Power and politics in Africa

A research design fit for purpose

David Booth

Discussion Paper

No. 3 Sept, 2008

Published on behalf of the Africa Power and Politics Programme (APPP) by the Overseas Development Institute, 111 Westminster Bridge Road, London SE1 7JD, UK (www.odi.org.uk).

The APPP Discussion Paper series is edited by Richard Crook, Director, Institute of Commonwealth Studies, School of Advanced Study, University of London, 27-28 Russell Square, London WC1B 5DS, UK (richard.crook@sas.ac.uk).

The Africa Power and Politics Programme is a consortium research programme funded by the UK Department for International Development (DFID), with additional support from the Advisory Board for Irish Aid (ABIA), for the benefit of developing countries. The views expressed in this publication are those of the author and not necessarily those of DFID, the ABIA or the Programme as a whole.

DFID Department for International Development

The Advisory Board for Irish Aid
An Bord Comhairleach do Chúnamh Éireann

A research design fit for purpose

David Booth

Contents

Figures and text boxes

Summary

This paper is one of a number of 'think pieces' commissioned during the nine-month design phase of the Africa Power and Politics Programme (APPP). Apart from an initial reaffirmation of the central research problem set out in the programme proposal, it is concerned with issues of method. The paper addresses the challenges implied by attempting to construct a unitary research design for a large multi-disciplinary consortium many of whose members are now eager to have hypotheses defined and issues resolved. It draws on a reading and re-reading of a significant literature on qualitative research methods in comparative politics, sociology and anthropology.

The programme's research problem is whether experience suggests that variants on Africa's existing stock of hybridised institutions might, if more widely adopted, produce better development results, particularly for poor people. The word 'institutions' is used in the established technical sense. Therefore, although the programme will include research on specific organisations, it is not concerned in a narrow way with improving organisations through better management. The focus is rather on the rules and norms that provide the context for organisational development. Having restated in this way what the programme's problem is, the paper asks: What sort of problem is it? How can these questions be answered? What kinds of outcomes should interest us? What kind of explanatory concepts do we need? And what empirical approaches are required?

A revisiting of the literature confirms that our problem is of a kind that is quite typical in exploratory research programmes that are taking the first steps in theory development. At this stage, we do not have well-honed research questions or refined propositions. However, this is perfectly defensible and even necessary in view of the kind of contribution to knowledge we are seeking to make. A proper appreciation of the role of empirical research in theory generation and the different stages in the life cycle of social science theories helps to make this point. The paper also discusses several aspects of the type of theory we are jointly seeking to develop, and finds no inconsistency across our disciplines or between academic researchers and policy makers in this regard.

The approach the programme has been taking in identifying topics for empirical study and thus selecting cases for case studies also stands up to scrutiny in the light of the best of the relevant methods literature. Both on the issue of case-study selection and bias, and on the matter of the number of cases needed for worthwhile comparative analysis, the literature is more diverse and sophisticated than might appear at first sight. It also enables us to dismiss the idea that comparative politics and anthropology are bearers of completely distinct traditions of case-study analysis. The range of types of causal inference is large on both sides of this disciplinary divide. Moreover, there is a wide consensus that the choice of cases, and the sequencing of other decisions about research design, needs to be highly responsive to the particular questions or propositions that are being pursued and their relations to existing knowledge about the relevant populations.

The paper responds to consortium members' hunger for 'definition' in regard to both outcome variables and explanatory concepts. In respect of outcomes, the case for adopting a programme-wide focus on the production of public goods is reaffirmed, with an emphasis on the need to distinguish between the theoretical concepts deployed in thinking about outcomes and their operationalisation. The paper argues additionally that the programme may find it useful to address the extensive literature on collective action problems in analysing the proximate causes of changes, or differences, in the under- or over-production of particular sorts of public goods. The assumptions about rational choice made by practitioners of this type of political economy may be too unrealistic to generate valid accounts of behaviour. Nonetheless, their work may provide a useful counterpoint in our efforts to generate better explanations. In respect of explanatory concepts, the paper draws attention to what is involved

in defining a theoretical concept. It argues that disagreements about the meanings of words may need to be treated as pointers to important empirical or theoretical questions, rather than as mere confusions to be settled by definition.

A research design fit for purpose

David Booth[*]

The design phase of the Africa Power and Politics Programme (APPP) was completed at a workshop held in Cape Town on 14-16 May 2008. Launched in July 2007, the programme committed the bulk of its first year to a series of activities – brainstormings, country scoping missions and debate around commissioned 'think pieces' – which were intended to feed into the necessary decisions about priorities and approaches for the forthcoming four years. This paper was written in April 2008 preparation for the Cape Town workshop. It was meant to set the scene, structure the agenda of discussion and, in particular, suggest realistic limits for the participants' expectations. It is addressed to the workshop participants. It can, however, also be read as an essay on the challenges likely to be faced by any consortium research programme starting, as the APPP does, with the ambition of addressing in a coherent way an important but only roughly formulated research hypothesis.

1 Introduction

On some issues, the APPP needs consensus. On others, it doesn't. Some points need to be agreed, or previous agreements reaffirmed. Others do not. It is important in the lead-up to the Cape Town workshop to be clear about which issues fall into which of these categories, so that we do not reproach ourselves unnecessarily if the workshop outcomes are less than neat and tidy. This paper makes a set of proposals concerning where these dividing lines should be drawn. It also offers a reflection and commentary on some of the central methodological and research-design questions that have been raised in our discussions so far, especially those that have been raised repeatedly without apparent resolution. My hope is that these comments will permit the conversation to move forward, removing the need to rehearse the same points again in exactly the same terms.

The APPP has a challenging agenda. We are committed to bringing research to bear on one of the largest policy issues of the early 21st century. We aim to do this with all the benefit of an experienced multi-disciplinary, multinational and multilingual research team. It is both unavoidable and one of the most exciting features of our set-up that we need hard thinking about how best to combine these forces to meet our shared objective. If we have not reached a comprehensive consensus already, this is hardly surprising and not a matter of particular regret. On the other hand, some team members do feel anxious at this point. They are hungry for resolution of a number of recurrent topics of discussion and impatient with the lack of definition that has been a feature of our extended design phase. They need some satisfaction.

The argument of the paper has two main strands, respectively reassuring and reaffirmative. On the one hand, it sets out reasons why close collaboration across our disciplines and intellectual traditions *is* going to be feasible within this programme despite some fears to the contrary. In a number if areas, differences of approach that might have appeared to involve matters of principle turn out not to do so. In particular, monolithic disciplinary differences do not exist. There are some universal precepts in the social sciences and it should not be too hard to make these prevail.

[*] Director, Africa Power and Politics Programme, and Research Fellow, Overseas Development Institute, London (d.booth@odi.org.uk). I am grateful to Jean-Pierre Olivier de Sardan for a helpful preliminary correspondence on some of the topics covered in the paper, and to Thomas Bierschenck, Patrick Chabal, Richard Crook, Tim Kelsall, Staffan Lindberg and Ole Therkildsen for comments on previous drafts. The responsibilities remain mine.

On the other hand, not all is sweetness and light. A number of 'red herrings' have been caught in the net of some of our previous discussions within the consortium, and it is time to get rid of them. Not all of the preoccupations that have been brought to the table are about choosing methods that are fit for the purpose of investigating the research questions set out in the APPP's programme proposal. In some cases, I believe they rest on some degree of misapprehension regarding the nature of our central research problem. In others, there is insufficient appreciation of the way good social science approaches this type of challenge. In both respects, I believe we need to draw some lines in the sand at this point in order to ensure that we are not diverted from what we are committed to doing.

The paper has six substantive sections. Each tackles one of the following questions (and, more formally, one of the topics in brackets):

- What is the programme's research problem, what isn't it and why? (on originality and contribution to knowledge)
- What sort of problem is it? (on the relations between theory, research and policy)
- How can our questions be answered? (on analytical and empirical strategies)
- What type of outcomes should interest us? (on 'defining the dependent variable')
- What kind of explanatory concepts do we need? (on definitions and their place in theory formation)
- What empirical approaches do we need? (on fieldwork coordination, timing and sequencing)

The paper draws on a reading of the quite substantial recent literature on methods, especially comparative and case study methods, in sociology, anthropology and political science. The recent literature in this field makes a major contribution in clearing away various myths that remain widespread among researchers who are not methods specialists, particularly those concerning the respective strengths and interdependencies of multivariate statistical analysis, comparative research using case studies and intensive ethnography. To a lesser extent, I draw on a re-reading of much older literature on research methods and the philosophy of social science. I have taken some reassurance from the discovery that although fashions change, and the scale and quality of research output tends to improve, there is much that is remarkably constant where the fundamentals are concerned.

2 What is the programme's research problem, what isn't it and why? (on originality and contribution to knowledge)

This paper is about research design and method, not about substantive matters. I do not want to repeat what we set out in our original programme proposal and have reproduced in several documents since then. Happily, others have taken on the task of exploring some key ideas through reflections on the substantive literature (Manor, 2008; Hyden, 2008b; Kelsall, 2008). It nonetheless seems essential to underline some features of the research problem that provides the programme's *raison d'être*. Otherwise, our methological discussion will not have a sufficiently firm anchorage, and may be inclined to drift at important moments.

To anticipate what is argued in the next section, our research problem is currently expressed in the form of a series of relatively loose questions, each using slightly different language (can we identify any hybrid institutional patterns that have been functional for developmental outcomes in Africa? how might results be improved by 'working with the grain' of African societies? etc.). This looseness is a reflection of the fact that the enterprise we have taken on is not to settle a specific empirical question, but rather to generate some new policy-relevant *theory* about how power structures and institutions influence the possibility and patterning of development in sub-Saharan Africa. The difference between these two types of enterprise may not be obvious to all readers. However, it is fundamental in my view.

With that caution in mind, three substantive features of our research problem need to be underlined:

1) The objective is to discover some salient *diversity* in African institutional patterns. It is not to provide further contributions to academic literature and policy understandings on the general question of why development outcomes are hard to achieve in sub-Saharan Africa. The literature's concentration on the general question is the point of departure of our programme proposal and the whole basis for our claims to be exploring something original and potentially important to policy.

   This does not mean that we do not need to have a thorough grasp of the generic analysis and its different variants (van de Walle vs. Chabal; Bayart vs. Olivier de Sardan, etc.). However, the challenge is to identify real-world differences that underlie and may be masked by these differences of interpretation, and not to offer a new analysis of an equally generic sort. For the same reason, we should not be trying to be comprehensive. Our explanatory frameworks need to be focused on explaining patterns of difference. They will not do this well if they take on the additional task of exploring all of the factors that are relevant to explaining all of the outcomes that matter to development or to people in Africa today.

2) Salient (or relevant) diversity means productive of *outcomes* that are different in some theoretically or practically important respect. The question of what outcomes should count as important, and how we decide that question, is discussed in Section 5. Here, the point to be underlined is that the interest of the programme is in whether and under what conditions different formal and informal institutions, or combinations of the two, have different results, not the intrinsic virtues or otherwise of particular intitutional types. The point is precisely to put a question-mark over the notion that certain kinds of governance structures are intrinsically 'good' or inevitably bad.

3) The differences that are of interest are those relating to power structures and political institutions, or alternatively politics and governance, where these terms are used in technical, not common-sense, ways. A key distinction here is between *institutions* and *organisations*, and thus between institutional change and organisational development. The differences between these concepts are quite well established in the social science literature and not the subject of major disputes (see Box 1). An important source of confusion, however, is that in everyday speech, formal organisations (banks, industrial firms, government departments, etc.) are commonly referred to as 'institutions'. For this reason, some institutional analysts prefer to avoid the term altogether, and simply speak about clusters of social norms or rules governing behaviour. However, in my view this is impractical, and a better solution is simply to be clear – at least when speaking to other researchers – that we are making use of a social-science concept defined in a particular way.

   This is not just a terminological question. Nor is it one that (as I will argue in connection with some other conceptual issues in Section 6) needs to be settled later on the basis of research. Our observation of the proper distinction between organisations and institutions will affect the consistency with which we pursue the specific gap in knowledge that the programme has targeted.

