

WHO, WHAT, WHEN, WHERE, AND WHY:
AN ANALYSIS OF USAGE OF THE GENERAL SOCIAL SURVEY, 1972-2000

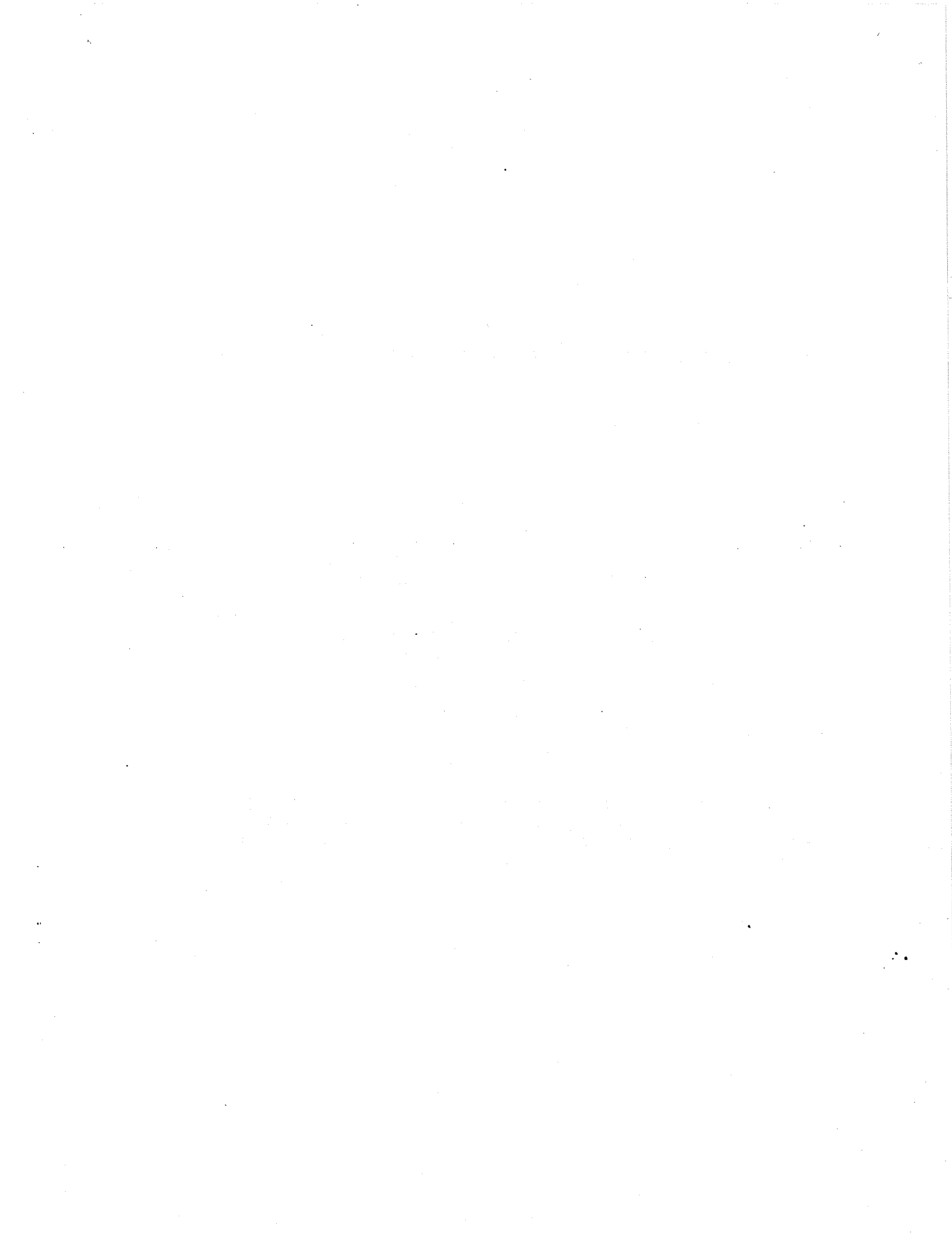
Tom W. Smith

GSS Project Report No. 22

National Opinion Research Center
University of Chicago

July, 2000

This research was done for the General Social Survey Project directed by James A. Davis, Tom W. Smith, and Peter V. Marsden. The project is supported by the National Science Foundation, Grant No. SBR-9617727.



