The Uneven Development of Quantitative Social Science in South Africa

One of the most striking contrasts between journals in South African studies (or African studies more broadly) and journals in the social sciences in the USA or Europe is the balance between quantitative and qualitative research. In the USA, especially, a high proportion of contemporary research in political science and sociology as well as economics and development studies entails the analysis of quantitative data using ever more complex statistical techniques. In South Africa (and Africa) this kind of social science remains significantly underdeveloped.

Since the mid-1990s, however, there has been a marked resurgence of quantitative research in South African social sciences, fuelled by the availability of new data, ready access to personal computers and more user-friendly statistical software, policy-makers’ urgent requests for quantitative analysis and scholars’ re-immersion in global academic networks. Economists have taken the lead in this new scholarship, but some political scientists, sociologists, criminologists and others have moved in a similar direction. Much of the new work has, however, been slow to cross disciplinary boundaries.

The first part of this article reflects on the parlous condition of quantitative social science under apartheid, so as to highlight more clearly the significance of the post-apartheid explosion in quantitative research. I then survey the development of quantitative social sciences over the past five to ten years. This development has been uneven, with rapid change in some fields (notably the study of poverty, inequality and public opinion) not matched in others (such as education). Finally, the article identifies some of the weaknesses of quantitative social science in South Africa, and suggests ways in which these weaknesses can be addressed.

Quantitative social sciences in pre-democratic South Africa
The development of the social sciences in South Africa under apartheid was profoundly influenced, directly and indirectly, by the political context. This is especially true with respect to the development of quantitative social science. Of the three major reasons why quantitative research developed slowly in apartheid South Africa, two are political: the state declined to produce the necessary data on the majority of South Africans, and political differences (including differences among opponents of apartheid) penetrated into academic life, such that the limited quantitative work that was done was often ignored by other social scientists. The third reason for the slow development of quantitative social science was the hostility to quantitative, empirical research among many social scientists, who preferred either theory-driven or qualitative research strategies. This can hardly be attributed to the political context.

In most countries quantitative social science depends very heavily on the generation of data by state agencies. In South Africa, the apartheid state invested considerable resources in collecting data on its white citizens, and to a lesser extent its coloured and Indian citizens, but it failed – or refused – to generate meaningful data on problems such as poverty and unemployment as they affected its African subjects.

On incomes, the state collected data on white but not black South Africans through its intermittent surveys of family expenditure. The official Population Census, conducted every ten years, collected no usable data on African incomes despite providing sophisticated statistics on other economic indicators. Researchers interested in issues of poverty or inequality in the population as a whole therefore had to rely on other sources outside of the state. From the early 1960s the Bureau of Market Research (at the University of South Africa) began to collect data on incomes and expenditures among coloured, Indian and African people, but the samples were small and restricted to certain urban areas. Otherwise, researchers had to rely on localised academic surveys, such as the survey of the Keiskammahoek district of the Ciskei supervised by the leading liberal economist Hobart Houghton in 1949 (see Houghton and Walton, 1952) and similar studies on farms (Roberts, 1958) and urban areas (including Wix, 1951; Brooks and Hurwitz, 1957; Houghton, 1960; Reader, 1961). In the 1960s the stream of such independent studies dried up, perhaps because of the exodus of so many younger social scientists abroad in the face of oppression. The state did collect detailed data on wages, broken down by race and economic sector, so that scholars could analyse racial discrimination and inequality in wages, but they could say little about income distribution as a whole. Despite the ‘need for reliable statistical information about income distribution’, Houghton wrote in his contribution to the 1971 *Oxford History of South
Africa, ‘accurate up-to-date data about earnings of the various racial
groups are difficult to obtain’ (1971: 43).

Similarly, the state collected no meaningful data on
unemployment among African people. The official register of
unemployment included only people who were eligible for
unemployment insurance benefits, which meant that most unemployed
African men and women were excluded from the register. The
Population Census collected data on the employment status of everyone
living within South Africa’s borders, but the instructions to enumerators
resulted in data that was of little or no use to researchers. In 1970, for
example, men and women who were not working but could give a prior
occupation or even the industry in which they had last worked were
classified as employed (Loots, 1978: 10). Definitions of who was
considered to be participating in the labour force also varied from census
to census. As John Knight pointed out, in the 1960 South African
census:

‘a male aged 16 or over in the homelands was classified as
“peasant farmer” unless another occupation was specified; a
female in the same position was classified as ‘housewife’.
In the 1970 census, the homeland wives of household heads
were classified as ‘housewives’ and thus ‘not economically
active’, but other females classified 16 or over were
classified as ‘peasant farmers’. Hence half the recorded
increase in economically active female Africans between
1960 and 1970 is the result of a change in the
definition…changes in definition from one census to
another render inter-census comparisons of the
The 1960 and 1970 censuses suggest that the unemployment rate fell by
two-thirds across the decade. This is nonsense, the result being primarily
because of the change in definition.

Scholars interested in income distribution or unemployment for
the population as a whole therefore had to resort to creative techniques.
Economists such as Charles Simkins, John Knight, Norman Bromberger,
Jill Nattrass and Michael McGrath pioneered new standards in empirical
research on poverty, inequality, migrancy and unemployment, but their
work remained constrained by the grave limitations of the data. Studies
of income distribution, for example, had to estimate aggregate African
incomes as a residual, subtracting the incomes of white, coloured and
Indian people (on which there was some data) from total personal
income. Even if this was done on the basis of sensible if crude
assumptions, estimates of the African income share were very sensitive
to errors in the estimation of other racial groups’ shares (McGrath, 1983:
And, in order to estimate the unemployment rate within the African population, researchers had to calculate the gap between the estimated African labour force and the number of jobs held by African men and women, making some allowance for self-employment in subsistence agriculture and other sectors (Knight, 1977; Simkins, 1978a, 1978b; see also Seekings and Nattrass, forthcoming).

In the late 1970s and early 1980s there was a marked resurgence of localised surveys examining poverty, incomes and employment among black South Africans. This provided impetus to, and was in turn pushed forward by, the Second Carnegie Inquiry into Poverty and Development in Southern Africa, and especially its conference on poverty held in Cape Town in 1984. Conference papers provided rich quantitative data on poverty in many parts of the country (see Wilson and Ramphele, 1989). The Centres for Applied Social Science (CASS) and of Development Studies (CDS), both at the University of Natal in Durban, were home to many studies in what is now the province of KwaZulu-Natal. Elsewhere, too, individual sociologists collected local data on topics related to poverty, such as migration histories. But the abundance of data from local studies was not collated to reveal the ‘big picture’, presumably in part because of the difficulties of such a project given the computing technology available at the time (although see May [1987] on rural areas). Nor was there any new national-level data.

The choice of topics by both economists and sociologists clearly reflected the political context. Poverty was political in that it was seen to be due in large part to state policies of dispossession, forced removal, influx control, restrictions on urban development, and so on. Many studies of migration histories were prompted by opposition to influx control. But the cost of generating credible and thorough countrywide data was prohibitive to researchers who were denied funds by the state, and the state was obviously not keen to fund its critics. The Second Carnegie Inquiry, indeed, was subjected to virulent verbal attack in Parliament by the State President (Wilson, 1997: 227).

