

AGENCY FOR INTERNATIONAL DEVELOPMENT WASHINGTON, D. C. 20523 BIBLIOGRAPHIC INPUT SHEET		FOR AID USE ONLY <i>Batch 87</i> ARDA	
1. SUBJECT CLASSIFICATION	A. PRIMARY Social sciences		SD00-0000-G100
	B. SECONDARY Political science--Africa		
2. TITLE AND SUBTITLE Data reliability in cross-national research; a test employing black Africa country experts			
3. AUTHOR(S) Hill, K.Q.; Von der Mehden, F.R.			
4. DOCUMENT DATE 1978		5. NUMBER OF PAGES 31p. 33p.	6. ARC NUMBER ARC
7. REFERENCE ORGANIZATION NAME AND ADDRESS Rice			
8. SUPPLEMENTARY NOTES (Sponsoring Organization, Publishers, Availability) (In Program of Development Studies. Paper no.83)			
9. ABSTRACT <p> This paper reports the results of an analysis of data reliability for a number of Black African nations based upon the judgmental assessments of political scientist nation specialists. In almost every area, the modal response category was either 10 or 20% error, and the bulk of all responses fell into these categories taken together. Scholars executing single nation studies would be most concerned about such error margins, since the reliability of single data points is critical in such research. For many aggregate cross-national analyses, however, this may not be a very damaging average error level. Of more concern is the character of some specific kinds of data for some specific nations. Data on population and on agricultural and general development policy were more subject to error and manipulation than that of other areas. Likewise, some nations were indicated to have especially weak data overall. Information drawn from these areas or nations should be treated with caution. The strengths and weaknesses of the "expert judges" approach for evaluating data quality are discussed and suggestions are made for future uses. Additional refinements in future studies might include questions about specific data estimates, queries regarding the accuracy of data for several different timepoints, and the estimation of both the extent and direction of errors, to discern whether the overall error for a sample of nations would be random or systematic. </p>			
10. CONTROL NUMBER PN-AAF-004		11. PRICE OF DOCUMENT	
12. DESCRIPTORS Africa Data Data acquisition		13. PROJECT NUMBER 93105200	
		14. CONTRACT NUMBER AID/otr-C-1394 Res.	
		15. TYPE OF DOCUMENT	

**PROGRAM OF DEVELOPMENT STUDIES
121 Sewall Hall
WILLIAM MARSH RICE UNIVERSITY
Houston, Texas 77001**

Paper No. 83

**Data Reliability in Cross-National Research:
A Test Employing Black Africa Country Experts**

by

**Kim Quaile Hill
and
Fred R. von der Mehden**

Spring, 1978

The authors are Associate Professor of Public Affairs at the University of Houston at Clear Lake City, and Albert Thomas Professor of Political Science at Rice University, respectively. The research reported in this paper was supported by the Agency for International Development under contract AID/otr-C-1394 on "Distribution of Gains, Wealth and Income from Development."

Program Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Discussion Papers should be cleared with the authors to protect the tentative character of these papers.

Abstract

**Data Reliability in Cross-National Research:
A Test Employing Black Africa Country Experts**

by

Kim Quaile Hill

and

Fred R. von der Mehden

This paper reports the results of an analysis of data reliability for a number of Black African nations based upon the judgmental assessments of political scientist nation specialists. The results indicate an average error range of 10-20 percent across all the nations and variables tested. Issue areas reported to have especially weak data, as well as individual nations with especially high or low data quality, are identified. The strengths and weaknesses of this methodological approach for evaluating data quality are discussed and suggestions are made for future uses.

Data Reliability in Cross-National Research
A Test Employing Black Africa Country Experts*

It is well recognized and frequently lamented that the quality of cross-national social, economic, and political data is highly varying and, for many countries, of dubious reliability.¹ Some observers of this problem view it as so staggering as virtually to obviate any cross-national aggregate research. Others argue that, while considerable progress has been made in improving reliability, equally considerable problems remain. Despite this widespread recognition of reliability problems, many comparative scholars continue to rely rather uncritically upon extant aggregate data--largely because of a lack of alternatives. The situation for empirical researchers is aggravated by the fact that estimates

*An earlier version of this paper was presented at the annual meeting of the Western Political Science Association, Los Angeles, California, March 16-18, 1978. The authors wish to thank Hugh Stephens and especially Crawford Young for assistance in identifying the respondent sample. Bruce Cain and David Klingman provided helpful comments on an earlier version of the paper. Also, Dorothy Johnson and Kathleen Quinlan provided invaluable research assistance.

¹For comments on the quality of cross-national data, see Ted R. Gurr, Politimetrics (Englewood Cliffs, N.J.: Prentice-Hall, 1972), pp. 49-59; Ted R. Gurr, "The Neo-Alexandrians: A Review Essay on Data Handbooks in Political Science," American Political Science Review, LXVIII (March, 1974), pp. 243-252; Michael C. Hudson, "Data Problems in Quantitative Comparative Analysis," Comparative Politics, 5 (July, 1973), pp. 611-630; and Richard L. Merritt, Systematic Approaches to Comparative Politics (Chicago: Rand McNally, 1970), pp. 36-44.

of the reliability of specific sets of data are just as difficult to acquire as is high quality data.¹

The dimensions of the data quality problem allow no single optimal solution, whether for assessing the extent of reliability or for generating highly satisfactory indices. Nonetheless, a variety of research strategies might be employed in pursuit of either of these goals. The present paper reports the results of one such exercise intended both to assess the reliability of a body of cross-national data and to test one particular research strategy for making these assessments.