Organisational analysis, organisational change and organisational capacity-development are strongly developed fields of development research and practice. There are standard propositions about the features of organisations that make them perform more or less well (e.g. Grindle, 1997). However, studies of organisational capacity development converge on the observation that – especially although not exclusively when the organisation is in the public sector – the wider institutional environment is the key determinant of possibilities for improvement (DAC, 2006; EuropeAid, 2005; Leonard, 1987; Teskey, 2005). Moreover, the exceptions tend to confirm the rule. In Africa, formal organisations outside the private sector that work well 'are typically enclaves or "protected zones" where they can hold off the pressures from informal institutions to the point that they allow for a purposive pursuit of particular policy goals' (Hyden, 2008b: 15, citing Leonard, 1991).[1]

Box 1: Some standard definitions

Institutions:

Social practices that are regularly and continuously repeated, are santioned and maintained by social norms, and have a major significance in the social structure (Abercrombie et al., 1984).

The prescriptions that humans use to organize all forms of repetitive and structured interactions including those within families, neighborhoods, markets, firms, sports leagues, churches, private associations, and governments at all scales (Ostrom, 2005: 3, 179).

The rules of the game in a society or, more formally, the humanly devised constraints that shape human interaction (North, 1990: 3).

Organisations:

A system of consciously coordinated activities of two or more persons that functions on a relatively continuous basis to achieve a common goal or set of goals (Robbins, 2004).

Groups of individuals bound by some common purpose to achieve objectives (North, 1990: 5).

Governance:

The manner in which power is exercised in the management of a country's economic and social resources for development (World Bank, 1991: 1).

The APPP will of course be studying organisations of various kinds. However, our primary task is not to add to understanding of what management styles, staff incentive systems, 'organisational cultures' and so forth make some organisations more effective than others within a given social environment. It is how the social environment might be changed in ways that would improve the aggregate or average performance of all those organisations that have a significant bearing on development outcomes.[2] The programme's *raison d'être* centres on a cluster of unresolved questions around with the way certain social and political legacies of Africa's history have affected the continent's possibilities for dynamic development – now elegantly reviewed in the Discussion Papers by Hyden (2008b) and Kelsall (2008). We are concerned with what, if anything, might be done to

[1] Correspondingly, so-called success stories and best practices 'are typically more contextually bound than [donor agencies] assume … [T]ransplantation of methods and practical experiences from one place to another is … more problematic than conventional thinking suggests' (Hyden, 2008b: 33). Cf. also Robinson (2006) on Uganda, whose findings on the role of 'insulation' in letting a limited number of organisations acquire good leadership and perform their formally designated functions correspond closely with those of more recent scoping work for the APPP by Golooba-Mutebi and Booth.

[2] I think the question of how and when particular organisational enclaves get 'protected' in such a way is an interesting question that might help in addressing our research problem. It involves the question of the room-for-manoeuvre there may be within rules of the game of running a neopatrimonial state – an institutional issue.

mitigate this problem. We want to discover ways to cultivate the positive side of the way social practices are currently structured and regulated.[3] These questions can only be understood as institutional questions.

The focus on salient institutional diversity in the above sense is what gives the APPP the potential to contribute in an important way to useful knowledge. At the same time, our resources are limited. For these reasons, we agreed at an early stage that we would construct a unitary research design focused on our central questions, rather than assembling a set of loosely related projects. The following discussion takes this commitments as its point of departure.

3 What sort of problem is it? (on the relations between theory, research and policy)

One of the reasons it is important to be clear about what the APPP's research problem is is that there are immediate implications for the types of method that are appropriate in attempting to address it. To begin with, the nature of the problem affects the way our undertaking fits into the typical life history of knowledge creation in the social sciences. A few simple observations on that subject are among the several points of clarification that seem to be needed on the broad topic of the relations between theory, research and policy in a field such as African development.

Some readers will find my treatment of these issues unduly abstract and will want to get on to the concrete design choices that concern them ('what is the dependent variable?' etc.). Others may find it elementary or pedestrian, and they may also have some justification. However, I am convinced that without this preamble we shall not be able to make wise and consistent choices about other issues that matter.

3.1 Where does theory come from?

It is a familiar feature of methodological debates in the social sciences, within as well as between the basic disciplines, that practitioners are distributed between two polar positions, usually characterised as 'deductivist' and 'inductivist' (e.g. Gerring, 2001: Ch 10). The former archetypically portray the scientific endeavour as largely concerned with deducing testable propositions from high-level propositions (theories) and organising empirical enquiries around these, often making use of large data-sets and multivariate statistical analysis as the best substitute for a fully fledged scientific experiment. The latter see the generation of new theory or the refinement of existing theories as the main task, and prefer to start from some more delimited empirical reality that can be investigated intensively (e.g. by means of a 'case study' – see Box 2).

As established by the philosophy of knowledge long ago, both purely deductive and purely inductive enquiry are impossible. On the one hand, there are no ways of approaching raw data that do not involve some prior assumptions or categorisations. On the other hand, the origin of theories has to be explained, and is not adequately explained by notions of 'inspiration' or 'insight' which imply an origin that is completely disconnected from evidence or experience. So the differences between the deductivists and the inductivists are only ever a matter of emphasis.

[3] For example, as Kelsall puts it, 'how to build on extant notions of moral obligation and interpersonal accountability to channel energies into activities that are developmental instead of predatory' (2008: 12).

The two archetypical visions are also caricatures in another important sense. In neither tradition is it seriously held that the 'facts speak for themselves'. All research involves what is nowadays usually called causal inference, where evidence is interrogated to establish the plausibilty or otherwise of a particular interpretation of the connections between things. Deductively inclined researchers do not 'prove' causal propositions, or even falsify them in a definitive way; they only shed new light on their plausibility within specifiable margins of certainty. To the extent they have theory-building and not merely historiographic ambitions, inductivists are involved in a conceptually equivalent type of exercise – inferring more or less likely general conclusions from a specific sort of evidence.

Box 2: What are case studies?

Case:

A spatially delimited phenomenon (a unit) observed at a single point in time or over some period of time. The type of phenomenon that an inference attempts to explain. (Gerring, 2007: 19)

An instance of a class of events. (George and Bennett, 2005: 17)

Case study:

The intensive study of a single case where the purpose of that study is – at least in part – to shed light on a larger class of cases (a population). (Gerring, 2007: 20)

Taste, professional inclination and other factors (e.g. competence in algebra and statistics) typically influence where researchers stand on the deductivist/inductivist continuum. But it is also the case that the relevance of the skills, procedures and perspectives generated by these different research traditions depends on the task in hand.

First, emphasis goes with subject matter. Other things being equal, it is easier to find interesting theoretical propositions to steer empirical work if the domain of one's research problem is a universal or very common phenomenon, such as some general aspect of human behaviour – market transactions, sexual partnership or voting. It is also easier to find a relevant and usable statistical data-set. At the other extreme, consider the nature of the interactions between customs officers and their clients in former colonies of France in West Africa as a research problem. This topic lends itself naturally to an interpretative style of research deploying a case study approach.[4] The researchers will necessarily come to the task with prior assumptions but they are unlikely to be able to draw on any body of formal theory, and pre-existing data of any sort will be hard to find.

The topics that have been the main focus of discussion in the literature on the methods of comparative politics – e.g., social revolutions, nationalist uprisings and democratic transitions – lie in the space between these extremes. There is relevant theory, because the subject matter is moderately generic, but there is also a lot of scope for generating new theoretical developments by means of relatively intensive, inductive exposure to the relevant evidence. As discussed further on, some of the best work on these topics is actually methodologically closer to intensive ethnography than is sometimes supposed. The subject matter of the APPP seems to me to be situated somewhere in the territory between this sort of comparative politics and the sort of largely inductive ethnography well represented – to take an example from the previous work of APPP researchers – by *Everyday Corruption and the State* (Blundo and Olivier de Sardan, 2006).

Second, in the life cycle of a given theoretical innovation there are more inductive and more deductive moments. New theories can and do spin off from innovative efforts to test old theories. However, more often they arise from a phase of exploratory research guided by no more than a 'hunch' or generalised dissatisfaction with received ways of organising the

4 '[A] proposition with a narrow scope is more conducive to case study analysis than a proposition with a broad purview, all other things being equal' (Gerring, 2007: 49).

evidence on a particular topic. I would argue that the APPP's research problem puts it squarely in the second category. One of the reasons we have half-a-dozen different formulations of the central question, using different words and metaphorical phrases (neopatrimonial, hybrid, 'with the grain', 'good enough governance', etc.) is that we are at the initial, hunch stage in the theory-building cycle. An exploratory approach, with a strong inductive element, is therefore required. The kinds of routines of causal inference that are appropriate when a theory has become well-established are absolutely not required, at least for the immediate future.

There are several reasons why this needs to be said, both clearly and boldly. The 'hunch stage' in theory building is perfectly respectable and extensively recognised in the methods literature. Almost by definition, it is the most creative and exciting. On the other hand, undergraduate research training all over the world tends to give rather short shrift to qualitative questions about theory development and research design, and to be dominated by the teaching of basic statistics, especially representative sampling approaches and confidence intervals. In some disciplines, references to Hempel or Popper, the high priests of deductivism, are balanced with citations of Glaser and Strauss – who coined the phrase 'grounded theory' in sociology – or references to the classics of anthropological fieldwork. In others they are not.[5] Those trained, even to PhD level, before the 1990s will not have benefited from the rich controversy incited by the rather statistically inspired handbook for qualitative researchers published some years ago by King, Keohane and Verba (1994).

Typical patterns of undergraduate training are one sort of worry. Others arise from quantitative (and thus deductivist) biases in the mainstream of disciplines such as political science and sociology, especially in the US, and the influence that this continues to exercise on notions about professional calibre and quality. Some other biases have their origin in the policy world, where – it is often claimed – the politicians who provide the lead are easily convinced by numbers, and not at all by unquantified claims (although, as discussed further on, this is probably an oversimplification).

In short, some consortium researchers may be forgiven for finding the whole idea of an exploratory research design both unfamiliar and unnerving. This may be a fair reflection of the way they have been exposed to ideas about good practice in research up to this point. The solution to this problem is, however, relatively simple, and involves dipping into the reference list that follows this paper, and returning on this basis to the specific requirements of the research task we have set ourselves.

3.2 What kind of theory do we need?

So far I have assumed that we know what we mean by 'theory'. This is not an assumption that it has always been easy to make in the social sciences, for at least two reasons. First, there are numerous strands of intellectual creation that are well established and have exercised considerable influence in our disciplines that do not comply with the standard meaning of 'theory' in the philosophy of science. In particular, they fall short on the criterion of being in principle testable (or to be more exact, falsifiable). Second, some doubts remain about the degree to which the specific subject matter of the social sciences – human action – is susceptible to high-level causal generalisation. To the extent that these doubts are valid, it is hardly surprising if, in general, the social sciences can show relatively little of the cumulative theoretical development that characterise the sciences of the rest of nature.

---

5 'Social science research involves a quest for new theories as well as a testing of existing theories; it is comprised of both "conjectures" and "refutations." Regrettably, social science methodology has focused almost exclusively on the latter' (Gerring, 2007: 39). Cf. George and Bennett (2005: 12, 21).

These two groups of issues do call for a brief discussion. However, I believe that we are well placed to handle both of them.

Testability

On the first, I propose we agree with Geddes (2003) in reserving the word 'theory' for high-level, conceptually deep, *causal propositions* from which testable claims can be deduced relatively easily and with definite implications for the credibility of the theory. This has two implications.

To begin with, it implies using a different word – Geddes suggests 'approach' – to refer to typical ways of going about the construction of theories, such as rational choice analysis or 'game theory'. This definition would also exclude a good part of what has constituted the field of 'sociological theory' or 'social theory' during the last 60 years, to the extent that this has been concerned with clarifying basic concepts or defining new concepts in ways that are potentially illuminating but do not involve the formulation of causal propositions.

Secondly, we should recognise that some of the perspectives that have dominated research in political science and sociology in past decades have been formulated in ways that have made them effectively untestable. The relative lack of cumulative theoretical development is a reflection of this fact. According to Geddes, the problem is principally that the level of ambition has been too high. The traditional 'big questions' underlying so-called modernisation theory and dependency theory, which dominated the teaching of development sociology and politics for several decades from the 1950s were simply too ambitious.[6] The solution proposed by Geddes is to break down the big questions, concerning compound outcomes such as 'underdevelopment' or 'democracy' into their constituent parts – the multiple processes that contribute to them, rather than the compound outcome as a whole (2003: 23, 40-88).

