

Zu

ZA – No. 0055 bis 0057

DATA BASES IN THE HUMANITIES AND SOCIAL SCIENCES

IFIP Working Conference on
Data Bases in the Humanities and Social Sciences
Dartmouth College, Hanover, NH, 23-24 August, 1979

Program Committee
Jerome M. Clubb, Bernard M. Fry,
John McCarthy, Judith S. Rowe

Proceedings of the IFIP Working Conference on
Data Bases in the Humanities and Social Sciences
Dartmouth College, Hanover, NH, 23-24 August, 1979

edited by

Joseph RABEN
*Queens College
City University of New York, U.S.A.*

and

Gregory MARKS
*Interuniversity Consortium for Political
and Social Research
Ann Arbor, MI., U.S.A.*



NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD



1980

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD

and are offered the options of (a) purchasing the newly released data tapes (including all corrections) or (b) receiving without charge revised documentation pages for all variables that have been changed as well as instructions for correcting the data tape(s) they already possess.

In the summer of 1978, a questionnaire was sent to users of the NLS data tapes (approximately 150) to measure reactions to the data and documentation. On the basis of a fifty percent response rate, overall user satisfaction with the structure of the data tapes, and especially with the documentation package, appears to be high. With respect to the fixed-length variable issue, over one-third of the respondents were strongly opposed to "compressing" the record, one-third were indifferent, and of the remaining one-third only about one-half had more than mild objections. Since even the largest of the four data files fits on three 2,400 foot reels (at 1600bpi density) and since the price is not related to number of reels, we expect to retain the current raw data format for both the in-house and the public tapes.

FOOTNOTES

¹ Under contracts with the Department of Labor and the Department of Defense, the Center has begun a new NLS study of a sample of approximately 13,000 young men and young women 14 to 21 years of age, with overrepresentation of blacks, Hispanics, economically disadvantaged whites, and youth in the armed services. The new sample was interviewed for the first time in early 1979 and will be interviewed annually thereafter for a period of five years. Responsibility for drawing the sample, conducting the field work, and preparing the data tapes has been subcontracted by the Center to the National Opinion Research Center.

² A more complete description of all aspects of the study may be found in the National Longitudinal Surveys Handbook, which is available free of charge from the Center for Human Resource Research, The Ohio State University, Columbus, Ohio, 43210.

³ 4 bytes = one word = 32 bits (IBM). A limited number of variables, by their very nature, require more than four digits (e.g., sampling weight, income, assets, etc.). In these few instances, two contiguous variable locations are used to contain the data.

GERMAN ELECTORAL DATA: 1961-1970

THE MANNHEIM VOTING STUDIES

Robert C. Danziger

Lehrstuhl für Politische Wissenschaft I
Universität Mannheim
Mannheim, West Germany

This paper describes a current data preparation project at the Universität Mannheim. A series of electoral studies undertaken from 1961 to 1970 are the target of this project. To this date, these studies constitute a relatively untouched data base. As the project develops, it will make the studies available for secondary analysis, providing as well a machine-readable Continuity Guide. When the data are ready, they will be stored at the Zentralarchiv in Cologne.

I. INTRODUCTION: THE PROJECT

Outside the United States, social science voting and public opinion research has taken hold in West Germany to a degree probably unmatched by any other western nation. Borrowing heavily from American experience, this research began slowly in the early and middle 1950's and blossomed fully in 1961 (Kaase (1978), Zentralarchiv (1975)). Since that time, in addition to work originating from a spate of private and semi-private polling institutes¹, a large body of university-based data has been collected. One of the first (and still) most active of these university centers was in Cologne. In the early 1960's, under the leadership of Rudolf Wildenmann and colleagues, some of the Cologne workers moved to Mannheim. In the period 1961 to the present, Mannheim has produced a large number of studies at the federal ("Bund"), state ("Länder"), and local ("Gemeinde") levels. Until recently, the bulk of these studies have been effectively unavailable for secondary analysis, as the data have been archived in their original multi-punch format, with extant documentation consisting often of little more than the questionnaire, various memoranda, and only occasionally a codeplan and univariate frequency counts. This paper briefly describes these data as well as a project in progress at Mannheim under the supervision of the author and Professor Wildenmann, and in cooperation with the Zentralarchiv für empirische Sozialforschung in Cologne and the Zentrum für Umfragen Methoden und Analysen (Z.U.M.A.) in Mannheim, to make the data accessible to the broader scientific research community.

The present focus of this project is some 39 representative surveys, made during the period 1961 to 1970, at the national, state, and local levels². All the samples consist of respondents selected by multi-stage, random procedures, and are predominantly oriented around electoral campaigns. The studies consist of single and multiple wave longitudinal samples as well as pre- and post-election research, series of items from larger, omnibus surveys, and several short-term panels. In addition, in a few cases involving state surveys, special subsamples were drawn from selected communities. The total cost of these surveys over the whole period amounted to an

estimated DM 2,000,000.