We have surveyed a set of country-specialists for Black African nations asking them to estimate margins of error in currently available data for these nations. Individual scholars were linked to specific nations for which they could be considered "expert" judges. The results of this study offer specific assessments of data reliability for a set of developing nations and provide an indication of the feasibility of an expert-judges strategy for improving cross-national measurement. We will consider the rationale for this particular error-assessment method, the procedures by which it was executed in our research, the findings, and then the implications of the overall project.

The Use of "Expert" Judges

Before discussing the results of our study, it is important to reflect briefly upon the methodological rationale and our expectations concerning the results. It must be emphasized that we consider this study an

¹There are some notable and commendable examples of data collection efforts which have demonstrated great concern for reliability problems. The most exceptional work of this character is certainly that of Donald G. Morrison, Robert C. Mitchell, John N. Paden, and Hugh M. Stevenson, Black Africa (New York: The Free Press, 1972). Also of note is Charles L. Taylor

exploratory test of the "expert" judges approach rather than a procedure of certain validity for arriving at reliability estimates. We were attracted to this approach for two reasons, because of previous related uses of presumed experts and because of some special characteristics of cross-national political research.

Expert judges have been employed for a variety of purposes in several other social science disciplines and, to a lesser extent, in political science itself. The greatest use of the approach has probably been in psychology, but the most typical applications there are not very analagous to our intentions.¹ Also, in the area of social forecasting, there exists a large number of "Delphi" forecasting studies which employ presumed experts to predict future occurrences in their fields of expertise.²

Another area of application, better known to political scientists, is in community power studies where local informants are sometimes asked to identify elites via the so-called "reputational" approach.³ Informants are presumed to have special expertise in estimating who are the most influential

and Michael C. Hudson's World Handbook of Political and Social Indicators (New Haven: Yale University Press, 1972). Even in these cases, however, it has seldom been possible to attach reliability estimates to specific data values in the collection.

¹See, as an example, Lee J. Cronbach, Essentials of Psychological Testing (New York: Harper, third edition, 1970).

²For a critique of the Delphi methodology, see Kim Quaile Hill and Jib Fowles, "The Methodological Worth of the Delphi Forecasting Technique," Technological Forecasting and Social Change, 7 (1975), pp. 179-192. For example applications and procedures, see Harold A. Linstone and Murray Turoff, eds., The Delphi Method (Reading, Mass.: Addison-Wesley, 1975); and J. P. Martino, Technological Forecasting for Decisionmaking (New York: American Elsevier, 1972).

³See Floyd Hunter, Community Power Structure (Chapel Hill, N.C.: University of North Carolina Press, 1953); and Raymond Wolfinger, "Reputation and Reality in the Study of Community Power," American Sociology Review, 25 (October, 1960), pp. 636-644.

citizens in the politics of their communities. The uses of expert judges most relevant to the present study, however, have been by Russell H. Fitzgibbon, Kenneth F. Johnson, and Raymond Nixon. Fitzgibbon and Johnson have carried out quinquennial surveys of Latin American area experts, requiring the judges to rate individual polities on a number of social and political characteristics.¹ The results of these studies are a number of judgmental scales for Latin American polities generated by averaging the responses of panelists in a given survey. Nixon² has, in a similar fashion, employed panels of expert judges to code nations in terms of freedom of the press.³ Since both these latter two enterprises have used area specialists to create new cross-national variables, it is only a short additional step to ask experts to estimate the reliability of existing aggregate data.

The second reason for our interest in this methodological approach concerns a special characteristic of cross-national political scholarship. There is one major and rather unique information pool for comparative politics which has remained largely untapped for purposes of systematic

¹See Russell H. Fitzgibbon, "Measurement of Latin American Political Phenomena," American Political Science Review, XLV (December, 1951), pp. 517-523; "A Statistical Evaluation of Latin American Democracy," Western Political Quarterly, IX (September, 1956), pp. 607-619; "Measuring Democratic Change in Latin America," Journal of Politics, 29 (February, 1967), pp. 129-166; Russell H. Fitzgibbon and Kenneth F. Johnson, "Measurement of Latin American Political Change," American Political Science Review, LV (September, 1961), pp. 515-526; and Kenneth F. Johnson, "Scholarly Images of Latin American Political Democracy in 1975," Latin American Research Review, XI (1976), pp. 129-140.

²See Raymond B. Nixon, "Factors Related to Freedom in National Press Systems," Journalism Quarterly, 37 (Winter, 1960), pp. 13-28; and "Freedom in the World's Press: A Fresh Appraisal with New Data," Journalism Quarterly, 42 (Winter, 1965), pp. 3-14 and 118-119.