As Geddes shows, and the rest of the recent literature in comparative politics and historical sociology confirms, we have made progress in this regard. Debate in our disciplines does now centre on more limited propositions, and research can and does lead to particular propositions being challenged and sometimes overturned.

The only possible exception to this among the immediately relevant disciplines is (social and cultural) anthropology. The anthropological mainstream was never, it should be said, as gripped by mega-theories of development as sociology and politics were in the 1960s and 1970s. Lévi-Straussian structuralism was never more than a minority current. On the other hand, it is arguable that in the last 30 years the influence of post-structuralism (Foucault) and postmodernism (Léotard, etc.) has had comparable effects. The effects include quite a strong aversion to the project of constructing general theoretical propositions, as well as to making analysis useful in any practical sense. As a consequence, contemporary anthropology is all about 'approach' and very little about theorising in the sense in which we have defined it above.

This is a statement about the mainstream, and there are other currents. In the middle of the last century, so-called Manchester anthropologists distinguished themselves from the then mainstream 'functionalism' by being more interested in actors and their networks than in social systems. In more recent times, those identifying with that school have been more inclined to participate in the central debates in development studies, including the critique of

---

6 Their over-overambitiousness was not really simple, however. Those traditions of thinking were also influenced by meta-theoretical commitments – to organicist and system concepts – which drove them down pathways that led to vacuousness and/or circularity as well as over-generality (Booth, 1985).

dependency theory for example (Long and Long, 1992). Relatedly, they are the ones that, in Europe at least (the US story is no doubt different), have been prepared to stand up for the idea of applied anthropology and defend the notion that making research relevant to policy and practice does not have to mean compromising academic integrity.

The principal figures in this tradition – Norman Long in the UK and the Netherlands (e.g. 2001) and J.-P. Olivier de Sardan in France and Niger (e.g. 2005) – have taken a stance that is both strongly fieldwork-oriented and committed to advancing substantive propositions about livelihoods, planning approaches, etc. They have largely succeeded avoiding both the Scylla of unprincipled consultancy and the Charybdis of postmodernist 'oppositionism'. The recent writings of Mosse (2005) and Blundo (2006; 2007) suggest that this tradition is alive and well among what may come to be regarded as the next generation of 'Manchester' thinkers.

A feature of this school of thought that is relevant for our programme is that it is quite cautious about its theory-building ambitions. Among the leading thinkers, this is a reflection of what is probably an accurate assessment of the degree to which the topics typically studied are susceptible to large generalisations that are both true and important. On the other hand, among some of their followers I have found what I would characterise as residual anti-theoretical prejudices. In these circles, it has sometimes been necessary to make the argument that premature generalisation about interactive social processes is not combated effectively by going to the opposite extreme and contenting oneself with merely documenting diversity and difference. It requires that we make the effort to uncover *patterns* of diversity and make these the subject of systematic enquiry and alternative theorising.[7]

Both Geddes' plea for a medium range of ambition for theory in comparative politics, and the cautions about theory-building offered by Long and Olivier de Sardan, seem to me to speak quite directly to the agenda of the APPP. The kind of theory we need will be theory that meets Geddes' strictures on testability and also captures the distinctive subject matter of relatively localised and specific social interactions.

Causal analysis and meaning

It may be important to spend some time on a different sort of objection to substantive theorising or causal generalisation which still has some followers in certain branches of sociology, anthropology and politics. This involves a particular point of view within the debate in the social sciences and philosophy about the implications of taking human action as a subject for research.

This debate is a very old one, well represented in some ways by the opposing positions of two of the founding fathers of sociology, Émile Durkheim and Max Weber, on the conditions for being scientific about the study of social life. Durkheim notoriously asserted that social facts need to be 'treated as things' if they are to be susceptible to scientific treatment. Weber, on the other hand, characterised the sociological method as inevitably two-stranded, with causal analysis always requiring to be accompanied by apprehension of the meanings that actors themselves give to their behaviour.

The implications of the basic fact that human action has meaning – or, to be slightly more precise, is not describable or explicable without some reference to actors' intentions and beliefs – has been the subject of one of the longest-running debates in the history of social thought. It has been one of the mainstays of the philosophy of the social sciences and a recurrent theme among leading figures in sociology and anthropology. The details need not detain us, fascinating as they can be. However, it is relevant and essential for our purposes to observe that the more extreme positions in the debate have not prevailed.

7 This was one of the main themes of Booth (1994; 2003).

Thus, for example, the position adopted by Winch (1958) – that the nature of social action makes both causal analysis and cross-cultural understanding impossible – was effectively critiqued on its home territory of Anglo-American analytical philosophy in the 1960s (MacIntyre, 1970). Anthropologists have, in general, not had a problem with the idea that research involves both 'emic' and 'etic' analysis (e.g. Pelto and Pelto, 1978: Ch 4, and for a sophisticated discussion, Olivier de Sardan, 1998). They do not generally demur from Geertz's (1973) account of the nature of cross-cultural research as an interpretive enterprise in which the researcher is obliged to move between the point of view of the actor and an observer's point of view which may include causal issues going well beyond the actors' reasons and beliefs. In sociology, Giddens (1984), Hindess (1989) and Granovetter (1985) among others[8] have provided a prospectuses for the discipline that easily accommodate actors, actor-networks and social structures or institutions, and both interpretation and causal analysis, within a single framework.

The essential subject matter of the social sciences is not, therefore, an obstacle to theorising about causal processes, although the scope for useful theorising may well depend on the type of topic that is being addressed. However, there is one issue that has not quite been settled. That is the degree to which cultural differences, and specifically differences in political cultures, are an obstacle, not to understanding but specifically to theorising across countries. The argument that a universal science of comparative politics, as espoused by some of the historic leaders of the American Political Science Association, is in fact not possible because what appear to be the same behaviours and institutions *mean* different things, was famously advanced by the philospher MacIntyre (1971). It has recently been stated in greater depth by Chabal and Daloz (2006).

Unlike some of the other issues reviewed in this paper, this seems to me to be one on which we do *not* need to have agreement, at least not in advance of doing the research. Our research problem is not about patterns of difference spanning the political cultures of the US, Sweden and Tanzania. Therefore, we do not have to confront the question of whether, say, parliaments or the act of voting are sufficiently 'the same' between those countries that they are illuminated by inclusion in a single body of theorising about legislatures or voting behaviour. This is an interesting question on which I confess to not having a firm view, but it is not an essential question for the APPP.

What is more essential is whether the hybrid, syncretic or otherwise localised institutional patterns that we find in, say, Senegal are strictly comparable with those in, say, Uganda. This is a different question – for Chabal and Daloz, in particular, a very different question, since they have few qualms about treating sub-Saharan Africa as a single cultural area for the purposes of political analysis. We will eventually have to take a view on this. However, this can and should be one of the questions we are asking ourselves in the research. I suspect no one would be in a position to give a firm answer here and now even if we needed one. What we can and must ensure is that the way we do the research has enough elements of standardisation that at the end of fieldwork we are in a position on what things are 'the same' and which are not, without too much concern that our impressions on this point reflect differences in field method.

Our theory, then, needs to be both of the testable, middle range variety, and serious about the need for both emic and etic perspectives on the data and the implications for generalisation. What other characteristics does it need?

[8] This is a personal list. Others would no doubt want to add names like Bourdieu and Habermas.

Theory, research and policy

The ultimate objective of the APPP is to create the intellectual basis for a different sort of policy approach to governance for development – a better basis for public action. The question obviously arises whether this objective is in any way in tension or contradiction with what has been said up to this point. I would argue that the tensions are slight or non-existent.

The trouble with existing theory from the point of view of development policy and practice is not that it is theoretical. It is rather to do with variable features of *how* it is theoretical. One problem is the over-generality discussed earlier, and in particular overgenerality focused on explaining what is, from a policy point of view, the negative type of outcome. This gives us theory that may be a sound basis for framing overall policies and educating practitioners on 'the way things are'. But policy is interested in differences within the general pattern, and particularly differences that might be traceable to variables over which policymakers have some leverage.

We argued in our programme proposal that, from a policy point of view, most of the qualitative Africanist literature is not sufficiently fine-grained, or not systematic enough in following up differences within the overall pattern. But the problem is not just with the qualitative work. Much of the quantitative analysis about development, on the other hand, generates broad probabilistic generalisations which are also problematic from the point of view of entry points for policy, both because they are broad and because they do not specify causal mechanisms in ways that suggest levers of change (cf. George and Bennett, 2005: 265-74). Think about the policy implications of the proposition 'institutions matter', for example.[9]

The idea of a theory-practice tradeoff has little foundation. In the much-quoted phrase of Keynes, there is nothing so practical as a good theory. The reason is that policy makers, like theoreticians, operate with big ideas that have wide application. In the specific case of governance and politics in Africa, the terrain is currently occupied by one extremely general big idea that claims to have roots in research-based knowledge – 'good governance'. If we want to displace that from centre-stage or give it more nuances, we should not be too timid about the level of abstraction and conceptual depth that we bring to the task.[10]

4 How can can these questions be answered? (on analytical strategies)

We are committed to doing research in a systematic way, so that we stand a good chance of detecting patterns of institutional diversity that are associated in regular ways with differences in outcomes. We have the ambition of developing propositions and concepts that explain the causal processes responsible for these relationships and specify the conditions under which they may be expected to hold. We need to start off in an exploratory way and develop some theory before changing gear and considering how the more interesting theoretical propositions might be subjected to more comprehensive testing.

It seems clear that, at least in the exploratory phase, we shall be in the business of using case studies, including some medium-level survey work and quantitative analysis but not large statistical datasets, as the basis for our causal inferences. The rationale is well explained in recent texts:

9 In many other cases, what is being offered is not theory but method masquerading as theory. There are no clear policy implications because there is in fact no commitment to theory building. This applies, for example, to the large body of 'oppositional anthropology' alluded to earlier.

10 For a good lesson in how to make theory relevant even to quite operational policy issues for donors, see Hyden (2008a).

> 'Case studies enjoy a natural advantage in research of an exploratory nature. … It is the very fuzziness of case studies that grants them an advantage at the exploratory stage, for the single-case study allows one to test a multitude of hypotheses in a rough-and-ready way. Nor is this an entirely conjectural process. The relationships discovered among the different elements of a single case have a prima facie causal connection: they are all at the scene of the crime'. (Gerring, 2007: 39, 41)

> 'Case studies can contribute to the inductive development of typological theories[11] in the early stages of a research program by identifying an initial list of possible theoretical variables. … Theoretical arguments derived through these inductive processes must of course be subjected to further testing to prevent "overfitting" and forestall the introduction of spurious variables'. (George and Bennett, 2005: 239-40)

But what sort of case studies do we need, how many of them, and selected in what way? In this section, I recall how we have been proceeding in our design process so far, and then place this in the context the rapidly expanding literature on good practice in case selection and causal inference in case-study and small-n comparative research.

4.1 The story so far

The position that we took in the programme proposal and have been following during the design phase is that macro-level country studies do not provide a sufficiently rich empirical basis for the propositions and questions we wish to explore. We are interested in differences in development outcomes, and our universe is sub-Saharan Africa. At the macro level, some of the relevant differences have been explored through cross-country regressions, but mostly by using crude proxy measures on world-wide data. This makes clear that institutions matter, but not *which* institutions are most helpful to performance at different stages of economic growth. Intensive studies of the few partial African success stories in building development-friendly institutions at the national level have helped us to understand certain things (e.g, Acemoglu et al., 2003). Additional applications of the method of the 'analytical narrative' at macro level are being produced for other research programmes (e.g. Golooba-Mutebi, 2008a; 2008b), and we can learn from them. However, the scope for further applications is quite limited and the payoffs from repeating them would be very doubtful.