The following tables provide an overview of the data collected during these 10 years. Because of inadequate documentation, the exact sample sizes are sometimes uncertain.

--Table I--
The Mannheim Voting Studies: 1961-1970

ZA-Nr.	1. National studies	(ZA-Nr.)			
0055/0056/0057	July*, Sept., Nov.* 1961		DFG (a)	1679, 1633, 1961(b)	
0064	-Jan./Feb.** 1964		ZDF	1592 ← 1762	
0695/0696/0697	March, April, May, Aug. 1968		ZDF	1948, 1896, 1854, 1618	
0556/0557	Sept., Oct., Nov./Dec. 1965/0558		ZDF	1411, 1305, 1425	
0524	-Oct. 1966		ZDF	1870	
	-Nov./Dec./Jan. 1967-68		Bka	2004 ← 1988	
	-Jan./Feb. 1968		Bka	1964	
0699/0700/0701	-March, May/June, Aug., 1970		ZDF	1699, 1737, 1568	
0525	Aug./Sept. 1969		ZDF	1945	
2. State studies					
North-Rhine Westphalia--					
0062	-June/July 1962		ZDF	1890	
0561/0562	-April, June 1966		SkD	1200, 1564	
0702	-April/May** 1970		ZDF	2219	
Bavaria--					
0472	-April 1966		SkB	1207	
	-Jan./Feb., Oct./Nov. 1969		SkB	2016, 1234	
	-May/June 1970 (2 waves)		SkB	721, 1026	
Rhineland-Pfalz--					
0563	-Sept./Oct. 1966		SkM	1597	
Baden-Württemberg--					
0519	-Sept./Oct. 1967		SkS	1704 ← 1966	
0520	-March/April, Sept. 1968		SkS	2013, 2027	
0703	Lower Saxony--				
	-April/May** 1970		ZDF	1097	
3. Local studies					
Dortmund-Stadt; Arnsberg-Söst; Heilbronn; München-Land--					
0058/0059/0060	-July/Aug., Sept., Oct./Nov. 1961		DFG	1223, 1179, 1504	
Dortmund-Stadt; Arnsberg-Söst--					
0061	-June/July 1962		DFG	617	
Dinslaken-Rees; Essen--					
0559	-May 1965		ZDF	1250	
Dinslaken-Rees--					
0560	-Oct. 1965		ZDF	915	
Freiburg--					
2. 0520	-March 1968		SkS	?	
Mannheim--					
2. 0520	-March 1968		SkS	499	

*Notes (to Table I)--

(a) Sponsoring institute:

DFG-Deutsche Forschungsgemeinschaft (German Research Foundation)
ZDF-Zweites Deutsches Fernsehen (Second German Television Network)
Bka-Bundeskanzleramt (Federal Office of the Chancellor)
SkD-Staatskanzlei Dusseldorf (State Chancellery in Dusseldorf)
SkB-Staatskanzlei Bayern (State Chancellery of Bavaria (Munich))
SkM-Staatskanzlei Mainz (State Chancellery in Mainz)
SkS-Staatskanzlei Stuttgart (State Chancellery in Stuttgart)

(b) Number of respondents

*, ** See text below, section III

As mentioned above, several of these single studies are parts of panels. The panel studies include the following data sets:

1. Dortmund-Stadt; Arnsberg-Söst; Heilbronn; München-Land--
July/Aug., Sept., Oct./Nov. 1961
--with continuation in:
Dortmund-Stadt; Arnsberg-Söst-- June/July 1962
2. Dinslaken-Rees-- May, Oct. 1965
3. Baden-Württemberg-- March/April, Sept. 1968
4. Bavaria-- Jan., Oct./Nov. 1969, May/June 1970 (2 waves)
5. National-- Nov./Dec./Jan. 1967-68, Aug./Sept. 1969

II. DESCRIPTION OF THE DATA

The content of these studies varies rather widely, but the main themes can be preliminarily summarized in the following manner³:

--Table II--
Summary of the Content of the Mannheim Voting Studies

1. Statistical, Demographic, Personal data--
Sex, age, occupation, education, religion, social "class", income, place of residence, marital/family status, type of residency, household size, place of business/work, group memberships, personal interests, family background
2. Government (national and state)--
Performance of, knowledge of, most important tasks facing, satisfaction with leadership of, differences between opposition party and, most capable party in, differing capabilities of national and state, importance of state, (general) satisfaction with, attitudes toward coalitions in, possibilities for citizen (personal) influence upon, importance of opposition in
3. Politicians (national and regional)--
Knowledge of, contact with, opinion about, behavior/characteristics of, appropriateness for ministerial (incl. Chancellor) responsibility of, should/should not be Chancellor, problem-solving (specific problems) capability of
4. Local candidates--
Knowledge of, nature of representative "role" of, opinion about, issues involving, performance/ability of, party affiliation of