The quantitative study of political behaviour (including attitudes) was also constrained by the availability of data. In the late 1960s academics and newspapers first began to conduct opinion polls among voters, but these polls were limited to white citizens as only they could vote. These polls allowed scholars to take a huge step forward, since until then studies of elections had been squarely within the British ‘Nuffield’ tradition of primarily descriptive or historical accounts, largely devoid of quantitative analysis of voting behaviour. But the innovative study of white South Africans’ attitudes (Lever, 1972, 1974; Schlemmer, 1973; Morse and Peele, 1975) was not matched by comparable research among black South Africans. This was in part due
to the practical difficulties and costs of research among black South Africans. It was easier and cheaper to conduct quantitative research among samples of white people. In a number of fields, therefore, research on the attitudes and experiences of white people was more advanced than that on those of black people. Thus Morse and Nel (1977) examined identities among a sample of Afrikaans university students, Booyzen and Kotzé (1985) studied political socialisation among a sample of white students, and Jubber (1988, 1994a, 1994b) studied the determinants of educational achievement among white school pupils in Cape Town.

Even the study of white South Africans, on whom there was data, was poorly developed. Elsewhere I have described South African election studies as a ‘disconnected’ tradition, developing in apparent ignorance of the work being published in international journals (Seekings, 1997). Even Henry Lever, who pioneered the use of opinion poll data, drew his inspiration from work done in the USA in the early 1940s and ignored subsequent work (including theories of partisan identification). Gouws and Gibson note that few political science departments offered the required training in research methodology and statistical analysis, and there was a widespread ‘bias against the supposedly “positivist” study of attitudes and use of statistics, under the mistaken assumption that these types of studies are ignorant of human complexity and context’ (Gouws and Gibson, 2001). Other key sociological topics were never explored even with respect to white South Africans. For example, there do not appear to be any studies of occupational change and inter-generational mobility among white South Africans, even though apartheid was in part a project to promote the interests of one section of the white population relative to the other.

Major survey-based research on the attitudes of black South Africans only really began in the late 1970s, with the prospect of political reform, accelerating in the 1980s as the National Party government debated and then implemented reforms designed to buy support or at least acquiescence among coloured, Indian and urban African constituencies. African people were included in surveys conducted in 1977 by Hanf et al (Hanf et al, 1981) and in 1981 for the Buthelezi Commission, investigating options for political change in Natal and KwaZulu. In 1982 the ‘Division for Group Interaction’ at the parastatal Human Sciences Research Council began to conduct annual surveys among white, coloured and Indian people (see Rhoodie and De Kock, 1983). In 1984 it conducted its first survey among a sample of African people in what is now Gauteng province (De Kock et al, 1985). The apartheid state’s interest in this data is not difficult to understand – although the quality of the data remains very suspect, given the political
climate in which respondents were being asked questions about political strategies, goals and loyalties. Little of this work was done by independent, critical scholars. Such scholars generally also lacked funds, even when they had an interest in quantitative data. The result, in the study of politics as of poverty, was that the best quantitative analysis was small-scale and focused on specific areas (see, as examples, studies conducted in Durban’s KwaMashu township by Schlemmer [1976] and Brewer [1985]).

This growth in the collection and analysis of quantitative data, on poverty, unemployment, migration and political attitudes, had less of an effect on South African social sciences than might have been expected. The key reason for this was probably the deep divisions in academic life that resulted from political differences. The deepest divide was between state-sponsored researchers, mostly in the Afrikaans-language universities and the parastatal research agencies, and opponents of apartheid in the English-language universities. In some disciplines (including Sociology) there were rival academic associations and journals. It is hard to overestimate the depth of the intellectual gulf that emerged. But these were not the only divisions.

Differences over how to oppose apartheid also translated into intellectual differences, with the new wave of radical or Marxist scholars reluctant to acknowledge positively any of the work of ‘liberal’ social scientists. In the early 1970s Marxist and neo-Marxist scholars mounted a sustained challenge to the prevailing liberal interpretations of South African history and society, and the political strategies that these implied. Whereas the liberals had emphasised Afrikaner nationalism and racial discrimination, the Marxists placed capitalism (and especially the mining industry) at the core of South African studies. Over time, however, Marxist structuralism became less a source of insight into the dynamics of South African society than a badge of intellectual and political correctness. Its empirical basis was often weak, and little or no use was made of quantitative data – especially since most of the existing quantitative political economy was done by non-Marxists who were dismissed as ‘liberals’ (see Nattrass, 1991). Marxist structuralism became an ever more theory-driven approach, rendering it vulnerable to refutation by critics who actually examined quantitative data on the key pillars of structuralist explanations (e.g. Simkins [1981] on agricultural production in the reserves, Bromberger [1983] and Moll [1991] on growth rates, and Nattrass [1992] on profit rates). The preoccupation with theory might have predated the Marxist turn (Jubber, 1983), but the embrace of Marxism surely gave theory-driven social science new vigour.
Marxist approaches rapidly became hegemonic, but they were far from homogeneous. In response to the empirical weakness of much (but not all) of Marxist structuralism, and especially to its apparent neglect of the agency of the black rural and urban population, there arose an alternative Marxist approach that the structuralists denounced as ‘empiricist’. These scholars, mostly in history departments, used archival sources and oral testimony to write richly-textured accounts of struggles over proletarianisation, on the land and in the towns, and of the political consequences of the transformation of South Africa’s economy and societies. The structuralist Marxists accused the neo-Marxist social historians of theoretical poverty amidst their headlong flight into empiricism, and of an incapacity to locate the experiences of ordinary people in a bigger picture of social forces and structural relations (Morris, 1987). The social historians countered that they had broken free of the ‘massive limitations’ implicit in the kind of evidence used by ‘abstract structuralist writers of the 1970s’ and were ‘exploring new ways of uncovering the past’ (Keegan, 1988: 3). Another strand of neo-Marxist research grew rapidly, in industrial sociology (Webster, 1985). Both the historical and the industrial research were empirically rich, but the empirical content was qualitative only. Structuralist Marxists, social historians and industrial sociologists shared an emphatic antipathy to quantitative work.

An antipathy among broadly Marxist scholars to quantitative research is not unique to South Africa. The leading American neo-Marxist sociologist, Erik Olin Wright, has noted the antipathy toward quantitative analysis that has long characterised Marxist social science: ‘Left-wing scholars, especially Marxists, are generally sceptical of quantitative analysis and have traditionally relied primarily on historical and qualitative methods in their empirical research.’ He attributes this scepticism in part to the fact that Marxists have been especially concerned with topics that ‘are not easily amenable to precise measurement and quantitative treatment’, but in significant part also to ‘a general hostility by many Marxists to anything that smacked of “bourgeois social science”’. Wright thus identifies his own objectives in his book Class Counts as ‘demonstrating the usefulness of class analysis to non-Marxists and the usefulness of quantitative analysis to Marxists and other radical scholars’ (Wright, 1997: 545-6).