³There have also been a number of other, more disparate, uses of panels of expert judges in cross-national analyses. For a review of these applications, see John E. Mueller, Approaches to Measurement in International Relations (New York: Appleton-Century-Crofts, 1969), pp. 249-252.

aggregate research. That information pool is, of course, the collective knowledge of area and country specialists. Based on close familiarity with indigenous data sources, some of these individuals have criticized aggregate comparative research for frequent reliance on data of dubious quality. While the differences in intent between these two research perspectives are often large enough to minimize the likelihood of close agreement, there certainly remains much in the knowledge of each group of scholars which can inform the other. Several previous reviewers of the state of cross-national data collection have urged the use of area experts in that enterprise.¹ The present study is intended to assist in such scholarly communication.

Despite the initial promise of this approach, there are a number of research problems which raise danger signals and which should be recognized at this point. Certainly, a central issue is just who is an "expert" on a given nation or region? This may ultimately be a question very difficult to resolve, but as discussed below, we have depended upon other experts to resolve the issue for the present study.

A second issue, related to the first, is whether "experts" themselves are likely to disagree on matters within their common area of presumed expertise. Given any expectation of heterogeneity among "experts," one would wish ideally to sample from a large number and to test for consensus in their responses. Some unique features of the present study disallow such a strategy, but we must remain attentive to its possible implications.

An ancillary problem may arise when one is concerned with "exotic" states which have received little attention from social science scholars.

¹See Alfred de Grazia, "What Indicates What?" American Behavioral Scientist, VIII (December, 1964), pp. 29-31; Michael C. Hudson, op. cit., p. 628; and John E. Mueller, op. cit., pp. 249-252.

In this situation there may be a severely limited pool of informants which may obviate the possibility of in-country comparison of expert opinion. This problem turned out to be of particular pertinence in Black Africa where we found only a few nations with a number of readily identifiable political scientist country-specialists. The latter difficulty can be aggravated further by the fact that many country experts may not be as knowledgeable of data reliability as they are of domestic politics per se. Older political scientists dealing with Third World countries were not often trained in quantitative methods. Some may even be antagonistic to such methods and the aggregate approach to comparative studies associated with them.¹ Thus, the interest of such individuals in data problems might be quite low. Just as important, some area experts may be knowledgeable regarding the quality of data on some topics but not on others.

The final concern one must have regarding the use of area specialists is whether they can, or will, submit to the kind of judgmental exercise demanded by our research questions. In one sense, this is merely a matter of the feasibility of employing other scholars in this fashion. It strikes, however, at the extent of comparability of research interests among different scholars and at whether country specialists would agree to the value of the questions that must be posed to them.

Despite the possible difficulties noted above, enough suggestions to pursue this approach exist in previous literature that it deserves systematic investigation. Our project was intended as much to provide insight regarding the method as to generate specific reliability estimates. Thus, our hope

¹See, for example, Robert E. Ward, "Culture and the Comparative Study of Politics, or the Constipated Dialectic," American Political Science Review, LXVIII (March, 1974), pp. 190-201.

is to develop both substantive findings of interest and guidelines for future uses of "expert" opinions for similar purposes.

Data Collection Procedures

Our intention in selecting a sample of respondents was to identify scholars who were specialists in the politics of particular Black African nations and require them to complete our questionnaire only for such nations. Some individuals were found to be suitable as respondents regarding two or up to three nations. Because our questionnaire was somewhat lengthy, was of uncertain difficulty, was expected to require some extended thoughtful consideration, and was to be executed by mail, we offered the respondents an honorarium for its completion. This fact, coupled with our desire to include a reasonably large number of nations in the study with only a modest research budget, restricted us to one expert per nation. Thus, we are unable to test for the extent of expert consensus regarding individual nations, but our data do allow for tests of several other matters of importance as outlined above.

To collect a pool of nation experts, we first searched the holdings of three university libraries for books published on the politics of single African nations or for multi-nation studies which indicated detailed knowledge of specific African nations. We supplemented the resulting list of authors with additional names taken from a search of articles on African politics published in four African studies journals in 1972-1976.¹ The complete list was subdivided by nation of presumed expertise and first

¹The journals were the Journal of Modern African Studies, the Journal of African Studies, the African Studies Review, and African Social Research.

choices for each nation were identified. Two African area experts were then asked to evaluate our lists and first choices, offering their own rankings of the scholars listed as well as any additional choices. Based on this multiple-step procedure we generated a final set of respondents for our initial mailing.

The survey instrument asked respondents to evaluate the extent of error in data for a number of specific aspects of three broad categories: demographic and social conditions, economic conditions, and government programs regarding economic growth and distribution. Respondents were asked to judge whether current publicly available official government data on a specific subject, such as total population data, were in error by 10 percent or less, 20 percent, 40 percent, 80 percent, or 100 percent. Respondents were also allowed to indicate whether particular kinds of data were available at all or were only partially available. A similar set of questions was posed regarding the data on certain topics reported in standard United Nations' publications. For the categories of official government data, the area experts were also asked whether the estimated extent of data unreliability was a result of poorly developed data collection systems or explicit attempts to "doctor" data and show more favorable conditions than actually exist.