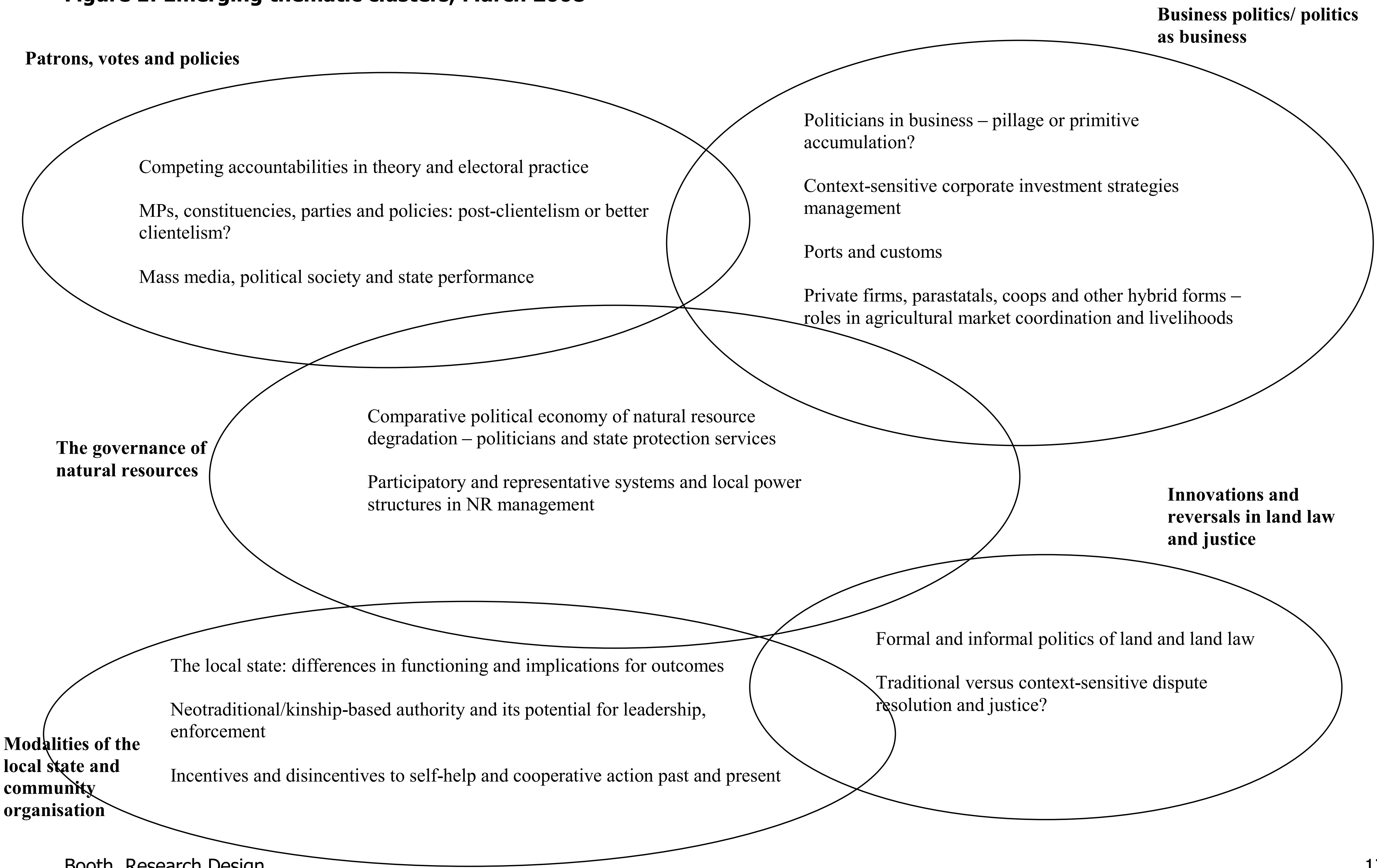
For these reasons,[12] we have been exploring which sub-national levels or spheres of institutional development in the present day or in the past would provide a richer fund of relevant 'cases'. We have identified several potential candidates, although finding even *prima facie* evidence of outcomes that are significantly better than the norm has been quite as hard as expected. This has been a moving target during the design phase of the programme, but as of the time of writing, the institutional fields (hereafter 'topics') that seem to be potentially interesting are those depicted in Figure 1.

---

[11] '[T]he development of contingent generalizations about combinations or configurations of variables that constitute theoretical types' (George and Bennett, 2005: 233).

[12] There are other reasons that might help to justify such an approach. George and Bennett favour a 'building block' approach to theory building with case studies: 'Better results are achieved if the "class" of the phenomenon to be investigated is not defined too broadly. Most successful studies, in fact, have worked with a well-defined, smaller-scope subclass of the general phenomenon. … Each block – a study of each subtype – fills a "space" in the overall theory or in a typological theory' (2005: 77-78)

Figure 1: Emerging thematic clusters, March 2008



These areas of interest have been derived in an inductive way from the knowledge that consortium researchers and key informants in selected countries have been able to bring to bear. We were guided by terms of reference for the design-phase activities which established relevance to the programme research problem as the key selection criterion. We searched for suspected differences in relevant outcomes arising at least in part from differences in institutional patterns. At the same time, the choices made so far are not uninfluenced by the topics that the researchers have worked on previously and find intrinsically interesting, and country selection is also based more on preferences and experience than on any principle of programme design.

At this point, we need to answer several questions about this, taking guidance from the literature on causal inference from cases and cross- case analysis. Questions include:

- Do five medium-level topics, or clusters of topics, provide a suitable basis for addressing the big question: what political institutions might work better for development and the poor?

- Is the number right? Or does it spread us too thinly, producing the risk that our case studies will be too few within each cluster to permit even a basic level of causal inference? How many cases are enough cases in contexts such as these?

- What is the universe? Is it sub-Saharan Africa in respect of all of these topics? Is there a case for restricting the breadth of the question and the population to which it applies?

- What sort of comparative analysis are we likely to need and/or be able to apply, assuming that we can settle (as discussed in Section 5) the question of the outcome variables to be considered?

Let us consider what the literature has to say that is relevant to the first three questions, before turning to the fourth.

4.2 Case selection: the state of the debate

An elementary but necessary preliminary is that selecting cases for study is *not* like choosing a representative sample from a statistical dataset covering a large universe. Selection is appropriately 'purposive', not random. Where the 'sample' is small or very small, as it is by definition in case study research, random selection does not ensure representativeness. Interest therefore centres on the kind of purposiveness that is to be applied. Depending on exactly what one means by it, representativeness is a secondary and possibly insignificant matter. The main concern is to obtain *causal leverage* – a selection that adequately captures the variations that are of theoretical interest (Gerring, 2007: 87-88).

Case selection and 'bias'

In one of the fullest available treatments, Gerring distinguishes at least nine techniques of purposive selection of cases taking into account the relevant aspects of representativeness and causal leverage. The one he considers of most general use, and particularly applicable where the population of cases is not (or not yet) well understood, is what he calls the 'diverse case' technique, where an effort is made to capture the full range of variation in the population. However, he also offers a strong defence of what he calls the 'extreme case' method, where – for exploratory purposes – priority is given to an outcome or type of outcome that is interesting because it is unusual in the context of a relatively well known overall pattern of outcomes (2007: 86-151).[13]

This seems highly relevant to APPP design decisions. It also needs discussion because, on the face of it, it is in conflict with some standard advice on what to avoid in case study selection.

Issues of *bias* in case study selection – that is choosing cases in ways that limit or predetermine the causal inferences that can be drawn – have been much discussed since in the early 1990s, when Geddes (1990; 2003: Ch 3) and King et al. (1994) launched parallel attacks on some of the best-known research in comparative politics and historical sociology. In particular, single and small-number case studies that are selected only from one end of the continuum of outcomes that is of interest (e.g. only countries where revolutions have occurred, and not those where they haven't) have been roundly condemned as insufficent and likely to mislead. However, a small army of recent writers on the subject have argued that matters are a good deal more complicated than these critics allow.

According to this rebuttal, the standard critique rests too heavily on the model of causal inference, and thus dangers of bias, that apply in large-N research and regression analysis. In fact, there are different bases for causal inference within the case-study and comparative traditions, only some of which are even loosely comparable with the logic of regression analysis. As discussed further on, the recent literature supports the conclusion that pathbreaking studies typically rest their inferences not on a single comparative technique but on:

1. various forms of cross-case analysis, using large-N or small-n samples;
2. various forms of 'within-case' analysis; and
3. 'process-tracing' exercises or 'causal process observation'.

Without getting into details for the moment, these techniques are increasingly seen as complementary and highly suited to iterative combinations (Collier et al., 2004c; George and Bennett, 2005: 18, Gerring, 2007: 207-10).[14] However, in different studies they will each be relied upon to a different degree.

The appropriate approach to case selection depends on what the research question is. It is also affected by the kind of causality that seems most likely to be involved – probabilistic or deterministic; simple, conjunctural or equifinal (George and Bennett, 2005: 161-78; Goertz, 2006: Ch 9; Ragin, 1987; 1994; 2000). And finally it depends on how causal leverage is being sought (which inference technique is predominant). There will sometimes be an argument for selecting cases so as to capture the full range of the outcomes that are to be explained. On the other hand, 'no variance' studies are not necessarily inferior bases for causal inference (Collier at al., 2004a; George and Bennett, 2005; Munck, 2004).[15]

We are left with the position that whether a particular selection implies bias or not depends very much on how exactly the research is aiming to contribute to the wider body of knowledge in its field. This is not equivalent to saying that it is an arbitrary matter, or that

13 The extreme case method is distinguished from what he calls 'single-outcome studies', which are not treated as cases of a more general phenomenon and thus implicitly comparative at all. The study of Botswana by Acemoglu et al. (2003) is discussed as an example (Gerring, 2007: 187-210).

14 There are some who disagree: 'Comparison is a competitor to probing study of a case' (Stake, 2006: 83). However, this is a viewpoint from applied social policy, and possibly not typical even in that field.

15 '[I]n the early stages of a research program, selection on the dependent variable can serve the heuristic purpose of identifying the potential causal paths and variables leading to the dependent variable of interest' (George and Bennett, 2005: 23).

bias is never a problem.[16] In the APPP, our approach to case selection should be driven by the nature of our research problem and the types of evidence and analysis that seem most likely to shed light on it. This sounds like common sense. I suggest it also has considerable academic backing.

How many cases is enough?

How many cases is 'enough' and how many would be too many is also one of those questions that the literature suggests depends on the purpose in hand. In the APPP we face an initial trade-off between the coverage of different topics that is desirable to give us enough building blocks for theory building, and the numbers of comparable case studies we can allow ourselves within each topic. Given that we have limited resources, the more topics or clusters of topics we have, the less likely we are to get enough variety in the cases (either within or across countries) for our causal inferences to be made with a high degree of confidence. However, the number of cases is important primarily where cross-case analysis is expected to play a major role, as in – for example – Ragin's prospectus for comparative analysis. Let us assume for the moment that it should.

Where cross-case analysis is centrally involved, the number of cases that is enough depends, first, on the breadth of the proposition or question that is being pursued or the population to which it refers. This is partly a given and partly something that the researcher can decide by the way s/he conceptualises the problem. One is less likely to be accused of having a weak empirical basis for inferences about a question if the scope of the question is narrow, and for this reason it is alleged that some comparativists artificially narrow the scope of their question in order to make their findings appear more robust.[17] However, there is a price to be paid. The proposition may become less interesting and less useful to theory building (and policy). It may also be hard to justify.[18]

Thus, in the APPP we could in principle narrow the scope of our theory-building ambitions in respect of local government to only members of the East African Community and only to the years since, say, 1980. We could consider politics and business only the CFA Franc zone since the devaluation. However, in both cases these restrictions would be hard to justify and would make our work much less interesting in a theory-building perspective.

Even some textbooks addressed to PhD candidates in comparative politics advise against being too cautious and protective. Young researchers with thesis examiners in mind are encouraged to go for a 'risky' level of generality, even though this means that the causal inferences that they are able to draw will be more tentative. This programme has much fewer reasons to be over-cautious about the scope of its propositions.

How many case studies is enough depends – less obviously – on the amount of variability of outcomes that is found in the relevant population and the number of causal factors that need to be considered (Goertz, 2006: Ch 7). This applies particularly if one is interested in the kind of representativeness obtained with Gerring's 'diverse case' technique (capturing the full range of variation in the population). The need for relatively large numbers of cases also arises in Ragin's (1987) discussion of how cross-case comparison can be used most effectively to

[16] According to George and Bennett, the standard, statistically influenced, treatment of selection bias understates the most severe and common kinds of selection bias in qualitative research, where attention focuses only on 'cases whose independent *and* dependent variables vary as the favored hypothesis suggests, ignoring cases that appear to contradict the theory, and overgeneralizing from these cases to wider poulations' (2005: 24-25).