One of the few topics in which there was quantitative research by scholars from different political backgrounds turns out to be an exception that ‘proves the rule’. When international pressures for economic sanctions and disinvestment intensified in the 1980s, state-sponsored and business-oriented researchers cited opinion poll data that ‘showed’ that black South Africans were opposed to disinvestment
Leftist scholars countered that there was broad support for sanctions, because they would lead to political change. In late 1985 the South African Council of Churches raised funds for the Community Agency for Social Enquiry (CASE) to conduct a survey of opinions on sanctions and disinvestment among African people in metropolitan areas. The CASE survey allowed respondents to choose between three responses: strong support for investment, strong opposition to investment, and an intermediate position specifying that firms should only invest if they actively pressurised the government to end apartheid. Approximately one quarter of the sample favoured each of the extreme positions, and one half the intermediate position. As Orkin (1986) pointed out, previous polls that posed the question in a binary either/or form forced the 50 percent who favoured his intermediate position to choose one or other of the extremes, and their choice varied according to the precise phrasing of the question. The CASE survey also probed political loyalties, preferred political strategies and political goals. This debate continued, encompassing attitudes to sanctions also (see Orkin, 1989). This is about the only area where leftist scholars engaged in serious survey-based research and analysis, generating a debate over both findings and methodologies. The existence of this debate was due entirely to political considerations.

Resources were clearly a constraint on the capacity of some researchers. But a hostility toward, or at least a marked lack of interest in, quantitative research was clearly also an inhibiting factor in the slow development of quantitative social science. The apartheid state had the resources, but not the inclination (at least not until the mid-1980s, when it sought some knowledge of what its oppressed subjects actually thought). Liberal researchers had the inclination, but rarely had access to the resources required. Marxist and neo-Marxist scholars generally lacked both resources and inclination.

New data on incomes, employment and unemployment

The years of transition from apartheid to democracy saw a transformation in the demand for and use of quantitative analysis. The discipline most profoundly transformed is undoubtedly economics, through the use of countrywide household survey data. Prior to 1994, as we saw above, the only available aggregate data was either produced by government agencies, in which case it excluded most of the population, or entailed crude modelling by independent scholars. An explosion of
independent studies at the local level had not been extended to the national level. After 1994, by contrast, researchers could work with a growing series of reasonably complete and reliable countrywide datasets. This new data had an extraordinary effect within South African economics, transforming large areas of teaching as well as research.

The first countrywide household income and expenditure survey in South Africa was conducted in 1993 by the South African Labour and Development Research Unit (SALDRU) at the University of Cape Town, together with the World Bank, under the title of Project for Statistics on Living Standards and Development (PSLSD). The PSLSD survey, also sometimes called the South African Living Standards Survey (SALSS), covered a sample of about 8,500 households. The PSLSD survey was followed by other surveys conducted by the government Central Statistical Service (CSS, since renamed Statistics South Africa): an annual October Household Survey (OHS), with a much larger sample of households, and a five-yearly, detailed Incomes and Expenditures Survey (conducted in 1995 and 2000).

The PSLSD survey resulted primarily from the need of the African National Congress, as the prospective future government, for better data, especially on the overall extent and dimensions of poverty, so as to be able to design effective strategies to combat poverty. In April 1992 a delegation from the ANC and allied trade unions met with officials of the World Bank in Washington. Funding was secured from the Danish, Dutch and Norwegian governments. Working closely with World Bank staff and leaders of the ANC and other organisations, a team of South Africans – mostly economists – designed and conducted the survey. The survey went into the field in the second half of 1993. Immediately after the elections of the ANC-led government in April 1994, preliminary tables presenting data from the survey began to be used in policy-making. In August 1994, after intensive cleaning of the data, a volume setting out key data, entitled *South Africans Rich and Poor: Baseline Household Statistics* (PSLSD, 1994), was released, together with a full database on computer disks (see Wilson and Horner, 1996; Wilson, 1996).

The PSLSD and subsequent CSS surveys have been used in reports explicitly written to guide policy-making as well as in a large and growing number of academic articles and papers. The first major official use of the data was in a report on *Key Indicators of Poverty in South Africa*, published by the government’s Reconstruction and Development Programme (RDP) Office in 1995. At about the same time the new South African government decided, after discussions with the World Bank and United Nations Development Programme (UNDP), to commission a more comprehensive report on poverty and inequality to
guide policy-making. This Poverty and Inequality Report was finally commissioned by the Office of the Vice-President in 1997, and completed in 1998. It combined statistical data from the PSLSD and OHS/IES data-sets with qualitative data from a participatory poverty assessment.

These new data-sets have also been used in an explosion of quantitative analysis, almost entirely by economists, including Haroon Bhorat, Justine Burns, Anne Case, Angus Deaton, Malcolm Keswell, Geeta Kingdon, Stefan Klasen, John Knight, David Lam, Murray Leibbrandt, Pieter Le Roux, Julian May, Peter Moll, Nicoli Nattrass, Dorrit Posel, Catherine Janisch, Servaas Van der Berg, Andrew Whiteford, Martin Wittenberg and Ingrid Woolard. Their studies cover a range of topics, including: the character and determinants of inequality and poverty; the extent, causes and consequences of unemployment; the relationship between education and earnings; and the distributional effects of government welfare systems (especially the old-age pension) and government taxation and expenditure more broadly. A good indication of the extent to which research has been transformed by household survey data is provided by the set of papers from a recent conference on Labour Markets and Poverty in South Africa, organised by the Development Policy Research Unit at the University of Cape Town (the papers are available on the website, at www.uct.ac.za/depts/dpru/fes01.htm).

This heavily economic research culminated in several recent volumes. Poverty and Inequality in South Africa: Meeting the Challenge, edited by Julian May (May, 2000), and Fighting Poverty: Labour Markets and Inequality in South Africa by Haroon Bhorat, Murray Leibbrandt, Muzi Maziya, Servaas van der Berg and Ingrid Woolard (Bhorat et al., 2001) both rely heavily on new quantitative data. Both volumes were generated by policy-oriented projects. The first was based on the Poverty and Inequality Report commissioned by the Office of the Vice-President, the second on a study funded in part by the Department of Labour. This research was also used extensively in the report South Africa: Transformation for Human Development, edited by Vivienne Taylor and funded by the United Nations Development Programme (Taylor, 2000). Articles using household data have been included in special issues of the journals Social Indicators Research (Møller, 1997), the Journal for Studies in Economics and Econometrics (Leibbrandt and Nattrass, 2000) and Social Dynamics (Seekings, 2001).

The impact of these new data-sets has been great, in part because of the rapid changes in computing hardware and software. Not only could the PSLSD data be distributed on disks, but very soon afterwards data could be sent or accessed through the internet. Slow and
cumbersome computers in dedicated laboratories gave way to ever more compact and powerful personal computers in individual researchers’ offices or even in their briefcases.