5. Grand Coalition--

Opinion about, most influential party in, knowledge of, goals of, success of

6. Political interest (general)--

Extent of, sources of, extent of among friends and acquaintances, extent of in family, area of greatest

7. Political information--

Sources of in mass media (TV, radio, weekly magazines, newspapers), attention to, satisfaction with extent of in mass media, other sources of, satisfaction with own level of, (*various questions probing extent of factual information about procedures and institutions in West Germany--e.g., voting procedures, who holds offices, institutional structure)

8. Economic situation--

Judgment of present German, future development of German, present personal, over-time change in personal, prospect for future personal, groups disadvantaged by, satisfaction with living standard, opinion about living standards

9. Political issues--

General knowledge of, possibilities to control outcomes of (by general citizenry and interest groups), most important, (*series of questions about specific political problems in the areas of-- defense (armed forces, NATO), general foreign policy, European policy, policy toward East Bloc, recognition of East Germany, educational policy, Emergency Powers Law, co-determination in industry, Berlin Crisis, atomic weaponry, danger of war, etc.)

10. General political--

Opinion of West Germany as world power, what world role for Germany, satisfaction with democratic system in Germany, best type of leadership, best form of party system, why/when one should vote, opinion on statements of democratic "groundrules", opinion about various social and economic groups

11. Parties, party identification--

Main goals of, opinion about role and behavior in political system, inclination toward specific, identification with, "strength" of identification with, inclination of various groups and institutions toward, opinions about followers of other, opinion of growth of NPD strength, opinion on financing of, opinion about specific, membership in, parents identification with, over-time change in opinion of specific, party best able to handle (specific) problems

12. Voting--

Effectiveness and importance of, meaning and importance of state elections, factors influencing decision, importance of electoral outcome, intent to participate in, vote intention (i.e., which party), why change vote/split ticket, when decided upon vote intention, personal voting history, for which party would never vote, opinion of conduct of campaign, influence of campaign, personal participation in campaign, interest in campaign, prediction of outcome, for whom voted (first and second ballots ("Stimme")), satisfaction with electoral outcome, political changes due to outcome

III. CONTINUING DEVELOPMENT OF THE PROJECT

Up to the present time, only a few of the studies listed in Table I are in finished OSIRIS format, with fully machine-readable codebooks. Those studies listed in the table with a single asterisk

have been prepared by the Zentralarchiv in Cologne (in conjunction with the Inter-University Consortium for Political and Social Research at the University of Michigan and Z.U.M.A.), primarily as part of the German Electoral Data Project (Kaase (1978), Zentralarchiv (1978)). In addition, those studies marked by a double asterisk have been made ready by the author and will be available through the Zentralarchiv by the beginning of 1980. Though currently handicapped by a shortage of trained personnel, just as soon as any of the other studies have been cleaned and documented in Mannheim, they too will be stored in Cologne.

For the past year, the major efforts at Mannheim have been focused upon the development of a Continuity Guide for these data. A first version of this guide will be finished in the spring of 1980. Presently, it is ready in machine-readable form only in German. The preparation of this guide at this early stage was undertaken both to provide the project with an overview of the task and to give other researchers preliminary access to the data. This means that many of these data, though currently only in very raw form, can be obtained from the Zentralarchiv by those scholars wishing to use them immediately. The Continuity Guide should at least serve to make the difficult task of data selection and preparation that much easier. The guide itself is simply organized with keyword headings, the exact question texts (with commentary or response categories where deemed necessary) and, of course, a list of the studies in which the items appear. Moreover, the guide is exhaustive in that every uniquely worded question is included in it. Each item is additionally coded according to its substantive content (Zentralarchiv (1970)). Thus, the investigator can not only access the information through the use of keywords, but may do so also according to topic. In its machine-readable form, the guide can be manipulated by textual processing or analysis programs, so that it serves any other specific needs of individual researchers (e.g., the development of different or multiple keywords).

Finally, the archiving of the data in Cologne has several advantages. Most important is the fact that the extensive service capabilities of the Zentralarchiv will ensure that the data are readily available to the scientific community. Secondly, the current information retrieval system in operation in Cologne will facilitate research with these data (Mochmann (1974-75)). The ZAR system now allows instant access to all items in OSIRIS dictionary/codebook format, and will eventually be provided with OSIRIS and SPSS data management/analysis capabilities as well. One by one, as the Mannheim studies are cleaned and documented, they will be integrated into ZAR. As this happens, interested scholars will be notified through Zentralarchiv publications, bulletins of the European Consortium for Political Research, or other sources.