[17] A now classic instance is Geddes' (1990; 2003) critique of Skocpol (1979).

[18] 'The breadth of an inference must make sense; there must be an explicable *reason* for including some cases and excluding others' (Gerring, 2007: 82).

explore complex conjunctural causation. For the reasons explained further on, it is less of an issue if the primary basis for causal inference is within-case analysis or process-tracing.

Gerring (2007) makes an important observation in this connection. The understanding of the relevant population that the researcher has in advance of initial case selection is imperfect, almost by definition, in exploratory research. Neither the extent of variation in outcomes nor the range of relevant causal factors are things that are typically known at the stage when decisions need to be taken. It is necessary to make guesses, and mistakes get made, only to be corrected much later. 'In short, the perfect case study research design is usually apparent only ex post facto' (2007: 149). It may be unrealistic to expect a perfect design – an ideal approach to case selection – at the outset.

I find this piece of advice particularly relevant and helpful for the APPP. It implies that comparative case-study design work does not have to be seen as a once-and-for-all exercise, particularly in a multi-year programme which offers opportunities for commissioning new work at various moments. If it is found, with greater familarity with the topic, that the range of case studies that has been focused on initially is inadequate in kind or number for making sound inferences, there is always the option of undertaking further case studies, selected with a view to correcting the initial biases and improving coverage.

I suggest this is relevant to our upcoming design choices. The most difficult challenge we face continues to be that of finding topics where there is enough relevant variation – outcomes that are significantly better than the norm. We shall not get the necessary causal leverage unless we find more. We should be steered primarily by that consideration, and not by adequacy of coverage. We can legitimately worry about that later, when we have a better idea of the set of variables that we are dealing with. There is wisdom in the remark of George and Bennett:

> 'Case selection is arguably the most difficult step in developing a case study research design. It is an opportunistic process of seeking the intersection between the extant cases that history provides and the kind of cases and comparisons that are likely to best test or develop theories' (2007: 234).

At the risk of repetition, whether the topics speak well to the theory-building exercise is more important than whether we can get a sufficiently rich set of cases for the purposes of causal inference. Any inadequacies on the latter score can be remedied over time though further research. Inadequacies of the other kind will be much more difficult to overcome. This has implications for sequencing and 'who does what' that are picked up towards the end of this paper in Section 7.

4.3 From case studies to comparative analysis

This brings us to the final issue to be considered in Section 4, which goes to the heart of the issue of inter-disciplinary collaboration within the programme. It concerns the question of whether there is a single comparative case-study method, or two with relatively distinct features (as initially suggested in my dialogue with Jean-Pierre Olivier de Sardan), or again a wide spectrum of approaches, capable of infinite combinations and recombinations (as I am now convinced is the case).

The notion that there is just one comparative method, as apparently implied by the title of Charles Ragin's 1987 book has been rightly challenged. The objection would no doubt have been acknowledged by Ragin himself had he not been so fully engaged in defending case-oriented comparative analysis against the 'variable-oriented' approach pursued in fields where large datasets are the norm and probabilistic causal models are considered *de rigeur*. I will get to the reasons why in a moment. But first let me make the case for paying serious attention to Ragin.

Comparative method according to Ragin

I have found Ragin's prospectus for comparative case research quite compelling and highly relevant for the APPP in the medium and long term. It responds well to features that the APPP research problem shares with the types of problems that have been tackled by some of the classics of historical macro-sociology and the political economy of development (say, Skocpol, 1979; Bates, 1981; or Evans, 1995).

The main alleged limitation of the 'Boolean' approach to causal analysis that is advocated by Ragin is that can deal only with the presence or absence of particular causal factors, and the presence or absence of the outcome that is of interest. Critics argue that that the reduction of all variables to a dichotomous on/off or 1/0 basis involves a loss of information, and that most variables are more realistically treated as scales and subjected to probabilistic analysis. However, to reiterate the theme of the recent literature on qualitative research, there are trade-offs in these matters. Which way one jumps has to be influenced primarily by the subject matter and the stage that has been reached in the life cycle of the theory or theoretical problem. Within the programme, there may be some scope for probabilistic regression analysis at some point. However, it is at this stage easier to be convinced that there is a major place for constructing and evaluating hypothese that are deterministic (that is, not probabilistic) and geared to dealing with causal complexity – Ragin's main concern.

In particular, we need to be able to accommodate what Ragin calls conjunctural causation and what he calls multiple causation and others call 'equifinality'. Where causation is conjunctural, we do not expect a single factor on its own, in all circumstances, to provide all or most of the explanation for the outcome we are interested in. We find it more likely that the outcome will be the joint product of certain combinations of factors which interact in positive ways in some circumstances but not all. This is the *forte* of small-n case-study-based comparative analysis as promoted by Ragin:

> 'Unlike multivariate statistical analysis, which tends to be radically analytic (because it breaks cases into parts – variables – that are difficult to reassemble into wholes), qualitative comparison allows examination of constellations, configurations, and conjunctures. It is especially well suited for addressing questions about outcomes resulting from multiple and conjunctural causes – where different conditions combine in different and sometimes contradictory ways to produce the same or similar outcomes. Multivariate statistical techniques start with simplifying assumptions about causes and their interrelation as variables. The method of qualitative comparison, by contrast, starts by assuming maximum causal complexity and then mounts an assault on that complexity'. (Ragin, 1987: x)

We may be wrong initially about the actual conjuction of factors that is needed to bring about the result, but we need a method that is at least sensitive to the possibility and allows us to assess it. To give an excessively simple example but one that is squarely in the sights of the APPP, mobilisation of ethnic identities and traditional forms of authority may be good for outcomes when they are accompanied by certain sorts of competitive politics but not others, and/or when the availability of donor funds is mediated in certain ways and not others.

And other ways of 'working with the grain' that depend less on neotradition and more on local incorporation of new global values may also play a role in causal conjunctures that produce the same or similar outcomes ('equifinality'; George and Bennett, 2005: 161-2; Ragin, 2000: 102-4). As Ragin originally expressed it (1987: 167):

> In case-oriented comparative studies, [the dialogue of ideas and evidence] centers on issues of divergence and causal heterogeneity. The problem is not to specify a single

> causal model that fits the data best (the question that dominates the research dialogue in studies using multivariate statistical techniques) but to determine the number and character of the different causal models that exist among comparable cases. … Thus, the goal in case-oriented comparative studies is not to explain variation but to account for the differences among instances of the same outcome'.

This seems to me to be the sort of perspective that is appropriate to the research problem we have selected. Not everybody in the APPP consortium will be equally convinced, but this seems to me to be one of the topics on which we can agree to disagree. The proof of the pudding may (and may not) be in the eating.

A wider perspective on causal inference from case studies

Another feature of the Boolean approach that should definitely recommend it to us is that, in the final reckoning, it is case-study based. This feature makes it easy to apply in conjunction with other uses of case studies that have proved their worth in comparative politics and anthropology, and which would appear to have promise for the APPP. To repeat what was said above above, recent debate in comparative politics has underlined the fact that case study research is almost never entirely single-stranded. It almost always makes use, in different degrees, of at least three bases for causal inference:

1. various forms of cross-case analysis, large-N or small-n;
2. various forms of 'within-case' analysis; and
3. what is variously called 'process-tracing' or 'causal process observation'.

Authors who have written about case studies and small-n comparative exercises from the vantage point of large-N statistical work are commonly charged with underestimating the degree to which illuminating comparative work draws upon methods 2 and/or 3.

'Within-case analysis refers to various techniques that do not have a particularly prominent place in Ragin's account of comparative method, but which, it is argued, are widely used, to good effect, in actual comparative work in politics and sociology. Gerring (2007: Ch 6) devotes a chapter to assessing the scope for four types of comparative exercise – dynamic, longitudinal, spatial and counterfactual – *within* a case study. He takes the view that these are all approximations to the experimental model. George and Bennett (2005: Ch 9) discuss what they call the 'congruence method' for assessing a theory within a case, placing the emphasis on procedures that provide causal leverage but rely barely or not at all on the model of controlled comparison.

In comparative politics and sociology, the phrase process-tracing was originally associated with the work of Alexander George in the field of international security studies. However, in recent years the term has been quite widely adopted to refer to what researchers do when the hypothesised linkages between associated factors or events are explored by in-depth observation or enquiry (e.g., Gerring, 2007: Ch 7). Others refer to it, perhaps more accurately and less pretentiously as 'causal-process observation' (Collier et al., 2004b). Nonetheless, the textbook by George and Bennett provides the most fully illustrated treatment.

According to George and Bennett, '[t]he process-tracing method attempts to identify the intervening causal process – the causal chain or causal mechanism – between an independent variable (or variables) and the outcome of the dependent variable' (2005: 206). It provides an alternative to controlled comparison where this is not feasible, and an essential tool for assessing the causal validity of the claims arising from cross-case work, whether of the large-N probabilistic or the small-n deterministic (Ragin) sort. It comes in various forms, corresponding to the different types of causality that are of interest, and it has some significant limitations, but it remains 'the only observational means of moving beyond covariation alone as a source of causal inference' (ibid: 224).[19]

The neologisms process-tracing and causal-process observation will not be recognised by most anthropologists or development practitioners. However, the substance to which it refers to should be familiar enough, since it accounts for a large part of what ethnographers and the more serious Participatory Rapid Appraisal (PRA) practitioners spend their time doing. The arguments offered by its advocates should also be recognisable.

In the development field, the standard defence of the anthropological case study, and to a lesser extent the PRA exercise, is that it is better at identifying causal mechanisms and debunking spurious associations than even the best econometric household-survey analysis can hope to be. Partly for this reason, there has been growing recognition of the value of new combinations of methods (Kanbur, 2003; Holland and Campbell, 2005). Multiple methods are particularly required for the part of the causal inference task that involves specifying the mechanisms behind the associations discovered and demonstrating that the association is in fact causal. This corresponds more or less exactly to what is argued in the context of comparative politics and security studies (George and Bennett, 2005: 131-45; Gerring, 2007: 43-48).

Specifying causal mechanism is in many ways the biggest challenge to be faced whatever the technique being employed. We need to keep an earlier point firmly in mind here: that all methods are equivalent in the sense of involving no more than more or less plausible inferences from data. However, different methods do have different strengths and weaknesses. Identifying causal mechanisms is a challenge that often defeats regression analysis because the tools that can be applied to testing for causality, excluding reverse causality and so on are quite limited and put large strains on data availability. Direct observation and various forms of story-gathering can sometimes help. There is a similar problem with Boolean causal analysis if it is relied upon exclusively. A Boolean truth table tells us what factors are associated with which outcomes. It may not tell us much about mechanisms. The detail of case studies on which the comparative analysis is based may be more helpful in this regard.

The recognition now being given to direct observation and in-depth interviewing within the mainstream political science community ought to provide reassurance, if reassurance is needed, that there is a solid meeting ground between the more anthropological and more PolSci members of our consortium

4.4 Some implications

For all of the reasons discussed in Section 4, it seems realistic and justified to set some quite firm limits on what needs to be decided at the conclusion of our design phase in order to set the APPP on a sound path in respect of case-study selection. Tidiness and completeness should concern us relatively little at this point. The primary consideration should be to identify experiences that give us *any* degree of causal leverage on the theoretical problem we want to address. We need to be quite tough-minded about this:

---

[19] Of the several examples of process tracing used for debunking spurious correlations and providing genuine causal explanations that George and Bennett include in their Appendix, my favourite will always be the demolition of Max Weber's 'protestant ethic thesis' by the historian Trevor-Roper. Trevor-Roper used archival research to uncover the reasons for the concentration of capitalists in the Protestant states of early modern Europe, which turned out to have nothing to do with Calvinist belief and everything to do with the intolerance of free enterprise and free thinking in Counter-Reformation states, and the resulting migration of business people out of those countries.