The National Data Program for the Social Sciences is a data diffusion project and program of social indicators research. Periodically since 1972 the National Opinion Research Center (NORC), supported by grants from the National Science Foundation, has conducted the General Social Survey (GSS). The GSSs contain a wide range of demographics covering respondent and parental characteristics; behavioral items such as group membership and voting; personal, psychological evaluations of happiness, misanthropy, life satisfactions, and other aspects; and attitudinal questions on such public issues as abortion, crime and punishment, race relations, sex roles, and spending priorities. Many GSS questions are replicated according to a fixed rotation pattern. Other items appear on a single survey as part of our topical or International Social Survey Program (ISSP) modules or as methodological experiments.

The National Data Program for the Social Sciences has made the GSS data available at cost to all interested parties for use in research, teaching, or other applications. Each survey is publicly available within a few months of collection and is distributed by the Roper Center, University of Connecticut and the Inter-university Consortium for Political and Social Research, University of Michigan.

This report analyzes the usage patterns of the 1972-1998 GSSs by examining the papers that have employed the GSSs in their research. The Annotated Bibliography of Papers Using the General Social Surveys (2000) lists 5,431 papers using the GSS. While this listing falls far short of all actual uses of the surveys, it is believed that it represents a sufficiently large, varied, and important body of uses to provide an informative source for analysis.¹

In examining the use pattern of the GSSs we ask the standard journalistic litany of inquiries:

Who uses the GSSs?
What variables are used?
When have the uses appeared?
Where have they appeared?
Why have the GSSs been used?

¹GSS usages are difficult to locate in part because of poor citation practices. Joan E. Sieber ("Investigators' Concerns About Data Sharing," Paper presented at the Public Health Service Workshop, Bethesda, Maryland, April, 1990) found that in a sample of published articles known to use the GSS 12% did not in any way identify the GSS and 87% did not name the principal investigators. Moreover, in our own study in 2000 we found that only 57% of a sample of the articles known to use the GSS were detected as GSS uses in an extensive search using various computerized databases (e.g. Sociological Abstracts, Dissertation Abstracts, Social Science Citation Index, ERIC, Medline).

Who

The 5,431 papers in the bibliography have a total of 8,436 authorships: 3,306 with single authors, 1,490 with two authors, 441 with three authors, 153 with four authors, 28 with five authors, 9 with six authors, and 3 with seven authors. Most authors (79.1 percent) are academics with college affiliations. They are followed by associates of research centers, foundations, or associations (10.4 percent), government employees (1.6 percent), journalists and editors (1.1 percent), others (1.4 percent), and unknown (6.3 percent). Among the academics, Sociology predominated (57.6 percent), followed by Political Science (13.3 percent), Medicine/Health (5.2 percent), Business Administration/Management (4.4 percent), Psychology (4.2 percent), Criminal Justice and Law (3.1 percent), Economics (2.3 percent), Communications (2.0 percent), religion (1.7 percent), Other Social Sciences, including anthropology, geography, education, etc. (3.5 percent), Non-social Sciences, including statistics, biology, and engineering (1.9%), and University-based, research centers rather than academic departments (0.9%). Within Sociology almost every specialty is covered including gerontology, public health, demography, religious studies, statistical and survey methods, rural and urban studies, cross national comparisons, social deviance, social change, social control, stratification, mobility, etc.