FOOTNOTES

1. The largest of these institutes, according to Kaase (1978) are: Infas-Bad Godesberg; Infratest-Munich; Institut für Demoskopie-Allensbach am Bodensee; Emnid-Bielefeld; Götas-Bremen; Marplan-Offenbach; Wickert-Tübingen. As well, the political parties, some government bureaus, and the major news media, while often using the above-mentioned (or other) agencies, sometimes support

their own polling organizations. Many of these agencies--esp. Allensbach--maintain their own archives, though these are not always fully accessible for research purposes.

2. Mannheim continues to produce voting and public opinion studies. Some of the post-1970 work is in finished format and is available through the Zentralarchiv in Cologne. Those studies, dating from 1971, which are not immediately accessible to researchers, will eventually be incorporated into the data project described here.
3. Not all questions in the list are posed in each of the surveys. Most of the items noted in Table II have been asked at least twice. This overview as well does not contain every question, but is rather designed to roughly sketch the most important themes. The Continuity Guide in preparation is complete in this regard.

REFERENCES

1. Kaase, M., Public Opinion Polling in the Federal Republic of Germany, in: Cerny, K. (ed.), Germany at the Polls: The Bundestag Election of 1976 (American Enterprise Institute, Washington, D.C., 1978).
2. Zentralarchiv für empirische Sozialforschung, Liste der archivierten Umfragen (Zentralarchiv für empirische Sozialforschung der Universität zu Köln, Bachemerstraße 40, 5000 Köln 4, West Germany, (1975(with supplement to 1977))).
3. Zentralarchiv für empirische Sozialforschung, Variables Over Time: Continuity Guide to the German Electoral Data Project (1978)
4. Zentralarchiv für empirische Sozialforschung, Klassifikations-schema zur Erfassung von Inhalt, Form, und Funktion von Umfragen der empirischen Sozialforschung, (1970(available in both English and German)).
5. Mochmann, E., Information Access at the Data Item Level: Approaches to Indicator Retrieval from Survey Archive Data Bases, SIGSOC (Special Interest Group on Social and Behavioral Science Computing) Bulletin 6, #2 and 3 (Fall/Winter 1974-75) 47-55.

Data Bases in the Humanities and the Social Sciences
J. Rabon and G. M. (eds.)
North-Holland Publishing Company, 1980
© IFIP

DATA RESOURCES AND SERVICES FROM THE CRIMINAL JUSTICE ARCHIVE AND INFORMATION NETWORK

Judith A. Marks
and
Michael W. Traugott

Inter-university Consortium for Political and Social Research
The University of Michigan
Ann Arbor, Michigan

The development of a data archive relevant to the study of crime and the criminal justice system reflects an awareness that greater advantage must be taken of investments which have been made in original data collections. Relatively immediate research returns and policy analyses are possible from the extended analysis of such resources, and such efforts can often be achieved without the lag and expense associated with launching new studies. With the increase in the number and size of computer-readable data files in the criminal justice area, the challenge is for archives to develop means for simplifying access to datasets on a routine basis and for training a new cohort of secondary analysts in their use.

The availability of computer-readable resources through a social science data archive serves both a scientific and an administrative function based upon the concept of secondary or extended analysis. The underlying notion of secondary analysis is that no principal investigator ever exhausts the full analytic potential of a dataset. As the process of theory development and testing occurs, and as additional disciplinary perspectives are brought to bear in the field, data collected for one purpose come to be viewed quite differently in conceptual and operational terms by other researchers.

From a scientific perspective, the significance of archival data rests upon the principle of replication, the notion that the original relationship between two or more variables should be reproducible by secondary analysts in the original data base. More importantly, data archives provide a central repository for multiple collections of data in which the same relationship may be tested in a variety of research contexts. Replication has even greater social significance where matters of public policy and planning affect the lives and well-being of individual citizens. The process which we call "policy evaluation" really represents successive replications of basic relationships in a time series analysis in which we are testing the hypothesis that desired goals have in fact been achieved as a result of a change in policy.

From an administrative point of view, archival resources represent a cost-effective means of increasing research opportunities and, implicitly, results. The greatest single cost involved in the conduct of research is usually for data collection and processing. In addition to financial resources required for research and support staff, field costs, and the like, this process often requires substantial periods of time, even years in the case of extensive longitudinal studies. The alternative, when appropriate archival resources are available, is construction of a research design based upon existing data, followed by a search through documentation for meaningful operationalizations of important theoretical concepts. One of the most important archival functions is to provide documentation of sufficient quality to enable a researcher who did not participate in the original data collection to understand the project designs and goals, field work, and analytic methodology. Although there is often an iterative process by which