The survey instrument was administered by mail initially with a letter inviting participation in the project, followed shortly by the complete questionnaire and then a follow-up mailing for tardy respondents. Invitees who demurred from the exercise (generally because they reported a change in research interest or because they deemed their interest too narrow to fit the needs of the survey) were replaced in the sample with other nation

specialists.¹

Estimated Data Reliability Levels

The first portion of our findings to be reported concerns the estimated margins of error in different kinds of data currently available in the public reports of these African nations. Tables 1, 2, and 3 report the assessments of our respondents for three broad categories of data-- on demographic and social conditions, on economic conditions, and on government programs related to economic growth.

The results in these tables support some interesting generalizations. Central among our findings is that the respondents seemed to encounter little difficulty in making these assessments. The scholars were quite willing to engage in this judgmental exercise, and visual inspection of the questionnaires gave us no obvious hint of systematic biases coloring the replies of the respondents. Many reported a continuing interest in the project and its ultimate results, as well. The general demeanor of the respondents, then, suggests that this usage of country-experts is feasible and falls within the bounds of worthy scholarship as perceived by such individuals themselves.

An examination of the actual distribution of responses in these tables is even more interesting. In all three of these tables the consistent tendency is for estimated reliability errors to cluster in the range of 10-20 percent. In each of the three tables, the first few specific variables-- like total population, population by states, and population by ethnic groups

¹Of those scholars originally invited to participate four declined the opportunity. Most of the nations involved were kept in the study by the choice of additional respondents. Six nations were dropped, nonetheless, because of our inability to find suitable respondents in a reasonable amount of time. These nations were the Congo (B), Dahomey, Togo, Chad, Zaire, and Guinea.

Those scholars who did participate in the survey, and who did not ask that their names be held in confidence, were: Robert Bates, Henry Bienen,

TABLE 1: Reliability for Demographic and Social Variables

Variable	Estimated Error Margin					Don't know
	10% or less	20%	40%	80%	100%	
Total population data (from the most recent census or survey)	9 (40%)	8 (35%)	4 (17%)	1 (4%)		1 (4%)
Population by administrative subunits such as states	6 (26%)	11 (48%)	4 (17%)	1 (4%)		1 (4%)
Population by major ethnic or religious groups	7 (30%)	8 (35%)	5 (22%)	1 (4%)		2 (9%)
Aggregate nation-wide education data as on enrollments, graduation rates, etc.	14 (61%)	7 (30%)				2 (9%)
Education data (as described above) by major ethnic or religious groups	5 (22%)	7 (30%)	2 (9%)			9 (39%)
Aggregate health and welfare data	2 (9%)	11 (48%)	4 (17%)		1 (4%)	5 (22%)
Health and welfare data by major ethnic or religious groups	2 (9%)	5 (22%)	4 (17%)			12 (52%)

TABLE 2: Reliability for Economic Variables

Variable	Estimated Error Margin					Don't know
	10% or less	20%	40%	80%	100%	
Aggregate GNP levels	6 (26%)	11 (48%)	1 (4%)	1 (4%)		4 (17%)
GNP growth rates (annual)	6 (26%)	11 (48%)	1 (4%)		1 (4%)	4 (17%)
Agricultural or primary products production data	6 (26%)	8 (35%)	3 (13%)	2 (9%)		4 (17%)
General industrialization data (such as energy output, factory output, coal and steel production, etc.)	12 (52%)	5 (22%)				6 (26%)
The general distribution of wealth in society, as shares of wealth or income by income groups		8 (35%)	3 (13%)			12 (52%)
Wealth or income levels by regional administrative units	1 (4%)	6 (26%)	5 (22%)			11 (48%)
Wealth or income levels by major communal or religious groups	5 (22%)	3 (13%)				15 (65%)

TABLE 3: Reliability for Government Program Data

Variable	Estimated Error Margin					Don't know
	10% or less	20%	40%	80%	100%	
Aggregate data on government "input" levels (program size, spending, level of government effort)	12 (52%)	9 (39%)	1 (4%)			1 (4%)
Data on "input levels by regional administrative units	6 (26%)	10 (44%)	2 (9%)			5 (22%)
Aggregate data on program "outcomes" (results, goal attainment, unanticipated consequences, etc.) in <u>education</u>	7 (30%)	12 (52%)	2 (9%)			2 (9%)
Aggregate data on program "outcomes" for general development programs	2 (9%)	14 (61%)	5 (22%)			2 (9%)
Data on "outcomes" by regional administrative units in <u>education</u>	5 (22%)	12 (52%)	3 (13%)			3 (13%)
Data on "outcomes" by regional administrative units for <u>regional investment</u> programs	2 (9%)	5 (22%)	8 (35%)			8 (35%)
Data on "outcomes" by regional administrative units for agricultural programs	3 (13%)	8 (35%)	6 (26%)	1 (4%)		5 (22%)
Data on "outcomes" by regional administrative units for <u>general development</u> programs	2 (9%)	6 (26%)	7 (30%)			8 (35%)
Data on taxation such as tax incidence by income groupings	5 (22%)	7 (30%)	2 (9%)			9 (39%)
Data on taxation revenues broken down by regional administrative units	7 (30%)	8 (35%)	2 (9%)			6 (26%)

in Table 1--represent the most prominent indicators in the general category. On these prominent indicators 15 to 21 of our 23 respondents estimate error margins of 10 or 20 percent.