There has been some diffusion of usage from its center in Sociology over time. Sociologists at universities accounted for 55% of the lead authors in 1972-1979, 51% in 1980-89, and 43% in 1990+. Within academia Political Science, Medicine, Criminal Justice/Law, Economics, and Non-social Sciences all showed monotonic gains and Communications and Business/Management lost share. Religion, Psychology, and Other Social Sciences showed no meaningful change. Use outside of academia also showed little directional change, moving from 17% in the 1970s to 18% in the 1980s, and 14% in the 1990s.

What

Both in terms of variables and survey years use has been quite varied and diverse. The counting of variables in surveys is a somewhat arbitrary matter depending on how filters are handled, whether several questions are used for determining one variable or one question coded to form several variables, and other matters. The cumulative GSS (1972-1998) contains 3,278 variables. Usage has been highest for replicating items that have been asked over a long period. Among these demographics have been the most heavily used (Table 1). The list is headed by education with the other major stratification variables, occupation and income, coming in fifth and tied for sixth. Coming right behind education are the three basic biological attributes: age, sex, race. The top cultural demographic --religion--ranks tied for sixth, followed by marital status at eighth, and the geographic variables--region and size of place--at ninth and tenth. The prevalence of demographics, of course, reflects their standard use as control variables and the first-place finish of education indicates that stratification and SES are still the central concerns of

Sociology. The other three types of items--personal evaluations, attitudes, and behaviors--all have been used less frequently. Personal evaluations are led off by the measures of personal happiness and job satisfaction, followed by other happiness and satisfaction items and self-rating of health. Attitude items are headed by the seven-item abortion scale and the top 16 groups of items also include five on race relations, two on crime and punishment, the 13-item confidence scale, the 15 Stouffer civil liberty items, the 9-item anomia scale, the spending priority items, three items on sexual morality, the four women's rights items, three items on misanthropy, and an item on legalization of marijuana. Behavioral and related items are topped by a wide edge by church attendance. Then come group memberships, TV viewing, gun ownership, political participation, socializing, and victimization.

Interest has been widely spread among a variety of attitudes, personal evaluations, and behavioral items and that attention has not been overly concentrated in any particular substantive area. Reflecting the smorgasbord content of the GSS, the use of variables has been wide-ranging and varied.

Variables that have appeared in the topical and cross-national modules have also attracted considerable use (especially considering that they usually appear only once). For example, the 1985 social networks module has been used 58 times and the 1985 and 1990 crossnational modules on the role of government have over 100 uses.

Only four types of items have had little use. First, various experimental variations (e.g. TRUSTY) which have appeared on a single form in a single year have usually been used less than 10 times. Second, various substantive items that have appeared in only a single year and which were not part of a supplement or module have been little used (e.g. PROWAR-2, SPDUE-1). Third, a number of minor demographics have been used only occasionally (e.g. the DOT characteristics of spouse's job). Finally, variables added in the last few years have little documented use so far (e.g. VOTE96-1).²

On average research has used 6.1 surveys. The two main factors restricting the number of years used are a) the number of surveys that were available at the time of the research and b) in how many surveys items of interest had appeared. Controlling for the number of surveys conducted at the time of the research, we find that on average researchers have used 42% of existing years. Among the available

²There has not yet been sufficient time for variables from very recent years to be widely used by the social science community. In addition, locating and abstracting the usages creates a lag. Based on abstracted usages the 1972 survey was used 1,447 times, 1973 - 1,875, 1974 - 1,970, 1975 - 1,884, 1976 - 1,862, 1977 - 1,940, 1978 - 1,722, 1980 - 1,604, 1982 - 1,421, 1983 - 1,292, 1984 - 1,349, 1985 - 1,336, 1986 - 1,160, 1987 - 1,160, 1988 - 1,186, 1989 - 989, 1990 - 885, 1991 - 707, 1993 - 373, 1994 - 284, and 1996 - 107. The more recent the survey the less opportunity there has been to use it and to detect and include it.

