19378.3	0.3	233598.8	2865.7	16376.4	3.4	226593.5
Share of pop. with no schooling (1996)	20.7	12.7	0.0	77.6	16.7	12.1	0.0	77.6	24.7	12.1	1.3	61.7
Low income (1996)	38.6	18.7	0.0	96.1	34.2	19.4	0.0	90.1	42.9	16.8	8.1	96.1
Median income (1996)	16340.0	28281.5	0.0	360000.0	24228.0	37846.7	2079.0	360000.0	8471.7	6827.7	0.0	41927.8
Households without electricity (1996)	4074.0	9136.6	0.0	72214.0	2325.8	5949.1	0.0	50579.0	5817.7	11204.8	0.0	72214.0
ΔNumber of households	3663.1	16286.2	-21450.7	267563.1	4029.7	19272.3	-21450.7	267563.1	3297.3	12638.0	-6731.8	127569.1
ΔPopulation density	2503.8	8862.0	-4700.5	88052.7	3005.0	9449.7	-4700.5	79469.7	2003.9	8215.6	-503.6	88052.7
ΔShare of population with no schooling	-2.0	6.2	-43.6	30.9	-2.0	6.3	-43.6	30.9	-1.9	6.1	-26.8	28.9
ΔMedian income	1001.3	25576.9	-355473.3	89634.6	-107.8	35689.1	-355473.3	89634.6	2107.6	5948.4	-23443.5	24200.7

Appendix Table B3: Robustness to excluding municipalities with high levels of electrification in 1996

	(1)	(2)	(3)	(4)	(5)	(6)
<i>a. Direct effects</i>						
Enfranchised	0.375*** (0.063)	0.475*** (0.093)	0.522*** (0.110)	0.658*** (0.132)	0.905*** (0.233)	0.813*** (0.278)
<i>b. Conditional effects</i>						
Enfranchised	0.130 (0.081)	0.239* (0.134)	0.299* (0.152)	0.474** (0.186)	0.558 (0.396)	0.488 (0.434)
ANC seat share	-0.535*** (0.142)	-0.370* (0.214)	-0.376* (0.219)	-0.118 (0.280)	-0.439 (0.485)	-0.435 (0.589)
Enfranchised × ANC seat share	0.006*** (0.002)	0.005* (0.002)	0.005* (0.002)	0.002 (0.003)	0.005 (0.005)	0.005 (0.006)
Including municipalities with 1996 levels of electrification <	90	80	70	60	50	40
Observations	667	517	405	308	237	181

Note: The dependent variable is the percentage share of households with electricity (difference 1996-2001) calculated from census data. All regressions include a constant, province fixed effects, geographic controls, population and socioeconomic controls (1996), and households without electricity (1996). Refer to Table 5 for a description of control variables, and the data appendix for full details. The pattern of results is not affected when we vary the combination of controls. The pattern of results is not affected when we vary the combination of controls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

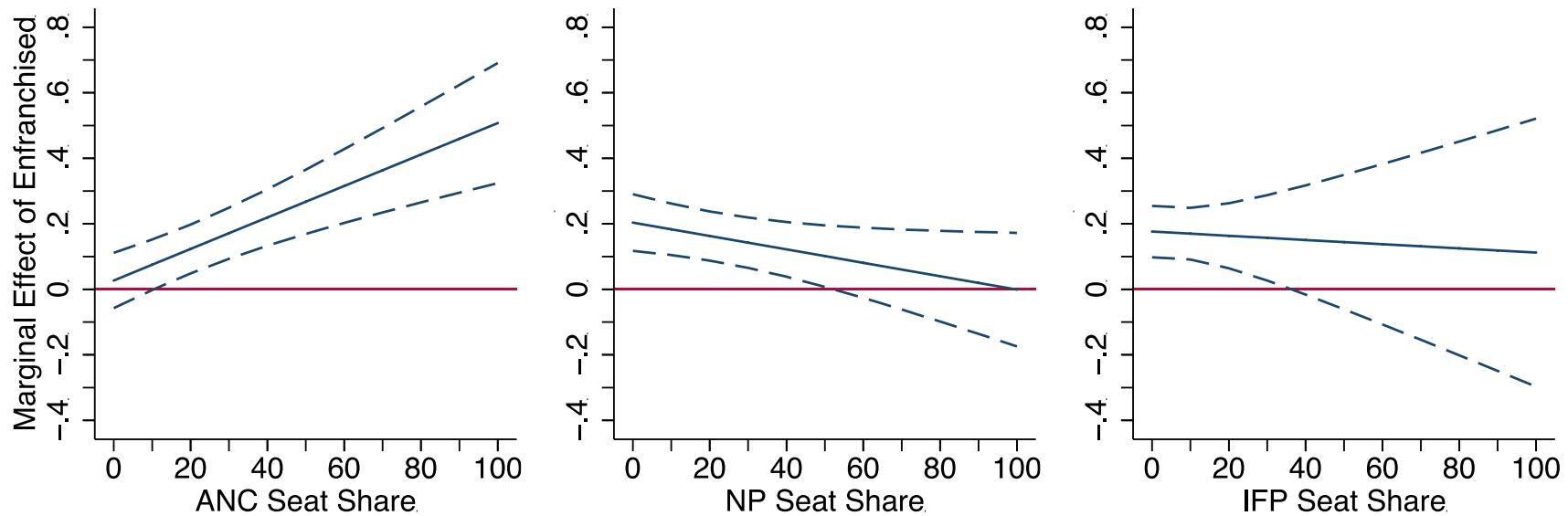
Appendix Table B4: Robustness to excluding municipalities in individual provinces