> 'One reason so many case studies of a particular phenomenon in the past did not contribute much to theory development is that they lacked a clearly defined and common focus. … [M]any case studies were not guided by a well-defined theoretical objective. … Collaborative studies must be carefully planned to impress on all participants the requirements of structure and focus'. (George and Bennett, 2005: 70-71)

If we can identify some individual cases which give us leverage on the problem, we should not be too worried for the moment about whether we can find a set of cases to compare it with. We should, certainly, try to be clear about what each case is a 'case' of, so that we are at least able to visualise the complementary studies that might give our interpretations greater 'external validity'. But we need to be equally conscious of the ways those future choices might be affected by what our initial explorations tell us about the relevant variables and their possible causal combinations.

As discussed further below, a high level of standardisation of case studies may also be both difficult and relatively unimportant in our first phase of empirical work. Standardisation becomes feasible when hypotheses are reasonably refined. It becomes essential when cross-case analysis becomes the primary method. But I would distinguish standardisation from the achievement of a minimum degree of intensiveness and empirical rigour. This will be important from the outset, so that the process-tracing and within-case analysis that we are able to do stands up to serious scrutiny. If we are able achieve this, and leave the door open to restudies, cross-case comparability should not become an overbearing problem if and when we tackle a full-scale exercise of the Ragin type.

5 What type of outcomes should interest us and why? (on 'defining the dependent variable')

5.2 Asking the question in the right way

I have put the subtitle of this section in quotation marks because it has connotations that may be unhelpful for us, in two way. First, it is potentially divisive. This language comes from the field of multivariate statistical analysis. It seems natural to those who have been well trained in that tradition, but is is positively off-putting to those who come from other traditions. Even though it is true in principle that any categorisation involves measurement, so that the language of variables is universally applicable, I am not convinced that it is good for the programme to impose a variable-oriented language on all our discussion. 'Outcome' is a perfectly good word that has fewer connotations of academic ethnocentrism, so I shall use it whenever I can.

Secondly, the expressions 'independent' and 'dependent' variable are associated particularly with the phase in the life cycle of theory development in which propositions are operationalised for the purposes of testing and refining them. Of course, operationalisation is important and needs to be well done at the appropriate moment. But it is important not to confuse theory-generation with operationalisation. One effect of doing so can be to skew conceptualisation away from what one is really interested in, from an intellectual and policy point of view, and towards what can be easily defined and measured with available data.

This happens, for example, when the debate about the contribution of foreign aid to development is reduced to one about cross-country regressions of economic growth on the aid/GDP ratio. Fortunately, academic and policy discussions are not yet so dominated by the cross-country regression industry that it is impossible to have a debate about 'development', which is a theoretically interesting concept in part *because* it is contested and not directly measurable. I would find it most unfortunate if the APPP were to get caught up in a reduction of theory to directly testable propositions.

This relates to a more fundamental point about the nature of theory, which was implied but not discussed explicitly in Section 3. In both the natural and the social sciences, there is more to theory and theoretical concepts than the testable predictions than can be derived from them. There is more to the modern theory of gases than the observed relationship between pressure, temperature and volume known as Boyle's Law. There is more to the concept of copper than the colour and other observable properties of the metal (Goertz, 2006: 27). And, closer to home, there is more to the theory of neopatrimonialism than comparative levels of the corruption perception index, and (in the work of Bräutigam, 2000, for example) more to the concept of aid dependence than a statistical association between aid/GDP and some measure of institutional quality.

Theories tell us about *mechanisms*. In the natural sciences, the mechanisms typically involve unobservable entities (atoms, molecules etc.), which is not always the case in the social sciences. But the common feature is that what gets tested is a logical implication of the theory, preferably an implication that has to be empirically true for the theory to be true (Popper's famous injunction), but not the theory itself.

The relevance of all this to the topic of this section is that it calls upon us to respect the multi-level character of theory-development in the social sciences. That means, among other things, paying due attention to how we may want to conceptualise the (differences in) outcomes that interest us from a theoretical point of view, and not turning this prematurely into a discussion about how to measure or otherwise categorise them for the purposes of empirical validation. In other words, our first task should be how to replace the rough-and-ready outcome concept used in the APPP blurbs – 'working/not working for poor people' – with a *theoretically* more robust outcome concept. That is, one that is capable of expressing abstractly the pattern of causal relationships or the mechanism that accounts for empirical findings relating to a range of different specific topics. Operationalisation, in the sense of defining indicators or criteria and locating or constructing relevant data-sets, is an important task that will come to the fore as we start to design specific empirical investigations on particular topics. However, it is not the *same* task.

This obviously relates to our earlier discussion of the stage in the life cycle of theory-development that we are at. If we already had a well-formed theory, there would be more reason to be concerned with operationalising it from the outset. It is the exploratory character of the programme that imposes this insistence on distinguishing the levels of the research enterprise.

5.3 Public goods and collective action problems

What is it about Africa's current development performance that we think is explained in part by the way power is exercised? What do we really have in mind? It is certainly not only one particular final development outcome. It certainly isn't, in 2008, a particularly low economic growth rate. A skewed sectoral pattern of growth is a more likely candidate, but even this varies between countries. A delayed demographic transition, reflecting abysmal child mortality rates is a big problem in Uganda but a much lesser one in Kenya, for example. Many social indicators, including AIDS prevalence and the availability of treatment, vary quite a lot between countries. War is happily no longer ubiquitous, and some countries have remained peaceful and united. What remains of the general picture is just a pervasive sense of disappointment – especially disappointment as compared with comparators in Asia – and the suspicion that there are some common factors at work, linking the unevenness of the final outcomes with the way the sub-continent is typically governed.

Do we have, or could we develop, a concept that captures the common factors that mediate between governance and final development outcomes? I don't think we have to start from scratch in answering this question. The answers are well understood, and they crop up commonly enough in ordinary conversation at every level of society in any country of the region one cares to think of. They also exist, in a more intellectualised form, in major published work that we would be unwise to ignore.

What is common to the ordinary and more intellectualised versions is the prevalence of private solutions to the problems people face, and the extreme weakness of the incentives and enforcement mechanisms associated with those essential goods that can only be provided collectively. The day-to-day examples are superabundant, and add up to a pretty good theory of the proximate causes of underdevelopment: public health measures not enforced; broken standpipes not mended; road construction quality not regulated; an absence of elementary personal security; prohibitive investment transaction costs; uncoordinated markets; public service reforms that fail, etc.

In short, development requires the production of a certain range of public goods, some of them tangible goods or services, others of a more institutional character. But public goods are typically under-provided when actors pursue their solutions independently (in whose interest is it to clear the drains or to deliver a civil-service reform that works?). What is usually missing is some arrangement – e.g. a set of rules and enforcement mechanisms – that enables solutions to be provided collectively or publicly, by the state or other national or local organisations.

The concept of public goods was originally coined within the field of public finance. There, it was largely bound up with the ideological debate between free-market liberals and social democrats about the proper role for state expenditure. But a far more relevant intellectualisation of the set of issues around the under-production of public goods is provided by the twin strands of collective-action political economy and New Institutional Economics.

I am aware that almost everybody associated with the programme has reservations about some of these literatures. There have indeed been some quite effective critiques, both from within each approach (Bates, 1988; 1989; 1995; North, 1990; Nabli and Nugent, 1989) and from outside (Hindess, 1989; Moore, 1995), as well as robust rejoinders (Geddes, 2003: 178-91). Göran Hyden rests his case for the APPP's research prospectus partly on the radical lack of realism involved in some of the simplifying assumptions associated with these approaches (2008b: 24-25).

Arguably, these reservations apply in full only to the wholesale importation of public choice concepts and modes of explanation – or New Public Management reform policies derived from them – into African contexts (e.g. Drazen, 2006, perhaps). They seem less obviously applicable to those sophisticated practitioners whose work has been developed primarily with a view to making sense of change in developing countries (as well as Bates, I would cite Boone, 2003; Geddes, 1994, Gibson, 1999; Gibson et al., 2005; Ostrom, 1990; 2005; and Shivakumar, 2005). However, this is not my main point. My main point is that, whether right or wrong, these interpretations of why particular public goods are typically under-provided (or in a perverse way over-provided; Hyden, 2008b: 31), only exceptionally being produced in a way that meets people's needs, supply a counterpoint (or punch-bag, if you will) that could profit us greatly in elaborating our own explanations.[20]

[20] For a good example, one need go no further than Göran Hyden's Discussion Paper for the programme (2008b: 31-33) which uses Olsnn (1965) as a counterpoint in advancing a proposition of central interest for the APPP, that whenever patronage dominates over policy, the free-rider problem impeding the production of collective goods tends to disappear – 'not because people are altruistic but because it is strategic for a "Big Man" to provide goods to others'.

What I would advocate, then, is not that the programme signs up in any global and uncritical sense to this particular analytical approach, but that we do agree:

1) to take the under-production of public (or club, or otherwise collective) goods as a central organising concept – which I take to be relatively uncontroversial; and
2) to at least take a serious look at the most comprehensive and serious body of work with which we share this interest.

What this is likely to mean in practice is that when we think we have discovered cases in which the under-production of some basic public good has been mitigated to a significant degree, we give some priority to considering the hypothesis that the proximate causes of the difference have to do one of more of the explanatory devices typically invoked by these literatures:

- free-rider problems,
- prisoner's dilemma,
- tragedy of the commons,
- Samaritan's dilemma,
- assymetric power,
- missing or assymetric information, and
- transaction costs, etc.

I take it as read that we would apply any such concepts in a culturally attuned and socially sensitive way, as I believe many of their advocates do. It would not be a question of taking on board any of the more academically imperialist and universalising variants of the rational-choice approach. I think that one of the more exciting things we shall do within the programme is probe the limits of this type of analysis and establish where it is strong and where it is weak. In any case, I challenge anyone to nominate a better body of work on which to hang our outcome analysis.

One of the attractions of the public goods concept for the APPP is that it is capable of playing a summative role across the topics (sectors) that we take as our empirical foci. Obviously, the relevant public goods will be different in different sectors, so there will not be any question of simply aggregating the results. However, the notion that theories might be produced just by adding up the results of empirical reseach would be very odd in the light of our discussion above.

Naturally, our middle-range analysis (case studies) will call for some agreement on the nature of the relevant outcome issues *in the sector* – what are the public goods that are under-produced to different degrees as a result of different institutional patterns? At a certain point, the question will arise whether the relevant production is best treated as a merely categorical difference (under-produced vs. not under-produced) or is capable of supporting an ordinal or interval scale. We may also want to consider in due course whether any proximate causes – e.g. the overcoming of a normally fatal principal-agent assymetric information problem – should be considered part of the outcome, rather than simply part of the explanation. This may be a minor matter but could be important to the programme's ability to formulate novel theoretical propositions in economical and appealing ways.

I hope I have shown reasons for approaching the question of the dependent variable in a way that takes us beyond the discussions we have had in the programme up to now. The key idea I have advanced – the importance of not reducing theory development to operationalisation – is also relevant to our next topic, when and how we refine our conceptual vocabulary for handling causal processes and developing explanations.

6 What kind of explanatory concepts do we need? (on definitions and their place in theory formation)

A frequent refrain in some of our discussions has been 'we must define our terms – otherwise we shall be talking at cross-purposes!' It has been suggested several times that we need a programme glossary. Candidates for agreed definition have included 'formal', 'informal', 'hybrid', 'clientelism', 'patrimonialism', 'governance', 'public goods', 'social embeddedness' and 'working with the grain'. This may seem like self-evident good sense, but I want to dispute it or at least subject it to heavy qualification.

It is hard not to view conceptual consistency as a virtue, and it is relatively easy to cut and paste some standard definitions for reference, as we have seen already for the case of 'institutions' and 'organisations' (Box 1 above). It is, however, important to know when to stop settling by definition issues that are actually empirical or theoretical, and which are therefore important signals of matters to be investigated.

Box 3: Examples of explanatory concepts

Neopatrimonialism:

A hybrid regime in which patrimonial practices coexist with modern bureaucracy – conveys the idea that, as in pre-modern states, the wealth and powers of the state are indistinguishable from those of the ruler (see e.g. van de Walle, 2002: 50-55)

Aid dependence:

Reliance of public revenues on aid rather than taxes – conveys the idea that this is a problem insofar as it discourages the development of mechanisms of domestic accountability, etc. (see e.g. Bräutigam, 2000).

Public (or collective) goods:

Any jointly produced good (material or otherwise) where if any member of the group consumes it, it cannot practically be withheld from the others in the group – now closely associated with the proposition that the infeasibility of exclusion eliminates any direct incentive for individuals to contribute to the cost of the good (see e.g. Olson, 1965: 14-16, Hardin, 1982: 19-20).