years, the number used depends on how many contained particular items of interest. Until 1988 most items appeared on two of every three surveys, putting a ceiling of .67 for most analyses and since 1985 the topical module have been mostly fielded only on a single survey and ISSP modules have been fielded once every 5-7 years, thereby severely restricting the number of usable surveys. At the lower end, 1.5% of uses have not employed any specific survey years. These consist of published replies or responses to articles using the GSS, descriptions of using the GSS for teaching, discussions of such matters as response rates, confidentiality, question wording, and survey methods, and unknown uses. At the upper end 22% have used 80+% of the available surveys. The increase in share of content devoted to ISSP and topical module has led to a slight decrease in the % of available surveys that have been used, from 46% in the 1970s and in the 1980s to 38% in the 1990s.

When

The GSSs have gained usages with time. There were only two uses in 1972, six in 1973, 24 in 1974, 51 in 1975, 80 in 1976, 82 in 1977, 155 in 1978, 161 in 1979, 191 in 1980, 195 in 1981, 179 in 1982, 127 in 1983, 160 in 1984, 188 in 1985, 153 in 1986, 166 in 1987, 206 in 1988, 270 in 1989, 306 in 1990, 289 in 1991, 272 in 1992, 384 in 1993, 361 in 1994, 382 for 1995, 275 for 1996, 157 in 1997, 239 in 1998, 264 in 1999, 85 in 2000 or forthcoming, and 19 with dates unknown. The increase in usage is even greater than these raw figures indicate. Recent usages are harder to locate. Based on past experience we estimate that we will eventually locate 470 for 1995, 330 for 1996, 239 for 1997, 364 for 1998, and 512 for 1999 (Table 2). The growth of uses reflects several factors: the natural lag time between data collection, analysis, and publication, the addition of more studies with increased opportunities for trend analysis or pooling, the addition of more questions (and especially the ISSP), and the spreading familiarity of the GSS throughout the social science community and beyond.

While GSS usage has expanded over time, it has also shown peaks and troughs. We believe that several of these short-term fluctuations can be explained by significant changes in the GSS. First, the surge in 1978 that peaked in 1981 comes from the creation of the first cumulative file in 1977. Second, the trough in 1983 results from the absence of GSS surveys in 1979 and 1981. Third, the surge starting in 1988 comes from the increase in cross-national analyses resulting from the start of the ISSP in 1985.³

³ISSP accounts for 5 usages 1986, 8 in 1987, 10 in 1988, 33 in 1989, 55 in 1990, 45 in 1991, 46 in 1992, 69 in 1993, 64 in 1994, 84 in 1995, 95 in 1996, 79 in 1997, 99 in 1998, and 133 in 1999. Since international usages are both harder to locate and take longer to unearth, these numbers are undoubtedly underestimates.

Where

Papers using the GSSs have appeared in a wide variety of places. Journal articles lead with 2,676 occurrences followed by books with 1,201, presentations at scholarly meetings with 959, reports with 442, and student dissertations and theses with 153. Distribution has been quite stable by type over time. The only two notable changes are the increase in reports from 4.5% of uses in the 1970s to 9% in the 1990s and the apparent decline in student dissertations/theses from 6% in the 1970s to 2% in the 1990s.⁴

As Table 3 details, most have appeared at conferences or in journals in the field of sociology. Virtually every major sociological and survey research journal has had articles using the GSSs with the Public Opinion Quarterly leading with 163 and followed by the American Sociological Review (113), Social Forces (94), and the Journal of Marriage and the Family (82). Outside of sociology, journals using the GSS include such interdisciplinary journals as the Annals and Daedalus, journals in other fields such as business administration (e.g., Personnel Psychology), communications (e.g., Journal of Communications and Journal of Broadcasting), political science (e.g., American Political Science Review), psychology (e.g., American Psychologist), and education (e.g., Journal of Research and Development in Education), journals of a general scientific nature (e.g., Science and Scientific American), religious journals (e.g., The Tablet), medical/public health journals (e.g., The New England Journal of Medicine and MMWR), economic journals (e.g. American Economic Review and Economic Inquiry), and journals of a general political or policy nature (e.g., American Enterprise, The Nation, and New Republic). In general, use has centered in sociology with a secondary concentration in the other social sciences and peripheral usages in a wide range of fields.