A second observation is that there are few estimates of very high error margins. In all three tables, only two estimates are in the 100 percent category and only a few scattered codings are for 80 percent error. Furthermore, if one examines those categories where relatively small numbers of nations are reported to have low error margins--the largest number of responses are not for high error estimates but are for "don't know" replies.¹ Most of these clusterings of "don't know" responses come in regard to highly specific kinds of data, often broken down by ethnic groups or regional administrative units. The pattern of responses suggests not that data are less reliable in these cases but that (1) we have simply gone beyond the degree of specificity at which our respondents can comfortably make judgmental assessments or (2) data are not available. For example, all three questions related to levels and distribution of income and wealth showed high percentages of both "don't know" and data unavailability.

Sources of Unreliability

In those cases where they estimated error margins to exceed 20 percent, we asked the respondents to judge whether such unreliability was a result of the quality of the data collection system or of governmental tampering

Robert B. Charlick, William J. Foltz, Harry A. Gailey, Fred Hayward, Rene Lemarchand, Victor Levine, James H. Mittelman, Louis A. Picard, Christian P. Potholm, Elliott Skinner, Richard E. Stryker, and Richard Weisfelder.

¹Most of the "don't know" replies were from scholars who reported that, for the particular variable under consideration, such data were either not available at all or were only partially available.

with the true figures. In Table 4 we report the responses to these questions for our three broad categories of data.¹

Across the board, the modal positive response was that the character of the administrative system was the principal source of data error. From a quarter to a third of all the respondents chose this interpretation for almost every category of data.

In terms of reported tampering with data, the results are more varied. In most areas, no more than one or two scholars responded affirmatively to this possibility. For population figures and for general development and agricultural policy outcomes, however, about a quarter of the respondents agreed there was "doctoring" of data. Apparently, these are areas of special sensitivity in these developing nations. Scholars working with information on those topics might have special concern as to the nature of such data.

The opportunity for manipulating data and for general data unreliability would seem especially high for population figures given the age of the most recent national census or survey in most of these nations. For eight nations it was reported that the most recent census or survey was over ten years old; for another eleven nations the most recent one was five to ten years old. Nevertheless, our respondents' estimates of population data reliability levels are apparently little influenced by the age of the most recent census or survey. The rank order correlation between the number of years since the last census and estimated population reliability level was .17. While five

¹The actual numbers of responses in Table 4 sometimes exceed the number of respondents making error assessments of 40 percent or more. Some scholars who reported lower error rates completed this portion of the questionnaire anyway, and some checked both possible sources of extant error margins.

TABLE 4: Sources of Unreliability--Numbers of Responses

Type of Data	Poorly Admin.	Developed Systems	Governmental to "Doctor"	Efforts Data
	Yes	No	Yes	No
<u>Demographic and Social Data</u>				
Population	7	2	6	4
Education	3	1	1	3
Health	7	1	2	2
<u>Economic Data</u>				
Gross National Product	5	1	1	3
Agriculture	7	1	4	3
Industrialization	0	1	0	3
Income distribution	8	1	2	3
<u>Government Programs</u>				
General government "inputs"	3	1	2	4
Government "inputs" by regional units	7	2	2	4
Education policy "outcomes"	3	2	2	4
Development policy "outcomes"	9	1	6	3
Agricultural policy "outcomes"	8	1	6	3
Regional investment "outcomes"	8	1	4	3
Taxation	7	1	1	3

respondents did report rather high error levels for such data in Table 1, most of the remainder estimate that error margins are not extraordinarily high. It may well be that the procedures for estimating year-to-year population growth are rigorous enough to merit this confidence. But it may also be the case that our respondents are too generous in their reliability estimates. Unfortunately, this is a matter we are unable to pursue further with the information available to this study.

Reliability in United Nations Data

The preceding results are based upon country-experts' assessments of country-specific data and sources. It is seldom the case, however, that the nonspecialist or the aggregate analyst makes direct use of indigenous sources. Instead, the latter is likely to employ statistics drawn from publications such as the United Nations Statistical Yearbook, the Demographic Yearbook, or other more specialized volumes. It is not always clear to the nonspecialist that these secondary sources are as accurate or as complete as the primary ones. To provide some evidence on this issue, we also asked our respondents to estimate margins of error in four broad categories of data (demographic and social, economic, general government programs, government programs relating to economic growth) as reported in standard United Nations publications. In general, the distributions of responses presented in Table 5 mirror those reported earlier for indigenous data. The replies cluster in the 10 and 20 percent error response cells, and the vast bulk of the remaining replies are "don't know." These results suggest that the major secondary sources reflect the same general level of data quality as those available to the country-expert. These findings are not too surprising given the fact that a considerable amount of United Nations

TABLE 5: Reliability for Standard United Nations Data

Variable	Estimated Error Margin					Don't know
	10% or less	20%	40%	80%	100%	
Demographic and social conditions	3 (13%)	9 (39%)	4 (17%)			7 (30%)
Economic conditions	3 (13%)	12 (52%)	4 (17%)	1 (4%)		3 (13%)
Government programs, budgets, etc. in general	4 (17%)	12 (52%)				7 (30%)
Government programs regarding economic growth and distribution	13 (57%)	3 (13%)				7 (30%)

statistics are provided by member governments. This, of course, means that such documents do not provide a cross-check on the reliability of national data. It is encouraging, nonetheless, that the non-specialist has access to data of about the same general quality as does the country-expert.