Much, perhaps even most, of this research has been highly technical, very focused and not readily accessible to most non-economists. As Bhorat writes of the study published as *Fighting Poverty*, ‘it did not translate the detailed, and sometimes extremely technical research, into information that was easily accessible to policy makers’ (in his foreword to Leibbrandt and Leibbrandt, 2000). Economists’ reluctance to explain in words the meaning of their equations, their affection for mind-boggling graphics (including, most recently, three-dimensional maps), and discipline-specific techniques all put off non-economists. For example, Leibbrandt and various co-authors have analysed inequality through the use of a series of econometric techniques that ‘decompose’ inequality. Few non-economists are familiar with the Theil-T, Theil-L and Atkinson measures of inequality, and many would be put off by multivariate modelling. Too much economics falls into what might be called a ‘technique trap’, when demonstrating mastery of technical modelling is deemed more important than providing clear answers to pressing questions. And economists often fail to acknowledge how sensitive their findings are to the specific techniques they choose to use. Yet their work should be of interest to many non-economists. Leibbrandt *et al*’s studies, for example, point to the enduring importance of inter-racial inequality in South Africa, relative even to other racially differentiated societies (such as Malaysia), but at the same time they indicate the rising importance of intra-racial inequalities. The precise balance between inter- and intra-racial dimensions is sensitive to the choice of technique, but the studies find that intra-racial inequality now explains a larger share of overall inequality than inter-racial inequality. Leibbrandt *et al* also calculate the relative importance of factors such as education, location, employment status and so on in the determination of income and the probability of poverty.

Some economists have tried to make their findings more accessible, especially to policy-makers. *Poverty and Inequality in South Africa* and, to a lesser extent, *Fighting Poverty* are written with non-specialist audiences in mind. In the case of *Fighting Poverty*, an easily assessed summary document was also produced (Leibbrandt and Leibbrandt, 2000). It is perhaps remarkable how some concepts have moved out of the realm of specialist jargon and become almost commonplace. The Gini coefficient, to take the most notable example, has become widely used in press reports concerning inequality in the distribution of income, whilst quintiles and deciles have become quite
familiar concepts. Some findings by economists have also entered into public discourse. A good example of this is the analysis by Whiteford and van Seventer (2000) of trends in income distribution. Their finding that overall patterns of inequality did not change in the 1990s, as a decline in inter-racial inequality was offset by an increase in intra-racial inequality, was widely reported in the press and has been widely cited by non-economists. But much of the work done by economists has been slow to enter wider debate, even when it is on closely related themes (such as Moll’s work on declining racial discrimination [Moll, 2000] and Hofmeyr’s on labour market segmentation [Hofmeyr, 2000]). The special issue of *Social Dynamics* (Seekings, 2001) was explicitly intended to counter these patterns, rendering recent work by economists more accessible to non-economists and showing how economic analysis can contribute to a broad understanding of contemporary South African society.

Non-economists’ neglect of economists’ work is especially striking in the case of fiscal incidence analysis, i.e. the study of the incidence (or distribution) of the costs of taxation and the benefits of public expenditure. The pioneer in this field in South Africa was McGrath (1983), who examined closely redistribution through the budget under apartheid. He showed that there was a low degree of redistribution from white to black people under apartheid, because white people paid an even larger share of taxation than they received as a share of public spending. In the mid-1990s one of McGrath’s students, Catherine Janisch, examined the extent of redistribution on the eve of democracy in 1993/94, showing that there was a very high level of redistribution from rich to poor, from white to black (Janisch, 1996; McGrath *et al.*, 1997). A key factor in this was South Africa’s surprisingly redistributive old-age pension system (see Ardington and Lund, 1995; Case and Deaton, 1998). Van der Berg’s paper in the special issue of *Social Dynamics* – part of a larger study commissioned by the Department of Finance – identifies shifts in the early post-apartheid period. Using income quintiles as well as racial categories, he argues that social spending became even more redistributive after 1994 (Van der Berg, 2001). Indeed, white people actually began to receive smaller benefits per capita than black people, as they removed their children from public schools, stopped using public hospitals and rarely qualified for means-tested welfare payments. This finding flies in the face of those left political economists who argue that the post-apartheid regime simply serves the interests of capitalists, and those cynics who believe that nothing has changed since 1994. If the post-apartheid state is serving the interests of any special interest group, it is teachers. As Van der Berg shows, the big shift in the incidence of social spending was
largely due to a big increase in spending on teachers’ salaries in poorer parts of the country.

One of the problems with using data at the household level is that intra-household inequalities are glossed over. Economists have extracted some important findings about intra-household relations from cross-sectional datasets (such as the PSLSD and OHS/IES datasets). For example, Duflo (2000) has found that children living in households are heavier and taller – and presumably therefore better fed – when household income takes the form of a pension paid to a resident grandmother than when income is received by other household members. Posel (2001) uses evidence on remittances to explore how resources are allocated within extended family networks. Remittances, mostly sent by urban workers to rural kin, have long constituted a private welfare system, previously more important but now less important than the public welfare system. Because household surveys typically treat non-resident migrant workers as separate households for the purposes of data collection, they collect data on remittances even when none is collected on any (other) intra-family transfers. Posel uses this data to explore ‘how households work’. She shows that migrants do not remit to ‘households’ as a whole but rather to individuals within households, and the value of their remittances depends on the closeness of the kin relationship. She concludes that surveys need to collect more information on who is sending remittances received by households in the sample, and conversely who receives the remittances sent by migrant respondents.

One recent survey in the Langeberg district of the Western Cape sought to focus directly on intra-household distribution by interviewing every adult member of the sampled households, collecting data on income and expenditure and transfers within the household, as well as on health issues (Case and Wilson, 2001). Analysis of the data remains incomplete, but the study promises to add substantially to our understanding of both the dimensions of inequality and the methodologies of interrogating it. A paper by Case (2001) uses Langeberg data to explore how the distribution of pension income in households affects the health of their members.

One of the key factors underlying poverty and inequality in South Africa, as we have already seen, is the very high unemployment rate. Kingdon and Knight’s article in the special issue of Social Dynamics summarises several larger research projects that add to existing analyses of unemployment (see especially Bhorat and Leibbrandt, 1996; Klasen and Woolard, 1999; Nattrass, 2000). In South Africa, as most famously in the U.K. when Margaret Thatcher was Prime Minister, the post-apartheid state has sought to reduce the apparent
severity of the unemployment problem by adjusting the definition of unemployment. One of the things that Kingdon and Knight (2001) do is to challenge the appropriateness of the officially-favoured ‘narrow’ definition of unemployment in South Africa today.

A full analysis of poverty and inequality requires some understanding of changes over time. The temporarily unemployed are clearly in a different position to the long-term unemployed. Cross-sectional data, i.e. from surveys that each use a different sample or cross-section of the population, can identify what factors correlate with unemployment, but are no substitute for longitudinal surveys that identify the factors that cause individuals to get jobs or remain jobless. This is where studies such as the KwaZulu-Natal Income Dynamics Study (KIDS), described by May and Roberts (2001), are really important (see also Koen, 1999, on the value of longitudinal studies).