As with affiliations, usage has diffused from the core sociological journals to those in other disciplines. Sociology journals accounted for 49% in the 1970s, 49% in the 1980s, and 40% in the 1990s. Also, having a decreased share are Journalism and Communications (2%, 2%, 1%), Business and Management (5%, 1%, 1%), Education (1%, 1%, less than 0.5%), Public Interest (5%, 2%, 2%), and Other Social Sciences (4%, 3%, 2%). Showing gains were Religion (5%, 7%, 7%), Political Science (4%, 5.5%, 7%), Interdisciplinary (1%, 3%, 3%), Economics (0%, 1%, 2%), Misc. (5%, 5%, 10%), and Medicine/Public Health (2%, 2%, 6%). Showing a less clear trends were Survey Research (13%, 15%, 13.5%) and Psychology (5%, 4%, 6%).

⁴ We believe this decline is artifactual rather than real. Until the mid-1980s Dissertation Abstracts were manually searched and abstracts scanned for possible uses. Later when Dissertation Abstracts became available on-line, the search was carried out by computer. We believe that because key terms like GSS and NORC are rarely used in these abstracts that the automated search procedures are much less effective than the previous manual approach.

Similarly, among conference papers those in Sociology declined from 58% in the 1970s, to 46% in the 1980s, and 45% in the 1990s. Miscellaneous social science conferences also dropped from 10% in the 1970s and 1980s to 2% in the 1990s. The biggest gainer was survey research conferences from 7% in the 1970s and the 1980s to 24% in the 1990s. This was primarily due to ISSP-related conferences. Small gains occurred for Religion, Medicine/Health, and Government. Others showed small changes and/or no clear direction in their trends.

Why

The GSSs have been used for a number of purposes. Most commonly the GSSs have been used to carry out scholarly research in the social sciences on contemporary American society. This research consists of proposing and testing various hypotheses about the organization, operation, and function of various social phenomena, inter-generational studies of mobility and transmittances, trend analyses of social change, cross-national comparisons, and related matters.

In addition, the GSSs have been used in social indicators reporting. Data from the GSS were used extensively in Social Indicators III and in publications by the National Center for Educational Statistics, the National Center for Health Statistics, the Bureau of Justice Statistics, the U.S. Bureau of the Census, and the National Science Foundation.

The GSSs have also been used to study various questions of survey methodology by using split ballot experiments on wordings and context, various item indexes to examine scaling, and GSS data to illustrate the application of new statistical methods.

Finally, the GSSs have been used in a wide variety of other ways. These include uses as guides of item wording, questionnaire construction, and codebook organization; as a teaching tool for sociology and related fields; as supporting evidence in judicial decisions; and as a source for discussions of politics and policy.

Conclusion

This analysis of the GSS usages demonstrates that the GSS is widely used by a large number of scholars in sociology and the other social sciences, by the government, and by researchers in many other areas. The GSSs have been used in an extensive and varied number of places including almost 550 different journals by an equally varied and even larger number of authors. While Sociology remains the core user community, usage has diffused to other social sciences and beyond. The set of scholarly hypotheses tested and trends analyzed has likewise been extremely diverse. With the exception of a few topics and groups where extensive usage has not developed, the GSSs have shown considerable success in providing useful data to the social science community. In addition, the collection and distribution of the cross national and topical modules has opened up many new research opportunities. The prognosis for the future, based on past trends and recent developments, is for even more extensive and diverse usage in the coming years.