The one major qualification regarding the assessments of United Nations data concerns availability. For each of the four kinds of data examined in Table 5, from a quarter to a third of our respondents reported that such information was generally not available or was of limited availability in standard UN sources. Evidently, and not surprisingly, there is either a time-lag in the transfer of information from primary to secondary sources or there is simply a significant screening in the amount of information so transferred. Most probably, both of these factors are operant for many nations.

Nation-Specific Results

The preceding analyses describe the findings when generalized for our entire set of nations. For some purposes, however, we might be concerned with the specific results regarding individual polities. Our survey results allow us to present several interesting "scores" for the individual nations regarding the quality of their data.

We have developed four discrete scales which indicate the relative quality of different aspects of data available from these individual nations. Three of these scales are simple additive indices for the number of times a respondent reported in the questionnaire that some kind of data was subject to "doctoring," was unreliable due to administrative system weakness, or was partially or completely unavailable.

We have also developed a fourth scale which indexes the average reported error margin across a number of specific types of data for a given nation. To construct this index, we chose 11 items from the questionnaire and assigned the numbers 1, 2, 3, 4, and 5 to the response categories of 10, 20, 40, 80, and 100 percent error margins. For each nation, we then averaged these assigned numbers for these items to generate an overall reliability score.¹ The exact interval distances separating nations on this index clearly should not be construed to indicate substantive variations. Neither the original data nor the index construction technique are robust enough to warrant such an inference. The rank ordering on this scale is, nonetheless, representative of the relative reliability levels reported for individual nations across the range of categories.

In Table 6 we report the nation scores on our four indices for average error margin, data non-availability, administrative weaknesses, and for "doctoring" of data. Examination of these results should be of greatest interest to the African-area specialist or to the individual considering the use of information from any of these nations. Visual inspection indicates that there are several nations which are reported to have especially satisfactory data overall and a few others reported to have relatively low data quality.

¹The items chosen to construct this scale were the error margin estimates for total population, population by administrative geographic units, population by major ethnic groups, aggregate national education data, aggregate health and welfare, aggregate GNP, agriculture production data, governmental "input" levels (program size, spending, etc.), educational program outcomes, general development program outcomes, and agricultural outcomes by administrative units. These particular items were chosen because of their substantive importance and because they exhibited some variance in the respondent's codings.

TABLE 6: Data Quality for Individual Nations^a -

Nation	Average Error margin ^b	Data Non-availability ^c	Admin. weaknesses ^c	"Doctoring" of data ^c
Tanzania	1.1	0	0	0
Kenya	1.3	0	0	0
Ghana	1.4	0	0	0
Rwanda	1.4	9	0	0
Botswana	1.6	5	6	0
Gambia	1.6	8	0	0
Swaziland	1.6	0	0	0
Upper Volta	1.6	0	0	0
Zambia	1.6	3	2	0
Burundi	1.7	6	2	1
Gabon	1.8	5	3	1
Senegal	1.8	3	6	2
Cameroon	2.0	2	2	3
Lesotho	2.0	4	6	5
Sierra Leone	2.0	0	3	0
Uganda	2.0	3	0	0
Nigeria	2.1	0	9	2
Somali	2.1	0	5	0
Ivory Coast	2.2	3	7	7
Mali	2.6	9	9	6
Niger	2.7	4	7	5
Central Afr. Emp.	2.9	7	8	5
Liberia	- ^d	0	0	0

^aRank ordered by "Average Error Margin" scores.

^bComputed from responses to 11 reliability questions as described in the text.

^cThe number of times the respondent noted such a deficiency in available data.

^dNot computed because of too high a proportion of missing responses.

For example, Ghana, Kenya, and Tanzania have especially high reliability (low scores on the reliability index) and were never judged to have weaknesses on the other three scales. On the other hand, the Central African Empire, Mali, and Niger have relatively poor reliability average scores (although none of these averages as much as 40 percent error margins in the items which make up this scale) and were judged to have large numbers of difficulties on the remaining three indices.

Additionally, Table 6 indicates that problems of one kind tend to co-occur with those of at least one other kind. High levels of data "doctoring" or administrative weaknesses do not occur independently but usually in conjunction with each other and with relatively low reliability and availability. It is only high non-availability of data--as in the case of the Gambia, Rwanda, and Uganda--which sometimes is reported independent of other problems.

The Correlates of Data Reliability

To this point we have considered levels of data quality both in the aggregate and for individual nations. It is also of interest to ask what characteristics of individual polities are associated with data quality variations within our sample. Knowledge of any conditions which are associated with reliability levels would be useful both to individuals choosing single nations for intensive study and to those executing large sample aggregate studies. Because of the limited systematic interest of most political scientists in data quality issues, however, there are few generally accepted hypotheses concerning the sources of data reliability. Nonetheless, we collected data on several background variables for our sample which one might presume to be related to data quality.