KIDS involved revisiting in 1998 households that had been interviewed for the PSLSD survey five years previously. The PSLSD sample was thus treated retrospectively as a ‘panel’ to be reinterviewed. With data from the two ‘waves’, researchers can begin to analyse the dynamics of household formation, the labour market, poverty and inequality. As May and Roberts (2001) report, the data shows a degree of flux that surprises some observers. One-third of the households that were below the poverty line in 1993 were above the poverty line in 1998, whilst a minority of richer households had moved in the opposite direction. Longitudinal data allows analysts to distinguish between the chronically poor and the transitarily poor (although it would clearly be useful to have data on more than just two points in time). May and Roberts also point to the wide range of other studies that have used the KIDS data to generate policy-relevant findings. Cichello, Fields and Leibbrandt (2001), also using the KIDS data, focus on dynamics in the labour market. They confirm that there is a high degree of flux in employment. They show that analysis of the two surveys as if the data was cross-sectional would lead researchers to conclude that individuals in formal employment enjoyed rising earnings between 1993 and 1998. But analysis of the data as a panel data-set reveals that there had been large changes in the composition of the formally employed, with a surprising proportion of the sample being employed in 1993 but not in 1998 (or visa-versa). Job loss meant that, on average, individuals in formal employment in 1993 actually experienced a markedly lower income in 1998.

In the short-term, losing or finding a job makes the biggest difference to earnings, but probably the single most important factor shaping changes over time is education. Education can serve either to reproduce or to mitigate the inequalities that children bring into the
classroom. Van der Berg (2001) shows that there has been a major shift in public education expenditure away from the rich and toward the poor, such that resources are distributed in pro-poor ways. But it is nonetheless unlikely that children from different social backgrounds face equal opportunities in the classroom. Crouch and Mabogoane (2001) show that formerly African schools in poor areas have markedly lower pass rates even when they control for school resources. Their analysis, using admittedly crude data for Gauteng, suggests that ‘schools in very poor areas tend to have matriculation pass scores some 20 points lower than schools in richer areas’. This implies that even a progressive allocation of public resources might be insufficient to offset the disadvantages that children in such schools suffer.

Family background clearly matters in South Africa, as elsewhere in the world. The PSLSD and OHS surveys show that there is a positive relationship between mothers’ education and the schooling of their children. African children whose mothers have twelve years of schooling are approximately two full grades more advanced at the ages of both thirteen and sixteen than African children whose mothers had fewer than four years of schooling. Family structure is also correlated with schooling outcomes, with children living with both genetic parents performing better than children living in other domestic arrangements (see Anderson et al., 2001). As we shall see further below, however, there are huge gaps in our understanding of education in South Africa, reflecting in part the inadequacies of the extant data and in part the deficient analysis of data that has been collected.

New fields of quantitative research in the social sciences

The use of household surveys of economic activity represents the most dramatic change in quantitative research after apartheid, but it is not alone. In several other fields there have been important shifts. The study of political behaviour is one such field. As we saw above, the study of political behaviour was very weak prior to 1990, but since then has been transformed rapidly.

The most important aspect of political behaviour to have been transformed is the study of elections and voting behaviour (see the review in Seekings, 1997). Before 1994, of course, the only elections of any importance involved white voters only. The holding of inclusive, non-racial elections in 1994 meant that, for the first time, the opinions of black South Africans were of considerable importance. In the run-up to
the elections, therefore, there was a proliferation of opinion polls with multi-racial, national samples, conducted for the competing political parties and for independent researchers (Mattes, 1994). The parties used data from opinion polls to design and modify their election campaigns. Academic researchers used the data to develop much fuller analyses of voting behaviour than had been possible hitherto. Some studies of the election made little or no use of polls, maintaining the existing, heavily descriptive tradition of election studies (see Reynolds, 1994). Some used polling data extensively, but with little attention to the theories of voting behaviour developed elsewhere in the world during the previous decades (see Johnson and Schlemmer, 1996). Others, however, used polling data to develop entirely new kinds of analysis in South Africa, using concepts such as partisan identification and issue-voting that are the mainstays of voting analysis elsewhere (see, especially, Mattes, 1995).

By the time of the second national and provincial elections, in 1999, the use of opinion polls had become far more widespread. A large part of the credit for this should go to the Opinion '99 consortium, involving the Institute for Democracy in South Africa (IDASA, through its polling expert, Bob Mattes), the Electoral Institute of South Africa (EISA), the polling company Markinor and the South African Broadcasting Corporation (SABC). Studies of the 1999 election made much fuller use of polling data (see Reynolds, 1999; Lodge, 1999; Taylor, 1999), and Mattes and others began to develop innovative models of voting behaviour (Mattes et al, 1999).

Other aspects of political culture were also interrogated in innovative ways using large-scale survey data. Amanda Gouws and James Gibson (2001) report on their research into tolerance and intolerance. Despite the supposed importance of intolerance and intimidation in South Africa, remarkably little research was conducted until Gibson and Gouws conducted countrywide surveys in 1996 and 1997. Their research included the use of experimental vignettes to probe how intolerance varies in different contexts, and an experiment to assess how easily people could be persuaded to change their minds and be more (or less) tolerant. Gibson and Gouws found high levels of intolerance in South Africa, as expected. More worryingly, they found it easier to persuade respondents to become less tolerant than to become more tolerant. Their work has been published in leading political science journals (including Gibson and Gouws, 2000) and is forthcoming as a book. They have also conducted survey research on attitudes towards justice, reconciliation and the Truth and Reconciliation Commission.

Hennie Kotzé (2001) reports on his participation in the cross-national World Values Study. Surveys were carried out in South Africa in 1981, 1990, 1995 and 2000 (although the sample for the first survey
was very limited). Kotzé has also been engaged in a longitudinal study of elite attitudes in South Africa, entailing six surveys between 1990 and 2000. Other aspects of political culture have also been explored by Mattes, including most recently as part of the cross-national Afrobarometer project (see www.afrobarometer.org). Somewhat related work is Møller’s ongoing project on trends in perceived quality of life (see Møller, 1998; also Brodie et al, 1999).

Another of the social sciences to have been transformed by the availability of quantitative data is criminology. Until the mid-1990s there was no credible data on crime, excepting perhaps murder. Since the end of apartheid the government has been committed to better recording of crime trends by the police, but it is unclear whether reporting rates really have risen and there remain doubts about the political manipulation of data. It is therefore difficult to analyse trends, although it is clear that crime rates, including especially violent crimes such as murder and armed robbery, are higher in South Africa than in all other countries for which there is any data. But, in the late 1990s, data on aspects of crime became available from surveys of victims of crime, and has become widely used (see, for example, Shaw and Gastrow, 2001). Between 1997 and 1999 the Institute for Security Studies organised victim surveys in four major cities as well as a sample of rural areas (see Camerer et al, 1998; Louw et al, 1998; Pelser et al, 2000). In late 1998 Statistics SA conducted the country’s first national Victims of Crime survey (Statistics SA, 1998). Data from these surveys allows researchers to identify how crime affects different social groups, as well as information on (for example) where and when crimes happen and whether the perpetrator is known to the victim. The surveys show that the poor experience more violent crime but less property crime than the rich, although both violent and property crimes are far from uncommon across the board. One of the more surprising findings was the very high rates of stock theft in rural areas: households owning stock are more likely to suffer stock theft than the owners of any other kinds of property, including car-owners. Criminologists also examined, through surveys, popular attitudes towards punishment (e.g. Schonteich, 2000). As further surveys are covered in future, they will allow analysis of changes over time.