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>a. Direct effects</i>									
Enfranchised	0.255*** (0.048)	0.180*** (0.043)	0.160*** (0.044)	0.204*** (0.043)	0.182*** (0.042)	0.177*** (0.040)	0.191*** (0.040)	0.167*** (0.043)	0.208*** (0.039)
<i>b. Conditional effects</i>									
Enfranchised	0.104* (0.058)	0.028 (0.048)	0.084* (0.050)	0.035 (0.044)	0.070 (0.044)	0.041 (0.043)	0.053 (0.044)	0.034 (0.045)	0.049 (0.043)
ANC seat share	-0.379*** (0.100)	-0.358*** (0.089)	-0.218** (0.088)	-0.367*** (0.090)	-0.366*** (0.091)	-0.397*** (0.086)	-0.370*** (0.086)	-0.350*** (0.089)	-0.450*** (0.097)
Enfranchised × ANC seat share	0.004*** (0.001)	0.004*** (0.001)	0.003** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.005*** (0.001)
Sample excluding	WC	NC	EC	KZN	FS	G	L	NW	M
Observations	671	693	623	732	704	754	799	746	722

Note: The dependent variable is the percentage share of households with electricity (difference 1996-2001) calculated from census data. Province abbreviations: WC = Western Cape; NC = Northern Cape; EC = Eastern Cape; KZN = KwaZulu-Natal; FS = Free State; G = Gauteng; L = Limpopo / Northern Province; NW = North West; M = Mpumalanga. All regressions include a constant, province fixed effects, geographic controls, population and socioeconomic controls (1996), and households without electricity (1996). Refer to Table 5 for a description of control variables, and the data appendix for full details. The pattern of results is not affected when we vary the combination of controls. Robust standard errors are in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Appendix Figure B3: Conditional effect with alternative measures of partisan representation



Note: The first graph is based on the results in column (5) of Table 7. The following two graphs show the results when we replace the ANC's seat share with that of the NP and the IFP, respectively, and run an otherwise identical model.

Appendix C

Interpretation of nighttime light images and definition of underlying data sources

In order to draw intelligible conclusions from comparing data derived from data sources as different as satellite imagery and census data, it is important to have a clear understanding of what each derived measure represents and where it comes from. The use of vague concepts such as ‘luminosity’ or ‘light density’ can easily lead to interpretations that are removed from the actual information contained in the nightlight images. To avoid misinterpretation, the here are definitions of the main data sources.

Nighttime lights. NOAA’s Earth Observation Group processes the nighttime light images acquired by the DMSP OLS and releases them in their digital archive, which is publicly available online and spans 21 years, from 1992-2012.⁴⁷ The images containing average, visible, stable lights are NOAA’s “flagship product”; they are also the most relevant for calculating electrification rates as they depict anthropogenic lights from places with persistent lighting and exclude ephemeral events, such as fires. Each of the 30 arc second grids contains over 700 million pixels, equivalent to nearly global coverage spanning from -180 to 180 degrees

⁴⁷ For each satellite-year, NOAA makes three image types available for download: a cloud-free coverage image, one which contains the average of the visible band digital number values with no further filtering and a cleaned up average visible, stable lights image. Further information is provided in the readme file that is included in each download as well as in Doll (2008).