One expectation was that the character of the political regime might influence the quality of reported data. We categorized the nation sample, therefore, in terms of a regime typology adapted from Crawford Young.¹ Likewise, hypothesizing that the age of the political system or the extensiveness of national political institutions might influence reliability, we gathered data on the number of years since independence and the size of the central government public sector relative to the gross domestic product.² Also, in recognition of the frequent observation by Third World area specialists that the character of the colonial regime--which it is also argued varied systematically with the identity of the colonial power--was often a significant determinant of institutional performance in the post-independence period, we categorized nations in terms of their former ruling colonial power.³ Finally, hypothesizing that societal "complexity" might complicate the administrative tasks associated with data collection, we added measures for population size, population density, geographic size, and ethnolinguistic diversity.⁴

¹Political Systems Development," pp. 452-472 in James N. Paden and Edward W. Soja, eds., The African Experience (Evanston: Northwestern University Press, 1970). The codings were taken from an independent, companion survey wherein we asked a large sample of African area experts to code nations on a variety of political scales. The results of that second survey will be reported in a later manuscript.

²Data on years since independence were derived from Morrison, et al., op. cit. Public expenditures as a proportion of gross domestic product were computed from World Tables, 1976 (Baltimore: Johns Hopkins Press, 1976).

³The codings were taken from Morrison, et al., op. cit.

⁴Taken from Morrison, et al., op. cit. and from the United Nations Statistical Yearbook, 1976 (New York: United Nations Department of Economic and Social Affairs, 1977).

Investigation of the rank order correlations of these background variables with the four indices of data quality reported in Table 6 was generally disappointing. While most of the associations were in the direction indicated by our implicit hypotheses, only occasionally did one or another rise to our criterion five percent level of statistical significance. Thus, most of these nation characteristics do not assist us in accounting for data quality variations within this nation set.

TABLE 7: Average Data Quality Scores by Colonial Background^a

	<u>Average Error margin</u>	<u>Data Non- availability</u>	<u>Admin. weaknesses</u>	<u>"Doctoring" of data</u>
Former British colonies (n=11)	1.7	2.1	2.4	0.6
Former French colonies (n=7)	2.2	4.4	5.7	3.7
Total Sample (n-23)	1.9	3.1	3.3	1.6

^aIndex construction is discussed in the text and in the notes to Table 6.

The one exception to the above generalization was for the identity of the former colonial power. On that one trait we found substantively significant variation in the estimates of contemporary data quality. Our nation sample only contains sufficient numbers of former British and French colonies for comparison purposes; yet, the results for those two groups, as reported in Table 7, are rather striking. When compared to the averages for the entire sample, the former British colonies show consistently high data quality and the former French consistently low quality. A possible

explanation for this difference may be the character of the French colonial administration in West and Equatorial Africa. Administrative units here, which later became countries, did not develop the same autonomous data gathering facilities as did the British colonies.

Conclusions

This paper has explored the quality of African social, economic, and political data by means of an "expert judges" approach. The adequacy of this particular method for estimating data quality is given support by the willingness of our sample of scholars to tackle this task, by their ability to make discriminating assessments of reliability for a fairly wide range of variables, and by their frequently expressed continued interest in this project. Certainly, our results do not provide unequivocal confirmation of the power and adequacy of this approach. They do suggest, nonetheless, that this can be a fruitful means for providing systematic information for comparative analysis. At a later point we will comment on possible improvements which might be incorporated into any similar future studies.

How one evaluates our specific results for African data depends upon his original expectations regarding such data and his research preferences for employing such information. In almost every area, the modal response category was either 10 or 20 percent error, and the bulk of all responses fell into these two categories taken together. Such error levels do not approach what is implied to exist by many critics who have condemned Third World aggregate data as near worthless. Depending on one's purposes, of course, this much error could be of great concern. Scholars executing single nation studies should be most concerned about such error margins. In such research, the reliability of single data points is obviously more critical.

For many aggregate cross-national analyses, however, this may not be a very damaging average error level. When analyzed by means of some of the standardly employed social science statistical techniques, error margins in this range for some of one's cases may not have an exceptionally detrimental impact on aggregate analyses. Rank order correlation methods would be resistant to much effect as long as one's sample of nations was not especially homogenous on the variables under analysis (with low variance, even moderate error levels could alter the rank orders). Likewise, Pearsonian correlation and linear regression methods can tolerate moderate error levels with little significant effect on the results.¹

Of more concern than the overall characterization of African data quality is probably the character of some specific kinds of data and for some specific nations. Our study indicated that data on population and on agricultural and general development policy were, on average, more subject to error and manipulation than that of other areas. Likewise, some nations were indicated to have especially weak data overall. Information drawn from these areas or nations should, then, be treated particularly cautiously. The possibility that these limited problem areas could have a broad impact is itself heightened when one recalls the extensive use in cross-national analyses of population estimates to create various per-capita measures. Errors in the denominator could, then, seriously bias measures for a wide range of other substantive topic.