Demography is a discipline on the verge of transformation as new data is generated, especially data from two ‘demographic surveillance surveys’. In Agincourt, in the Northern Province lowveld, and Hlabisa, in northern KwaZulu-Natal, extraordinarily detailed demographic data is being collected, on a quarterly basis in the case of Hlabisa. Not only will this provide extraordinarily rich evidence on
fertility, morbidity and mortality trends, but it will also revolutionise quantitative analyses of household formation and dissolution.

Sociology is probably the discipline that has been least transformed by the availability of new data. Indeed, the most notable quantitative work in post-apartheid sociology has involved the use of existing, not new, sources. Owen Crankshaw’s *Race, Class and the Changing Division of Labour Under Apartheid* (Crankshaw, 1997) represented a breakthrough in the study of class in South Africa. Using the government’s Manpower Surveys, which collected detailed data on many African as well as white, coloured and Indian workers, Crankshaw plotted the changing occupational division of labour under apartheid and the changing racial composition of different occupations. He documented the extent and pattern of African advancement under apartheid, explaining it in terms of the relative power of workers, employers and the government. This work combines with the work of economists in explaining changes in the distribution of wages and incomes in society under apartheid.

Prior to Crankshaw’s study, most work on class was devoid of quantitative content. There were minor exceptions to this. First, some Marxist scholars had attempted to study the size of particular classes, but limited their attention to those ‘atypical’ classes that defied a crude identification of class and race – i.e. the white working class (Davies, 1973) and the African middle class or petty bourgeoisie (Wolpe, 1977). Secondly, Simkins and Hindson (1979) had conducted some valuable preliminary work on the occupational class structure, but this had been largely ignored, in part no doubt because it employed a Weberian emphasis on skills at a time when the mainstream of the social sciences at the leading South African universities was resolutely Marxist. It is remarkable how little empirical work there was on the class structure at a time when class was central to so many scholars’ analytic frameworks.

The importance of Crankshaw’s work lies not only in its empirical quality. It is important also because it draws on approaches that had hitherto remained largely separate. The book was explicitly concerned with the class analysis championed by Marxist structuralists, and engaged with the bigger picture with which they were concerned. At the same time it drew on a qualitative research tradition to grapple with changes in the workplace as the labour process was itself transformed. Also it followed in the footsteps of scholars internationally (Wright being an example) in a systematic engagement with quantitative data. Crankshaw was perhaps the first sociologist working within a Marxist tradition, however loosely, to engage fully with the work of the so-called liberal political economists (including Simkins, McGrath and Hofmeyr).
Using Manpower Surveys meant, however, that the analysis remained incomplete, because the surveys covered only the population working in the formal sector, and excluded farm and domestic workers as well as subsistence farmers, other self-employed workers (including in the informal sector) and the unemployed. He was therefore unable to examine the class structure of society as a whole. Using household survey data, which covers a representative sample of the entire population, I have mapped quantitatively the class structure of South Africa at the end of the apartheid era (Seekings, 2000; Seekings and Nattrass, forthcoming). Household survey data has the disadvantage of much less precise data on occupations, but its wider coverage allows analysis of the class position of the unemployed (including especially households comprising only unemployed men and women), households with members in different occupations, and people who are themselves employers, are engaged in entrepreneurial activity or have income in the form of rent, profit or interest on assets (for a more economistic class analysis, see Carter and May, 2001). Using this data Nattrass and I extend the argument that there has been a shift from race to class as the central feature in patterns of inequality in South Africa (Nattrass and Seekings, 2001; Seekings and Nattrass, forthcoming). We also argue that South Africa has a large ‘underclass’ of households where no one works and no one is likely to find work, i.e. a class defined by unemployment and its exclusion from the wage relationship.

The use of household survey data also allows for a preliminary analysis of inter-generational mobility, i.e. the extent to which inequality and stratification is reproduced across generations as the children of disadvantaged parents end up disadvantaged themselves whilst the children of advantaged parents end up in privileged positions. One might expect that the end of apartheid, and legislative restrictions on the upward mobility of talented black South Africans, would lead to high rates of upward mobility (matched, perhaps, by downward mobility among some white families). The first surveys of the 1990s – such as the PSLSD survey – allow a preliminary analysis only. Most strikingly, they show that children in higher-class households progress significantly faster and further through school than children in lower-class households, even when we look at the African population only (Seekings and Nattrass, forthcoming). This suggests that intra-racial inequalities are being reproduced across generations. A full analysis of mobility requires more recent and preferably longitudinal data. More information should be forthcoming from the Khayelitsha-Mitchell’s Plain survey conducted in late 2000 in Cape Town, which asked retrospective questions about inter- and intra-generational change, and from
longitudinal studies such as KIDS (and other, imminent surveys such as the Cape Area Panel Study, see below).

Where sociology has more obviously if slowly been transformed is in the use of small-scale survey research, including especially studies of inequality and stratification in selected urban areas. Crankshaw’s study of the African township of Bekkersdal, for example, revealed a growing coincidence between occupation, income and residential area within the township (Crankshaw, 1996). In 1997, Crankshaw was also a key member of the team of sociologists from the University of the Witwatersrand who organised a pioneering survey in Soweto, the massive complex of townships south-west of Johannesburg. The survey collected data on incomes, employment, household structure, migration, health, crime and the perceived quality of life (see Morris, 1999). With the exception of data on housing (see, for example, Crankshaw et al, 2000), much of this Soweto data remains regrettably under-used.

Perhaps the fastest-growing field of social research has been tied to medical research on health. Social and medical research have been combined in a variety of very different studies. One kind of study is the one-off survey with a large countrywide sample, focussing on public experiences, attitudes and needs on health issues. At the beginning of 1994 CASE conducted this kind of survey for the National Primary Health Care Network (Hirschowitz and Orkin, 1995). In late 1994 and early 1995 CASE conducted a much larger survey, with a nationally representative sample of 4 000 households (CASE, 1997). Four years later, at the end of 1998, a third survey was conducted, with a similar sample of 4 000 households (CASE, 1999). The first two surveys were intended to establish the need for health policy reforms, and provide a baseline against which change could be measured. The third survey was intended to assess what changes had been effected in the initial years of post-apartheid government. It found that South Africans thought that public health facilities had improved since 1994, but at the same time they thought that the health of both children and adults had deteriorated over this period. All three surveys were conducted with funding from the Henry J. Kaiser Family Foundation in the USA. The Kaiser Family Foundation also commissioned a major survey of 2 000 South African teenagers, conducted in September 2000 (Lovelife, 2001). The survey focused on health issues, including attitudes to sex, drugs and alcohol, and AIDS.

A second kind of study is the innovative Birth-to-Ten in the Johannesburg area. This was a longitudinal study, tracking a birth cohort over a period of time. The study focused on babies born in public hospitals between April and June 1990. The babies themselves were
subjected to a range of medical and psychological tests and their parents were interviewed, building up a remarkable database on child development. The study was initially driven primarily by medical researchers, and quickly generated a mass of papers in scientific journals. However, as the children grew older, they became more and more interesting to social scientists. In 2001 a book – *Mandela’s Children: Growing Up in Post-Apartheid South Africa* (Barbarin and Richter, 2001) – was published, summarising the non-medical data on the children to the age of six. In 2000 the study was extended as the Birth-to-Ten-to-Twenty study. The study promises to generate extraordinarily rich data as the participating children proceed through school, begin to experiment sexually, perhaps get pregnant, and face and make possibly difficult decisions about leaving school, looking for work and so on.