longitude and -65 to 75 degrees latitude. Each pixel represents slightly less than 1 square kilometre on the earth's surface and has a value between 0 and 63. This integer is referred to as the digital number (DN) and indicates different levels of luminosity. The higher the digital number in a given pixel, the higher the luminosity.⁴⁸ In this context, luminosity refers to the level of brightness of the observed anthropogenic or manmade lights. Grid cells without average, stable light observations have the value 0, which makes areas without lights clearly identifiable. This is the context in which I refer to a pixel or an area being unlit. Further general information about the data is provided in the readme text file accompanying each download of the data. Doll's aforementioned guide (2008) provides additional technical information about how the sensors work and how the data are collected.

Gridded population data. Even if we assume that the absence of light in a given area indicates the lack of electricity in that area, this observation alone does not allow any direct inference about whether or not people are affected by this lack of electricity. To generate per capita measures, it is possible to combine nighttime light imagery with population data, ideally using the same resolution. This enables differentiation between pixels that are unlit and unpopulated, such as forests and lakes, and pixels that are unlit, but populated. More specifically, it allows one to estimate the percentage of population living in lit pixels, as shown in Elvidge et al. (2010).

The US Department of Energy's Landscan data, hereafter referred to as Landscan, can be used to combine information about where people live

⁴⁸ Luminosity is not directly proportional to the digital number, but it is proportional to the square root of the digital number cubed: $Luminosity \propto DN^{2/3}$ (Doll 2008: 9).

with the presence of nighttime lights. Landscan data are constructed using an algorithm to disaggregate census counts within an administrative boundary to generate population counts for a resolution of 1 square kilometre. Unlike census data this dataset aims to represent the ambient population. It measures daytime counts in terms of 24h averages, thus showing populations on roads and other public areas such as airports where people work, but no one lives. Subnational-level Landscan population counts may differ considerably from census data, particularly in regions where daytime activity is different from nighttime activity. Unlike nighttime lights, Landscan data are not available free of charge for non-US government agencies. Data requests for the latest available data, which is currently 2011 data, can be submitted via email and “license fees are determined on a case-by-case basis” (Oak Ridge National Laboratory 2014). For applications in combination with nighttime lights data see Elvidge et al. (2010) or Ghosh et al. (2010a).

The Center for International Earth Science Information Network (CIESIN) also publishes gridded population datasets that can be used in combination with nighttime lights. While these data are free of charge, the standard gridded population product, the Gridded Population of the World (GPW) is only available for a resolution of 2.5' (approximately 4.5 square kilometres at the equator). For the purpose of calculating nighttime lights, this is problematic because the less-fine-grained resolution of the GPW dataset does not match the resolution of the nighttime lights dataset. Moreover, the datasets are largely derived from an earlier version of nighttime lights (Doll 2010: 194). In this context, another problem with the GPW is that the aim of this dataset is global breadth, rather than local accuracy, which may make it less appropriate for use in local-level analyses. Doll and Pachauri (2010) compare GPW

and Landscan data in relation to nighttime lights and show that there are vast differences between observable populations in different parts of the world, depending on which data source is used.

Overall, Landscan data are generally recognised as the most accurate gridded population dataset. They also match the resolution of nighttime lights images in all available years and are used in Elvidge et al. (2010), which is a reference point for most social science applications. I therefore draw on Landscan data for the construction of population based measures of socioeconomic outcomes in the subsequent analysis.

Census data. The method for collecting census data is in many ways the opposite of remote sensing. The aim of a census is generally to collect information about the universe of a population in a given country. Traditionally, census takers would aim to visit and complete surveys for every registered dwelling in a country. These efforts are usually coordinated by a national statistics office. Availability is usually restricted to one census per decade and the resolution depends on each country's enumeration methodology. The sophistication of statistics offices varies greatly across countries and statistical capacities tend to be particularly poor in sub-Saharan African countries (Jerven 2013). This limits the comparability of census data and yet they generally present the only available data source of household level characteristics for international comparisons.

With respect to electricity, census data are the main data source across countries because it is usually the most accurate if not the only data source on household characteristics. The IEA's World Energy Outlook derives most country national statistics from census data. In the absence

of a consistent definition of access to electricity across countries, the IEA's database is constructed using the definitions of each respective national government (Doll and Pachauri 2010: 5662). From the point of view of achieving consistency across countries, remote sensing technologies therefore offer a clear advantage. The same logic applies to other socioeconomic indicators such as income or educational attainment, which are published on an annual basis in the World Bank's development indicators.