Our examination of nation specific reliability levels and of the correlates of data quality levels should help to extend scholarly concern for the amount and sources of error in cross-national data. We hope that

¹ See Bruce M. Russett, "A Note on the Evaluation of Error and Transformation in Data Analysis," American Political Science Review (June, 1965), pp. 444-446.

area experts might be intrigued enough by the nation-specific results to explore these matters in more detail. Similarly, our brief examination of the correlates of data quality variations was intended to pique the interest of aggregate cross-national analysts. Systematic comparative study of both the levels and sources of data quality could provide continuing dividends by means of its impact on subsequent research.

In addition to the preceding remarks, some comments are in order concerning possible future uses of this particular methodological strategy for estimating data quality. We recognize fully the limitations of our own work imposed, for example, by the use of a single scholar per nation. An important issue which could be resolved by subsequent analysis would be the likelihood of disagreement on these matters by samples of experts for individual nations. We should point out, however, that the experts would have to disagree quite radically in order to shift significantly the general characterization of African data equality supported by this paper. That is, since we have found consistent error margin estimates of 10-20 percent, alternative "experts" would have to rate similar data as of especially low reliability to shift this overall average very far. If the debate among the experts was between estimates of 10 or of 20 percent error, we might conclude that for many purposes the ultimate resolution of the issue would not be of great consequence.

There are, as well, some additional improvements which could be incorporated into any future uses of this "expert judges" approach. Perhaps, of most benefit would be the collaboration of a specialist in the study of the region for which experts are to be sampled. Both sample selection and the selection of items for the questionnaire would be facilitated and would

likely be focused more closely on some specific matters of interest in the particular region. Additional refinements in future studies might include questions about specific data estimates, perhaps, including the actual figure in the questionnaire itself. In similar fashion, country experts could also be queried regarding the accuracy of data for several different timepoints, such as population figures at 1960, 1970, and so on. If judges were asked to estimate both the extent and direction of errors, one could discern whether the overall error for a sample of nations would be random or systematic.

In conclusion, we see the preceding study as having two separate but equally useful products. On the one hand, whatever their possible limitations, our findings offer the only systematic evidence of which we are aware regarding the accuracy of a specific set of cross-national variables on a nation-by-nation basis. On the other hand, it is our hope, at least, that this work will heighten the inquisitiveness of comparativists regarding data quality. Regardless of whatever approach one takes for addressing this issue--and we hope several different ones will be explored in the future--the potential gains for subsequent research would seem to be exceptional.

PROGRAM OF DEVELOPMENT STUDIES
Discussion Paper Series

65. "Public Expenditures and Services as Policy Outcome Predictors: A Two-Nation Analysis for Health and Education" (1975), 39 pp.....Kim Q. Hill
66. "Low-Income Housing Development in Bogota"(1975), 32 pp.....Stephen Bender
67. "Consumption and Earnings Patterns and Income Redistribution" (1975), 26 pp..... J. Gregory Ballentine and Ronald Soligo
68. "Income Redistribution and Its Effects on Factor and Import Demand in Taiwan: A Simulation Approach" (1975), 33 pp..... Yhi-Min Ho
69. "The International Tin Agreement: A Reassessment" (1975), 19 pp. Gordon W. Smith and George R. Schink
70. "Ethnography of Migration: Breaking out of the Bipolar Myth" (1976), 15 pp..... Douglas Uzzell
71. "Education, Communalism and Income Distribution" (1976), 33 pp..... Fred R. von der Mehden
72. "From Play Lexicons to Disengagement Spheres in Peru's Urbanization" (1976) 21 pp..... Douglas Uzzell
73. "Savings Behavior of Poor and Rich in Taiwan: 1964, 1966, 1968, and 1970" (1976), 30 pp..... Marian Krzyzaniak
74. "Communalism and the Gains from Development: The Case of Nigeria" (1976), 55 pp..... Gaston V. Rimlinger
75. "The Relevance of the New View of the Incidence of the Property Tax in Developing Countries" (1976), 38 pp.....Charles E. McLure, Jr.
76. "Price Policy and Agricultural Development in Ecuador" (1976), 34 pp..... Wayne R. Thirsk
77. "The Incidence of Jamaican Taxes, 1971-72" (1976), 37 pp..... Charles E. McLure, Jr.
78. "The Production Structure of the Manufacturing Sector and Its Distribution Implications: The Case of Taiwan" (1976), 33 pp.....Yhi-Min Ho
79. "Income Distribution and the Composition of Final Demand: Some New Findings for Colombia" (1977), 41 pp..... Wayne R. Thirsk
80. "Commodity Instability and Market Failure: A Survey of Issues" (1977), 45 pp..... Gordon W. Smith
81. "An Analysis of the Saving Behavior of a Group of Colombian Artisan Entrepreneurs" (1977), 41 pp..... Donald L. Huddle
82. "U.S. Commodity Policy and the Tin Agreement" (1977), 31 pp..... Gordon W. Smith
83. "Data Reliability in Cross-National Research: A Test Employing Black Africa Country Experts" (1978), 28 pp..... Kim Q. Hill and F. R. von der Mehden

NOTE: Discussion Papers are available upon request to individual scholars and researchers and libraries of educational institutions.