The observation was made earlier that theoretical concepts have *depth*. To a greater or lesser extent, they incorporate propositions about the world, usually about causal mechanisms – e.g. the particular atomic structure that gives copper its colour, melting point, etc. Concepts are not to be confused with the indicators associated with them. This makes definition more than a mere classification exercise. Established social-science concepts, including some of those 'defined' in Box 3, are established in the sense that they signal in a recognised way a particular body of analytical thinking, including some substantive claims about causal linkages (Goertz, 2006). [21]

For this reason, one can accept a definition but also add the rider that one does not *like* the concept it describes, meaning one disagrees with what it says about the way the world works. Several concepts associated with Max Weber (e.g rational-legal authority) are cases in point. 'Civil society' is surely another – as a theoretical concept it is badly compromised by the combination of its original usage in early modern Europe and its appropriation in recent years by the aid and democratisation businesses. These associations bring with them a series of

[21] '[A] concept involves a theoretical and empirical analysis of the object or phenomenon referred to by the word. …A purely semantic analysis of concepts, words, and their definitions is never adequate by itself. … To develop a concept is more than providing a definition: it is deciding what is important about an entity' (ibid.: 4-5, 27). Or, with Gerring (2001: 59-60): 'One can think of concepts whose existence is almost wholly dependent on their utility within a theoretical framework. … The broader point is that concepts rest within propositions, and propositions rest within research designs'.

implicit, and largely contradictory, claims about social and political change now and in the past.

That applies to established concepts. Several of those in the definitional wish-list above are not in that category, but rather words in common usage in search of a more satisfactory, theoretically useful, formulation. That certainly includes formal, informal, hybrid, and metaphors like ‘working with the grain’. It would be possible in all of these cases to take the dictionary definition, which reflects common usage, and harden it up, so that any particular observed phenomenon falls clearly on one side or the other of a categorical divide. That might stop us arguing about the meaning of words, but the danger is that it would stop us having arguments that are actually about the structure of reality or about how to explain it better.

For example, in our discussions Giorgio Blundo – who heads the APPP team in Marseille – has more than once made the statement that it is wrong to distinguish formal from informal behavioural rules within public organisations on the basis of whether or not they are written down, in the law, or the organisation’s statutes or rule book. In organisations with which he is familiar, the unwritten rules governing much of the behaviour of members of the organisation are so regularised and ingrained that it would be quite inaccurate to call them informal. My English dictionary does not settle the matter. In fact, it tends to support Giorgio’s usage by associating formality with regularity of some kind, and not particularly with codification in writing. Others in the team find this absurd. But at least partly, this seems to me to involve, first, an empirical question, and then a substantive theoretical one.

The empirical question is whether the way institutional hybridity has been handled in the main literature in English on neopatrimonialism – where actual practice *is* generally contrasted with written rules, by the actors themselves as well as by observers – is overgeneralising from a limited and maybe distinctively Anglophone type of organisation. How common is it, one would want to know, that the rules governing corruption, are so regularised and binding that they might as well be written down?

The theoretical question is whether this language is adequate, at all, to the task of explaining the particular outcome variations that interest us? What do our difficulties in using common speech to capture the relevant differences tell us about the conceptual gaps that need to be filled?

Hybridity provides another example. Thomas Bierschenk – APPP Advisory Group member – in a comment on the APPP Inception Report, says that he disapproves of the usage ‘hybrid’ to distinguish African institutional patterns, because all institutions everywhere are hybrids. My first response to this was to think: that is of course true but it is also a matter of degree, and the *degree* of hybridity in the institutions of postcolonial Africa is beyond dispute.

It would not be hard to provide a definition that would prove me right and him wrong. But Bierschenk is a serious thinker with a lot of field experience, and my suspicion therefore is that there is more to this than meets the eye. What exactly he means by this would be worthy of further discussion, again with both empirical regularities and theoretical accuracy or parsimony in mind. The concerns expressed by another Advisory Group member, Ole Therkildsen, about the public management literature’s over-reliance on the concept of neopatrimonialism probably need to be the subject of a similar discussion (Therkildsen, 2005).

‘Embeddedness’ provides a third example. Peter Evans’ (1995) account of the factors that made certain states in Asia developmental and not predatory famously revolves around the claim that they, or their bureaucracies, were both autonomous and socially embedded. It is famously true, too, that while this expression appears a reasonably good way of conveying certain characteristics acquired by the state bureaucracy in Japan and South Korea at a certain point, applying it prospectively to other state-society relationships is challenging. 'What do we mean by embeddedness?' involves the question of how to distinguish between forms of interaction between state officials and economic interest-groups that remove institutional barriers to progress and those that keep them in place. It is a question about how best to make sense of evidence on differential performance.

I would argue, therefore, that defining the concepts that will need to figure as explanatory constructs in eventual theoretical propositions should be approached as a *process*, a task to be undertaken over time. This does not do away with the need to define our terms as carefully as we can in each piece of work we undertake. In fact, unless we do that, it will not be apparent which differences of definition contain the germ of a theoretical discovery or refinement, and which are simply careless or accidental. Neverthless, conceptual elaboration should be undertaken in close relationship with the analysis of empirical material, and not subjected to *a priori* rulings. If a particular usage appears to violate common sense or established usage in a discipline, we should take care to ask why. It will often be useful to make reference to precedents in the literature. It may even be helpful to make some use of an English or French dictionary. But it will normally be a mistake to try to concoct a theoretical definition in the abstract.

7 What empirical approaches are required? (on fieldwork coordination, timing and sequencing)

Several points from the previous sections are relevant to setting ourselves appropriate objectives for the design of the programme's empirical reseach. They include the exploratory character of the programme's immediate research tasks, and the priority to be given in the first phase to identifying topics for study that promise to illuminate the programme's research problem and contribute to theory development around it.

Also relevant is the argument about the possibility and desirability of drawing on the full range of forms of causal inference from case studies, not just cross-case comparative analysis, even if the latter remains an essential medium-term objective. This has the implication that we do not require a well-selected *set* of cases on each topic in order to make the kind of headway we most need to begin with. At least in the first phase of work, it is more important to have case studies that speak well to our theory-building objective than to have many potentially comparable case studies. Standardisation of approach may not be feasible until we are in a position to construct some more elaborated causal stories. On the other hand, what we do in the first phase should not compromise our ability to make rigorous comparisons when we do have more cases and more refined hypotheses. So where we are studying the same topic in several places we should pay attention to empirical rigour and seize any opportunities for convergence in fieldwork approaches from the outset.

These arguments articulate retrospectively part of the rationale for the iterative, top-down/bottom-up, approach to design issues which we have been taking in the APPP up to now. I think they also have implications for what should be considered a sufficiently satisfying outcome from the programme's design phase. As indicated at the beginning of the paper, it seems important that we do not reproach ourselves unduly if the initial empirical design we come up with is lacking in elegance and tidiness. There are more important matters at stake.

If all of the above is true, what we definitely need is agreement on:

- a set of middle-range topics (e.g. some of those suggested in Figure 1) that seem *both* researchable *and* likely to provide some kind of causal levage on the programme's research problem;

- a shared concept of *how* – in an exploratory way – the proposed studies of these topics will illuminate relationships between institutional variations and relevant outcomes (public goods);
- what the relevant public-goods outcome variations are for each topic, and the ways these can be recognised, documented and compared;
- a continuing dialogue on how best to combine emic and etic perspectives and different observational, survey or action-research techniques in each of the topic areas; and
- a *mechanism* for coordinating the fieldwork design between exponents of each topic, such that the cross-case comparison that we aspire to is not hindered by the use of different field methods or incompatible interview guides.

We are not planning to settle details of our empirical approaches at the Cape Town workshop. As well as reviewing some of the general methodological issues discussed in this paper, we shall be focusing on a set of outline proposals to assess whether they meet the basic criteria above (particularly, leverage on the APPP research problem, researchability and arrangements for coordination). Those outlines that pass these basic tests will be elaborated with particular attention to empirical design and resourcing during the next three months. The conversations that have been taking place between those proposing empirical work on the same topics will need to change into a higher gear at this stage, so that coordinated fieldwork becomes the main theme of discussion.

At the programme launch meetings, we agreed that one of the criteria to be applied in allocating funding to individual projects would be the extent to which they exploit the potential for fruitful collaboration between APPP partners. We expected that, as a minimum, this would involve close coordination of fieldwork to be undertaken in different countries on the same topic, with a view to standardisation (or at least convergence) of field methods. I hope that it will also involve, in cases where it is practical on language and other grounds, the construction of mixed teams that work jointly across more than one country. This will happen if we create a culture of collaboration in the programme, and not if we don't.

The signs are encouraging in this respect. However, on this too we need to stay realistic about our ambitions for Cape Town. We should not expect the details of collaborations to emerge fully armed during the workshop – not least because many of us will be meeting face-to-face for the first time. On the other hand, a reasonable ambition is to agree mechanisms – by which I mean identified people and a timetable and method of communication – which give us confidence that these matters will be hammered out during the following few months.

8 Conclusion

I said that this paper would be both reassuring and reaffirmative. Having reviewed the main research design issues that we face against the background of some of the best of the relevant literature, I will now try to say what I think has emerged that is either importantly reassuring or usefully reaffirmative.

It is importantly reassuring, I suggest, that the comparative case-study approach does not imply radically different and potentially incompatible things across our disciplines. There is no sharp divide on this issue between comparative politics and macro-sociology on the one hand and social/cultural anthropology on the other. None of these disciplines is as monolithic as it may appear from the outside. Recent debate among leading US-based political scientists has not just established that case-oriented comparative methods have strengths that variable-oriented approaches do not. It has also revealed that case-study research often makes effective use of techniques of within-case analysis and causal-process observation that will be familiar to anthropologists in relying relatively little on cross-case analysis. At the limit, the forms of causal inference employed by some macro-comparativists resemble quite closely those involved in ethnographies of the sort undertaken by anthropologists in the broad 'Manchester' tradition.

This is reassuring because it means there is no reason for the barriers to integration of analytical approaches across the consortium to be insuperable. That does not mean that we shall never disagree. But it does mean that any disagreements we shall have will not be discipline-based or bipolar, and that therefore they can assume the constructive form of debates about how best to explore the substantive questions before us. I find this an encouraging conclusion to the dialogue I began with Jean-Pierre Olivier de Sardan of LASDEL some months ago, the starting point of which was a sense that Ragin's account of *the* comparative method might be hard to reconcile with the case-study approach that has been pursued with some success by his team.

Another group within the programme who, I hope, will feel reassured are those who have felt uncomfortable about the relatively loose process followed during our design phase, with think pieces, brainstormings, scoping missions and bilateral conversations feeding into a large melting pot without any decisive resolution. I believe I have shown that those who suppose that at this point we should have well-articulated hypotheses, clearly defined terms and an overarching plan for testing them are starting from a vision of the typical reseach process that is seriously oversimplified at best. A sound appreciation of the nature of theory and the life cycle of theoretical development in the social sciences points in the other direction. In the exploratory phase, research needs to have a strong inductive dimension. Premature refinement of concepts and operationalisation of propositions is likely to lead a programme to take its eye off the ball, and end up investigating something other than the research problem it started with. Looseness can be a virtue in this context, particularly where – as in our case – the discovery of cases that give us real causal leverage on the research problem is itself rather challenging.

As many of the recent textbooks say, the teaching of research methods at undergraduate level in some subjects has a lot to answer for. Together with the misleading image of the process of discovery in the natural sciences that prevails in the popular mind, it has created a context in which programmes dedicated to developing new theory on the basis of inductive research have a hard time making themselves understood, even within their own ranks. I do not suppose that this paper will be sufficient to settle the matter, but I would encourage those who remain unconvinced to consult some of the sources I have used and reconsider in that light.

What exactly has been reaffirmed? I began by devoting a few pages to what the programme's research problem is. This may not seem to have been entirely necessary, since the key features were present from the very first of our internal discussion papers and have been reproduced in our documents several times since. But given that we are committed to a unitary research design, it seems best to err on the side of caution and insist a bit on what the basic issue is. The basic issue is whether Africa's complex institutional heritage might be harnessed more effectively to the achievement of development and poverty reduction, and if so how. This means focusing empirically on institutional differences that seem to be associated with important differences in development outcomes.