In South Africa, there are three available censuses for the years 1996, 2001, and 2011, in each of which there are three possible variables that are relevant for electrification. All three are derived from the following question on the census surveys: "What type of energy/fuel does this household mainly use for cooking, for heating and for lighting?" In addition to electricity, there are eight other possible answers to this question: gas, paraffin, wood, coal, candles, animal dung, solar, or other (Statistics South Africa 2001: A11). Based on the phrasing of the question, the answer 'electricity' captures use rather than access to electricity (Statistics South Africa 2005: 144). A household may have access to electricity, but not use it. The difference between access and use is likely to be particularly pronounced where households can use a suitable substitute for electricity, for example by cooking with firewood rather than an expensive electric stove. Cooking with energy sources other than electricity, such as gas, may also simply be a preference and thus not indicative of a lack of access to electricity. This distinction is important because energy consumption patterns are not only a function of availability, but also preferences. When we interpret data in empirical analyses, it is important to do so with a good understanding of where the data comes from and what precisely it entails.

Appendix Table C1: Selected empirical applications of nighttime light satellite imagery in descending order

Paper	Year	Socioeconomic outcome studied	Level of analysis	Regional focus	Nighttime light variable	Journal/Publication
<i>a. Social science applications</i>						
Min	2014	Electrification	Subnational	India	Dichotomous variable: 0 if DN at village center equals 0, 1 otherwise	American Journal of Political Science (revise and resubmit)
Storeygard	2014	City-level economic activity	Subnational:	Sub-Saharan Africa	Digital number	Unpublished manuscript
Michalopoulos and Papaioannou	2013	Local economic activity	Subnational: pixel	Sub-Saharan Africa	Average digital number	Econometrica
Pinkovskiy	2013	Economic activity	Subnational	Global	Digital number	Unpublished manuscript
Alesina et al.	2012	Income per capita / spatial inequality	Subnational: districts and ethnic homelands	Sub-Saharan Africa	Average digital number	CEPR Discussion Paper
Henderson et al.	2011	Real income growth	National	Global	Digital number	American Economic Review
<i>b. Remote sensing and geoscience applications</i>						
Min et al.	2013	Rural electrification	Subnational	Senegal and Mali	Digital number	International Journal of Remote Sensing
Elvidge et al.	2012	Human development	Subnational	Global	Various	Social Geography
Elvidge et al.	2010	Electrification	National and Subnational	Global	Share of population living in lit pixels	Book chapter
Doll and Pachauri	2010	Rural electrification	Subnational	Global	Digital number	Energy Policy
Agnew et al.	2008	Electrical infrastructure	Subnational	Iraq	Digital number	Environment and Planning
Doll et al.	2006	Regional economic activity	Subnational	Global	Digital number	Ecological Economics
Doll et al.	2000	Socioeconomic parameters and greenhouse gas emission	National and subnational	Global	na	Ambio

Appendix Table C2: Tracing changes in socioeconomic outcomes over time

Census-based indicator	Digital number (total)	Digital number (mean)	Share of pop. in lit pixels	Share of lit pixels
(1)	(2)	(3)	(4)	(5)
<i>a. Change in electrification (1996-2011)</i>				
Census: share of hhs with access to electricity	1.00			
Digital number (total)	0.08	1.00		
Digital number (mean)	0.11	0.20	1.00	
Share of population in lit pixels	0.12	0.11	0.18	1.00
Share of lit pixels	0.18	0.10	0.28	0.64
				1.00
<i>b. Change in income (1996-2011)</i>				
Census: share of hhs with average income	1.00			
Digital number (total)	0.10	1.00		
Digital number (mean)	0.01	0.20	1.00	
Share of population in lit pixels	0.04	0.11	0.18	1.00
Share of lit pixels	0.08	0.10	0.28	0.64
				1.00
<i>c. Change in education (1996-2011)</i>				
Census: share of pop. without any schooling	1.00			
Digital number (total)	0.11	1.00		
Digital number (mean)	0.03	0.20	1.00	
Share of population in lit pixels	0.04	0.11	0.18	1.00
Share of lit pixels	0.06	0.10	0.28	0.64
				1.00

Note: Figures calculated for South Africa's municipalities in 2011 and 1996, using South African census data and nighttime light images acquired with satellites F15 and F18. The total and average digital numbers are greyed out because of comparability concerns over time due to sensor sensitivity.