There is a further, third kind of study that combines medical and social research. The HIV/AIDS pandemic is clearly one of the most pressing challenges facing post-apartheid South Africa. Substantial funds have been made available for quantitative as well as qualitative research. One of the major projects in this field is located in the gold-mining area of Carletonville, on the West Rand. Infection rates are very high in the area, and not only among prostitutes and migrant mineworkers. The highest infection rates were found among adolescent girls. The research project was longitudinal, trying to track the effects on sexual attitudes, behaviour and health of a health care intervention directed at HIV/AIDS. A baseline survey was conducted in 1998, with further surveys in 1999 and 2000, and a related programme of qualitative research (see Williams *et al*., 2000). Another, longitudinal study, focused on adolescents and HIV/AIDS, was conducted in Durban, with surveys taken into the field in 1999 and 2001.

Whilst this mixed medical/social research has the potential to generate rich data for social scientists in a range of disciplines, to date it has been primarily psychologists who have teamed up with the medical researchers. Sociology departments routinely teach quantitative methods, but surprisingly few sociologists themselves conduct primarily or even partially quantitative research. New quantitative data has been slow to penetrate into sociology courses and journals.

Sociology may have been slow to embrace quantitative analysis but there remain other fields in which quantitative research has barely begun. One of the most striking cases of this is the field of education. A shortage of data continues to be a huge constraint on analysis of education. Unsurprisingly, the apartheid state collected little data on the causes of educational performance among black students. More surprisingly, it collected little data on white students. The post-apartheid state inherited an education system that collected very little data on any
education indicators. Luis Crouch, the leading quantitative social scientist to have worked within the national Department of Education, wrote as recently as 1999:

‘What one normally thinks of as ‘education indicators’ have not been widely used in South Africa’s past or, when used, were often used without sufficient critical understanding of their purposes or limitations. The tradition is hard to break. With a few (key, but often misused) exceptions, such as the Std. 10 pass rates, per learner expenditures and the learner-educator ratio, the system proceeds largely unaided by the use of indicators.’

Crouch (1999) muses on the effect of this on education management:

‘The experience must be something like that of flying a Boeing 747 with, say, only a fuel gauge and a weight gauge prior to take-off: no altimeter, no speed gauge, no radar, no engine temperature gauges, no compasses, no gyroscopes, no oil pressure gauges, etc. Exhilarating, but prone to surprises. Not an enviable position for the [national] Minister and the MECs [i.e., provincial ministers] to be in – one expects that if they want roller-coaster rides they would rather get them at an amusement park than in the sector they are accountable for.’

In the case of education, a hostility to quantitative analysis born out of the anti-positivist stance of South African social sciences, which had existed for many years, (see Muller, 1999) seems to have combined with the self-interest of teachers and educational administrators in opposing data collection that would allow them to be held to account.

The only credible data on student performance comes from the Standard Ten (or matriculation) examination. However, more than half of the country’s children do not reach Standard Ten. The Department of Education has unveiled plans for a range of assessment policies, but these are yet to be implemented and in any case will provide an incomplete set of data. Researchers have therefore had to use poor proxies for student performance. Economists examining returns to education use the highest school grade attained. Researchers interested in the effects of family and community factors use grade attainment by age. Yet we know that pupils are promoted into higher grades with extraordinary variation in their numeracy, literacy or other skills.
Students in the same grade record very different results when they sit standardised numeracy and literacy tests.

There are some data-sets that include literacy and numeracy test results for samples of students, along with information on the schools they attend and their home and community environments. At least three studies – the Third International Mathematics and Science Study (TIMSS) in 1995, a limited repeat in 1999, and the Monitoring Learning Achievement (MLA) project – were conducted in South Africa as parts of international studies. Very little use has been made of any of this data (see the descriptive reporting in Howie, 1997, 2001; Strauss and Burger, 2000 – but see also the preliminary analysis of TIMSS data in Howie and Pietersen, nd, and Howie and Wedepohl, nd). Research on the importance of family background seems to have been characterised by a glaring dichotomy: one group of economists has been trying to do complex work with very limited data-sets, whilst richer data-sets have barely been used. Education policy units seem to have little interest in quantitative research, and the national and provincial departments of education seem to lack the capacity to process fully, yet alone analyse fully, the data they already collect.

The most promising quantitative analysis of education is the generally small-scale work by consultants such as Eric Schollar, conducting evaluations of education reform programmes. In order to assess the effects of three- or five-year interventions, Schollar (among others) combines qualitative and quantitative research. The quantitative component typically entails literacy and numeracy tests at the outset, mid-way and at the end of the reform programme, with the tests administered both to a sample of students in classes covered by the programme and a control group of students not covered. Such test data is combined with data from questionnaires on the community and family background, data on school resources, and qualitative research including observations of lessons (see, for example, ESA, 2000). Such research, on a larger scale, would generate very valuable information on schooling, of immediate value to policy-makers and academics alike.

**Building capacity in quantitative social science**

The growth of quantitative social science in South Africa over the past decade has led to some exciting research. Much of the initial survey work involved the collection of data to fill gaping holes in our understanding of society, economy and policy, especially in response to
appeals for data from government departments in the post-apartheid state. More recent survey work has begun to refine the way in which data is collected, and what kind of data. Three examples of surveys conducted by researchers at the University of Cape Town indicates the shift. In 1999, together with economists from Princeton, Francis Wilson conducted a survey in the Langeberg district that collected data from all adult members of the sampled households, in order to explore intra-household inequality and different perceptions of overall household income. In 2000, an inter-disciplinary team conducted a survey in Khayelitsha and Mitchell’s Plain collecting data from all adult household members on labour market issues, to interrogate the significance of alternative definitions of unemployment and to probe the effects on individuals’ decisions in the labour market of the employment status of other members of the household. In 2002, together with researchers from the University of Michigan, the first wave in a multi-wave Cape Area Panel Study goes into the field. The overall study will collect data on a panel of young people over several years, as they progress through school and make the transition to the worlds of parenthood and work (or unemployment, as the case may be).

Of course, quantitative data requires critical analysis, and it is fitting that research centres such as the newly-established Wits Institute for Social and Economic Research (WISER) and the Centre for Social Science Research (CSSR) at the University of Cape Town have committed themselves to exploring critically the relations between qualitative and quantitative research. If quantitative data is to be understood fully, and surveys designed well, it is imperative that researchers draw on the insights of qualitative research.

Despite the considerable progress made over the past decade, quantitative social science remains very uneven in South Africa. Outside of economics, too few scholars are actively engaged in quantitative research to sustain serious intellectual debate over either findings or methodologies. Even within economics, it is striking how much of the new research is done by scholars from outside South Africa. In the 2001 special issue of Social Dynamics alone, six contributors (Anderson, Case, Cichello, Crouch, Fields and Lam) are American, and another two (Kingdon and Knight) are based in Britain. KIDS was a collaboration between researchers at the University of Natal and the International Food Policy Research Institute in Washington DC. Foreign expertise and links are also integral to quantitative work in political science: Gibson, who collaborates with Gouws, is American; much of Kotzé’s work is based on a study directed from the USA; Mattes, the pre-eminent quantitative scholar working on voting behaviour and other aspects of political culture, is also American (although resident in South Africa), and the
Afrobarometer project is driven from Michigan State University. Most large-scale surveys have been financed by grants from foundations or other sources in the USA and, to a lesser extent, Europe, and American or European researchers have often played the leading role in accessing these funds. Quantitative social science research on South Africa is still heavily dependent on skills and finance from elsewhere.