Another theme that has been reaffirmed is the harmony between our theory-building and our policy-transforming objectives. Theory and policy are not antithetical in principle. What can make theory not very useful for policy purposes is that it is too general – not focused enough on systematic differences, especially of the kind where there is some scope for policy to exercise an influence. That has been one of our selling points, and it remains true. One of the additional things that my reading has brought into stronger view is that this stricture applies to over-general quantitative results as well as to the more qualitative overviews.

Lastly, I hope I have helped prepare the ground for the programme design workshop in Cape Town by moderating expectations in areas where they may have been inappropriately high. I have leant heavily on the best of the recent methods literature to assist me in this effort. The literature has helped to put concerns about representativeness and selection bias in their proper place. It has led me to recognise that best practice in building sets of cases for cross-case comparison is quite process-oriented. Case study selection for comparative purposes often *cannot* precede the start of fieldwork, does not *need* to, and probably *ought not* to. The possibility and even likelihood that the first studies to be completed will suggest new causal factors or different ways of conceiving the relations between the variables, and thus the need for a different set of comparative cases, needs to be taken seriously.

It follows that, rather than trying to set up a perfect cross-case design, our focus at this stage should be on getting an initial cohort of individual case studies that are of the right quality. Without discounting empirical rigour, a key criterion of quality should be the ability to speak directly or indirectly to our central research problem.

References

Abercrombie, Nicholas; Hill, Stephen; and Turner, Bryan S. (1984) *The Penguin Dictionary of Sociology*. Harmondsworth, Middlesex: Penguin Books.

Acemoglu, Daron; Johnson, Simon; and Robinson, James A. (2003) 'An African Success Story: Botswana', in Rodrik (2003).

Bates, Robert H. (1995) 'Social Dilemmas and Rational Individualism: An Assessment of the New Institutionalism', in John Harriss, Janet Hunter and Colin M. Lewis (eds) *The New Institutional Economics and Third World Development*. London: Routledge.

Bates, Robert H. (1989) *Beyond the Miracle of the Market: The Political Economy of Agrarian Development in Kenya*. Cambridge: Cambridge University Press.

Bates, Robert H. (ed.) (1988) *Toward a Political Economy of Development: A Rational Choice Perspective*. Berkeley: University of California Press.

Bates, Robert H. (1981) *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies*. Berkeley, CA: University of California Press.

Bierschenk, Thomas; Blundo, G.; Jaffré, Y., and Tidjani Alou, M. (eds) (2007) *Une anthropologie entre rigeur et engagement: Essais autour de l'oeuvre de Jean-Pierre Olivier de Sardan*. Paris: Karthala/IPAD.

Blundo, Giorgio (2007) 'Un "afro-réalisme" réformiste: L'anthropologie politique de Jean-Pierre Olivier de Sardan', in Bierschenk et al. (2007).

Blundo, Giorgio (2006) 'Corruption in Africa and the Social Sciences: A Review of the Literature', in Blundo and Olivier de Sardan (2006).

Blundo, Giorgio, and Olivier de Sardan, Jean-Pierre (eds) (2006*) Everyday Corruption and the State: Citizens and Public Officials in Africa*. London: Zed Books.

Boone, Catherine (2003) *Political Topographies of the African State: Territorial Authority and Institutional Choice*. New York: Cambridge University Press.

Booth, David (2003) 'Patterns of Difference and Practical Theory: Researching the New Poverty Strategy Processes in Africa', *Journal of International Development* 15: 863-77.

Booth, David (ed.) (1994) *Rethinking Social Development: Theory, Research and Practice*. London: Longman.

Booth, David (1985) 'Marxism and Development Sociology: Interpreting the Impasse', *World Development* 13(7): 761-87.

Brady, Henry E., and Collier, David (eds) (2004) *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* Lanham, MD: Rowman & Littlefield.

Braütigam, Deborah (2000) *Aid Dependence and Governance.* EGDI Study 2000:1. Stockholm: Almqvist and Wiksell International.

Chabal, Patrick, and Daloz, Jean-Pascal (2006) *Culture Troubles: Politics and the Interpretation of Meaning.* London: C. Hurst and Co.

Chabal, Patrick, and Daloz, Jean-Pascal (1999) *Africa Works: Disorder as Political Instrument*. Oxford: James Currey, and Bloomington, IN: Indiana University Press.

Collier, David; Mahoney, James; and Seawright, Jason (2004a) 'Claiming Too Much: Warnings about Selection Bias', in Brady and Collier (2004).

Collier, David; Brady, Henry E.; and Seawright, Jason (2004b) 'Sources of Leverage in Causal Inference: Toward and Alternative View of Methodology', in Brady and Collier (2004).

Collier, David; Brady, Henry E.; and Seawright, Jason (2004c) 'Critiques, Responses, and Trade-Offs: Drawing Together the Debate', in Brady and Collier (2004).

DAC (2006) 'The Challenge of Capacity Development: Working Towards Good Practice?. Paris: OECD, Development Assistance Committee, Network on Governance.

Drazen, Allan (2006) 'Is There a Different Political Economy for Developing Countries? Issues, Perspectives, and Methodology'. Paper Presented to Plenary of the African Economic Research Consortium, Nairobi, May 28.

EuropeAid (2005) *Institutional Assessment and Capacity Development: Why, What and How?* Brussels: European Commission.

Evans, Peter (1995) *Embedded Autonomy: States and Industrial Transformation*. Princeton, NJ: Princeton University Press.
Geddes, Barbara (2003) *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.
Geddes, Barbara (1994) *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley: University of Califormia Press.
Geddes, Barbara (1990) 'How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics', *Political Analysis*, 2: 131-50.
Geertz, Clifford (1973) *The Interpretation of Cultures: Selected Essays*. New York: Basic Books.
George, Alexander L., and Bennett, Andrew (2004) *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
Gerring, John (2007) *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.
Gerring, John (2001) *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.
Gibson, Clark C. (1999) *Politicians and Poachers: The Political Economy of Wildlife Policy in Africa*. New York: Cambridge University Press.
Gibson, Clark C.; Andersson, Krister; Ostrom, Elinor, and Shivakumar, Sujai (2005) *The Samaritan's Dilemma: The Political Economy of Development Aid*. Oxford: Oxford University Press.
Giddens, Anthony (1984) *The Constitution of Society*. Cambridge: Polity Press.
Goertz, Gary (2006) *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.
Golooba-Mutebi, Frederick (2008a) *Collapse, War and Reconstruction in Rwanda: An Analytical Narrative on State Making.* Working Paper 28.2. London: London School of Economics, Crisis States Research Centre.
Golooba-Mutebi, Frederick (2008b) *Collapse, War and Reconstruction in Uganda: An Analytical Narrative on State Making*. Working Paper 27.2. London: London School of Economics, Crisis States Research Centre. (www.crisisstates.com)
Granovetter, Mark (1985) 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology* 91(3): 481-510.
Grindle, Merilee S. (1997) 'Divergent Cultures? When Public Organizations Perform Well in Developing Countries', *World Development* 25(4): 481-95.
Hardin, Russell (1982) *Collective Action*. Baltimore, MD: Johns Hopkins University press.
Hindess, Barry (1989) *Political Choice and Social Structure: An Analysis of Actors, Interests and Rationality*. Aldershot, Hants: Edward Elgar.
Holland, Jeremy, and Campbell, John (eds) *Combining Quantitative and Qualitative Methods in Poverty Analysis*. London: Intermediate Technology Publications.
Hyden, Göran (2008a) 'After the Paris Declaration: Taking on the Issue of Power', *Development Policy Review* 26(3), May.
Hyden, Göran (2008b) *Institutions, Power and Policy Outcomes in Africa*. Discussion Paper 2. Africa Power and Politics Programme. June.
Kanbur, Ravi (ed.) *Q-Squared: Qualitative and Quantitative Methods of Poverty Appaisal*. Delhi: Permanent Black.
Kelsall, Tim (2008) *Going with the Grain in African Development?*. Discussion Paper 1. Africa Power and Politics Programme. June.
King, Gary; Keohane, Robert, and Verba, Sidney (1994) *Designing Social Inquiry*, Princeton, NJ: Princeton University Press.
Leonard D.K. (1991) *African Successes: Four Managers of Rural Development in Kenya.* Berkeley CA: University of California Press.
Leonard, D.K. (1987) 'The Political Realities of African Management', *World Development* 15(7): 899-910.
Long, Norman, and Long, Ann (1992) *Battlefields of Knowledge: The Interlocking of Theory and Practice in Social Research and Development*. London: Routledge.

MacIntyre, Alasdair (1971) ‘Is a Science of Comparative Politics Possible?’, in MacIntyre, *Against the Self-Images of the Age: Essays on Ideology and Philosophy*. London: Duckworth.
MacIntyre, Alasdair (1970) ‘The Idea of a Social Science’, in Bryan R. Wilson (ed.) *Rationality*. Oxford: Basil Blackwell.
Manor, James (2008) ‘An Asianist’s Perspective on the Africa Power and Politics Programme’. Draft Discussion Paper. Africa Power and Politics Programme. February
Moore, Mick (1994) ‘How Difficult is it to Construct Market Relations? A Commentary on Platteau’, *Journal of Development Studies* 30(4), July.
Mosse, David (2005) *Cultivating Development; An Ethnography of Aid Policy and Practice*. New Delhi: Vistaar Publications.
Munck, Gerardo (2004) ‘Tools for Qualitative Research’, in Brady and Collier (2004).
Nabli, M.K., and Nugent, J.B. (1989) ‘The New Institutional Economics and Its Applicability to Development’, *World Development* 17(9), Sept.
North, Douglass (1990) *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
Olivier de Sardan, Jean-Pierre (2005) *Anthropology and Development: Understanding Contemporary Social Change*. London: Zed Press.
Olivier de Sardan, Jean-Pierre (1998) ‘Emique’, *L’Homme*, 147: 151-66.
Olson, Mancur (1965) *The Logic of Collective Action*. Cambridge MA: Harvard University Press.
Ostrom, Elinor (2005) *Understanding Institutional Diversity*. Princeton: Princeton University Press.
Ostrom, Elinor (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
Parsons, Talcott (1951) *The Social System*. London: Routledge & Kegan Paul.
Pelto, Pertti J., and Pelto, Gretel (1978) *Anthropological Research: The Structure of Enquiry*. 2nd edn. Cambridge: Cambridge University Press.
Ragin, Charles C. (2000) *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
Ragin, Charles C. (1994) *Constructing Social Research*. Thousand Oaks, CA: Pine Forge Press.
Ragin, Charles C. (1987) *The Comparative Method: Beyond Qualitative and Quantitative Strategies.* Berkeley: University of California Press.
Robbins, Stephen P. (2004) *Organizational Behaviour*. Upper Saddle River, NJ: Prentice-Hall.
Robinson, Mark (2006) *The Political Economy of Governance Reforms in Uganda*. Discussion Paper 386. Brighton: Institute of Development Studies.
Rodrik, Dani (ed.) (2003*) In Search of Prosperity: Analytical Narratives of Economic Growth*, Princeton: Princeton University Press.
Rodrik, Dani; Subramanian, Arvind; and Trebbi, Francesco (2004) ‘Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development’, *Journal of Economic Growth* 9: 131-165.
Shivakumar, Sujai (2005) *The Constitution of Development: Crafting Capabilities for Self-Governance*, New York: Palgrave Macmillan.
Skocpol, Theda (1979) *States and Social Revolutions: A Comparative Analysis of France, Russia and China*. New York: Cambridge University Press.
Stake, Robert E. (2006) *Multiple Case Study Analysis*. New York: Guildford Press.
Teskey, Graham (2005) ‘Capacity Development and State Building: Issues, Evidence and Implications for DFID’. London: Department for International Development.
Therkildsen, Ole (2005) ‘Understanding Public Management Through Neopatrimonialism: A Paradigm for all African Seasons?’, in Ulf Engel and Gorm Rye Olsen (eds) *The African Exception*. Aldershot, Hants: Ashgate.
van de Walle, Nicolas (2001) *African Economies and the Politics of Permanent Crisis, 1979-1999*. Cambridge: Cambridge University Press.

World Bank (1991) 'Managing Development: The Governance Dimension'. Washington, DC: World Bank.