The limits to and constraints on capacity within South Africa in quantitative social science research and analysis are easily identified:

- the number of university-based social scientists engaged in quantitative research and analysis remains very small, especially in disciplines other than economics;
- quantitative research and analysis is weakly integrated into routine undergraduate and graduate teaching (i.e. outside of dedicated courses in research methodologies);
- there are few opportunities for graduate students to develop practical and analytical skills through ongoing participation in a series of major surveys;
- university-based researchers have insufficient time to analyse data-sets in depth, due to teaching commitments or the pressures of conducting surveys;
- too much data is either not publicly available, or is available only after long delays and in user-unfriendly forms;
- there has been too little accumulation of skills within institutions over time, so that surveys have not benefitted enough from the methodological and analytical lessons learnt in other, previous surveys;
- there is too little inter-disciplinary collaboration and exchange.

The dedicated research institutions (such as Statistics South Africa and the Human Sciences Research Council) are clearly in a somewhat different position to universities. They may have more resources for quantitative work, but do little in the way of capacity-building in terms of providing students with integrated course-based training and practical experience. They are also generally more skilled in the collection than the analysis of data, and are under pressure to produce policy-oriented results fast without time for critical reflection.

There is a number of ways in which the quality of quantitative social science research can be strengthened through building capacity. First, and most important, there needs to be a synergy between teaching and research. Much survey research (and indeed a great deal of social science research) in South Africa has been done under the auspices of dedicated research units staffed by full-time researchers. In the USA and UK, most major survey research projects are undertaken by scholars who
‘buy’ themselves out of teaching for a certain period. These models have clear advantages, especially given the special difficulties of combining survey research with a heavy teaching load. However they have a disadvantage in that they isolate quantitative research from teaching and graduate students from quantitative research, and they can leave teaching departments devoid of researchers who are actually involved in quantitative research and analysis. If students are to become fully proficient in the critical analysis of quantitative data, it is not sufficient for them simply to take dedicated courses in research methodology that cover survey work. They need also to encounter routinely the critical analysis of quantitative data in standard courses. For this to happen, quantitative social scientists must be involved in teaching beyond research methods courses. One way forward is for research institutes to ‘buy’ teaching staff out of some but not all of their teaching commitments. This will provide researchers with the time to conduct and especially to analyse quantitative research whilst maintaining the vibrant connection between teaching and research. This is the model adopted by the Centre for Social Research (CSSR) in Cape Town.

Secondly, special attention needs to be paid to the development of future research capacity through developing skills among graduate students. Many social science university departments require that their undergraduate and graduate students take a research methods course, often including practical projects. But the skills developed in these kinds of courses need to be strengthened through close participation in the design and analysis of larger surveys. The emphasis must be on quality more than quantity. The objective should be to ensure that more and more of the applicants for junior posts in university departments in disciplines such as sociology and politics have strong backgrounds in quantitative social science research.

Thirdly, continued efforts need to be made to strengthen quality and capacity through strong international links. Because funds and skills are more available in the USA (and Europe) than in South Africa, most major quantitative studies in South Africa entail partnerships between South African and foreign scholars. These partnerships contribute to strengthening capacity in South Africa, especially if they entail a component of training for graduate students and junior faculty as well as co-operation in research. Among the programmes funded by the Mellon Foundation in South Africa is a demography training programme, involving faculties from the universities of Michigan and Cape Town.

Fourthly, research surveys should be designed and conducted by the scholars who will analyse the data. Too often, data collection and
data analysis are done by different people: one team designs the questionnaire and organises the collection of data, which is later used by a largely different and often disparate set of researchers. For survey research to generate new understandings of social reality, it is important that surveys are designed in closer collaboration with social scientists wishing to answer specific questions (whether these are ‘problem-inspired’ or ‘curiosity-inspired’). This will help to ensure that teaching and research are synergistic, and that surveys do not collect data that sits, un- or under-analysed.

Fifthly, *inter-disciplinary research and interaction needs to be encouraged.* Surveys might be managed so as to provide opportunities for participation by diverse scholars, each contributing their own (but preferably interactive) modules. The CSSR in Cape Town is planning a regular series of surveys in the Western Cape, modelled on area studies in other parts of the world (most notably the Detroit Area Study, conducted annually since 1951 by the University of Michigan).

Sixthly, *wider use of existing and new data-sets should be promoted through more user-friendly data resource centres.* Too many data-sets are used by only a handful of scholars, because they are difficult to access or even because they are unknown. Improving access to data not only ensures better value for money, in that more analysis is done for each dollar spent on survey costs, but also promotes better analysis through facilitating fresh ideas and fostering debate. Improved access to data-sets will also boost the use of quantitative data in teaching. One of the units within the CSSR is the Data First Resource Unit, dedicated to improving access to quantitative data.

Finally, *better communication and even co-ordination within and between institutions should help to ensure that there is an ongoing accumulation of skills and knowledge, that lessons are learnt from mistakes, and unnecessary repetition in the collection of data is avoided.*
References


1990s. Unpublished report. Johannesburg: University of the Witwatersrand,
Department of Sociology.
Morris, M. 198. Social History and the Transition to Capitalism in the
South African Countryside. *Africa Perspective (New Series)* 1(5 &
6), 7-24.
Muller, J. 1999. Reason, Reality and Public Trust: The Case for
Educational Research for Policy. In Taylor, N., & Vinjevold, P.
(eds), *Getting Learning Right*. Johannesburg: Joint Education Trust.
Nattrass, N. 1991. Controversies about Capitalism and Apartheid in South
Studies*, 17(4).
-----, 2000. The Debate about Unemployment in the 1990s. *Journal for
----- & Seekings, J. 2001. Two Nations? Race and Economic Inequality in
-----, 1989. Politics, Social Change and Black Attitudes to Sanctions. In
Orkin (ed.). *Sanctions Against Apartheid*. Cape Town: David
Philip.
South Africa’s Rural Areas*. Monograph No. 47. Pretoria: Institute
for Security Studies.
Posel, Dorrit. 2001. How Do Households Work? Migration, the Household
and Remittance Behaviour in South Africa. *Social Dynamics* 27(1).
PSLSD (Project for Statistics on Living Standards and Development).
University of Cape Town: South African Labour and Development
Research Unit.
Conditions in the Native Locations of East London, Cape Province*. Cape
Town: Oxford University Press.
Future Prospects*. Cape Town: David Philip.
David Philip.
Rhodie, N., & de Kock, C. 1983. Coloured and Indian Perceptions of Socio-
Political Change in South Africa. In van Vuuren, D., Wiehahn, N.,
Lombard, J., & Rhodie, N. (eds). *Change in South Africa*. Durban:
Butterworths.


