THE SAGE Encyclopedia of the
SOCIOLOGY OF RELIGION
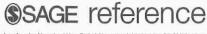
The SAGE Encyclopedia of the
SOCIOLOGY OF RELIGION

2

Editors

Adam Possamai
Western Sydney University

Anthony J. Blasi
Tennessee State University (Retired)

$SAGE

STATA is mostly used by advanced researchers who are interested in time efficiency and analysing a large data set. Although the software is mostly used for advanced and complex quantitative procedures, it can also perform simple data analyses just like SPSS. STATA is a more interactive (easily generates graphs of results) and time-efficient data analytical software that can run on multiple platforms with the right command.

Four analysis commands can be used in STATA: (1) command window, (2) result window, (3) review window, and (4) variable window. The analytical command window is used to record what command should be performed via the review window, while the available variables in the data are listed in the variables window based on the recorded command. The results of these analytical commands are then presented in the results window. STATA may seem a bit daunting for most researchers because of its complex organisation of command syntax.

SAS is another quantitative research software used by most social scientists to perform sophisticated and customised analysis for assessing religious patterns through forecasting, graphics, quality improvement, and planning. SAS is, by far, a more advanced software for performing statistical analysis because of its sophistication and effectiveness in performing advanced statistics using large data sets. Compared with other software packages, SAS is most effective for estimating strata, group effects, and weights and largely runs by a programming syntax rather than a point-and-click command. Social scientists interested in using this software require some above-average programming skill.

Social scientists need to familiarise themselves with one or two quantitative research software programmes to help improve the sophistication of their scholarly work and strengthen research outputs emerging from the field. Such expertise could help social scientists' study religious phenomena more efficiently, assist in managing their data sets, and perform more accurate and advanced statistical analyses that would shed some light on complex religious issues.

Victor Counted

See also Psychology of Religion; Qualitative Data Analysis Software; Society for the Scientific Study of Religion

Further Readings

Bryman, A., & Cramer, D. (1995). *Quantitative data analysis for social scientists* (Rev. ed.). New York, NY: Taylor & Frances/Routledge.

Cramer, D. (2003). *Advanced quantitative data analysis*. New York, NY: McGraw-Hill International.

Knoke, D., Bohrnstedt, G. W., & Mee, A. P. (2002). *Statistics for social data analysis* (4th ed.). Itasca, IL: Peacock.

Suen, H. K., & Ary, D. (2014). *Analyzing quantitative behavioral observation data*. New York, NY: Psychology Press.

QUANTITATIVE RESEARCH

Compared with disciplines that rely heavily on either qualitative (e.g., cultural anthropology) or quantitative research methods (e.g., economics, psychology), sociology boasts a wide array of methodological approaches. This methodological openness and diversity applies to sociology of religion, too, indeed much more so than to fields like the sociology of stratification and mobility, where quantitative methods predominate, or cultural sociology, in which qualitative ones prevail. Although distinguishing quantitative from qualitative research is not as easy as it may seem, the former can be defined as research that generalizes statistics (e.g., averages, means, correlations) from a sample to the larger population that it represents, which is done by applying standardized measurement procedures to the units of analysis that compose the sample. This entry provides an overview of the procedures used in data collection, the forms of data analysis, and the types of variables and their respective roles in survey research and hypothesis testing.

Data Collection and Units of Analysis

Although in 'data collection' a representative sample of respondents from a national population constitutes the most typical example, many different types of samples and populations are imaginable. One may for instance also draw samples from subpopulations such as clergy, members of a religious congregation, or theology students; or even draw samples from nonhuman populations

such as church sermons, television programmes about religion, tombstones, or whatever. The units of analysis that constitute the sample (e.g., respondents) are subjected to a series of standardized procedures (e.g., interview questions) to measure the features deemed necessary to answer the researcher's question. 'Measurement' here involves not only the assignment of quantitative scores such as respondents' ages or scores on a scale but also the coding of the units of analysis into nominal categories (e.g., whether or not a respondent self-defines as 'religious', 'spiritual but religious', 'Protestant', or 'Catholic'; or whether or not a tombstone refers to religion, and if so, then how). This produces scores for each of the measured features ('variables') and for each of the units of analysis (e.g., respondents) that together constitute the 'data matrix' or 'data set'. The latter is then first manipulated so as to construct the new variables that are needed to answer the research question (e.g., by combining a series of (dis)agreements with statements into one single scale or by combining categories like 'Protestant' and 'Catholic' into a new 'Christian' category). Subsequently, these new variables are then analyzed to obtain the statistics the researcher is interested in (e.g., means, percentages, correlations).

Forms of Data Analysis and the Roles of Variables

In terms of data analysis, quantitative research generally aims for either (a) estimations of how widespread particular characteristics are in the population from which the units of analysis have been sampled ('univariate analysis') or (b) estimations of the degree to which two ('bivariate analysis') or more ('multivariate analysis') of these characteristics tend to occur together. Univariate analysis is central to what is popularly known as 'opinion polling'; for example, estimating the percentage of a population that does or believes something (e.g., believing in God, going to church at least once a month, conceiving of oneself as 'spiritual but not religious', conceiving of religion as an outdated phenomenon). Bivariate and multivariate analyses are instead aimed at estimating relationships between variables. They are as such crucial for theory testing; in other words, determining whether or not a theory's predictions

about interrelations between variables ('hypotheses') are borne out by actually existing relationships between variables in the real world. Secularization theory, for instance, predicts a decline in religion and hence that younger birth cohorts will be less religious than older ones (bivariate analysis). Grace Davie's theory of an increase in 'believing without belonging', to give another example, predicts that the relationship between religious belief and church attendance will be stronger for the older than for the younger birth cohorts (multivariate analysis).

The most widely used statistical methods in the multivariate analysis of survey data (e.g., multiple regression analysis) require the researcher to impose a causal order on the data; that is, to distinguish between a 'dependent' variable and a series of 'independent' ones that can potentially explain it. It is important to underscore that such distinctions between dependent and independent variables are typically based on nothing but causal assumptions on the part of the researcher, typically informed by the theory she wants to test. Because causality can as such typically not be 'proven' statistically, survey researchers tend to be cautious in making strong causal claims on the basis of their findings. Yet, there are special situations in which causality ceases to be a matter of speculation or assumption. The most obvious one is a panel survey design which measures the same variables for the same respondents at multiple moments in time. In some instances, the problem of causal order does however not even exist in 'one-shot' surveys. A first example is a situation like the one mentioned above about a statistical relationship between birth cohort and religiousness, as typically found for Western countries. For the fact that younger birth cohorts are less religious than older ones must logically indicate that birth cohort somehow affects religiousness rather than the other way around. Indeed, 'somehow', because this still leaves open the important question of how and why this influence comes about in the first place (see below). A second example would be a relationship between respondents' religiosity and the religiosity of their parents, for it is of course very unlikely that the causal path runs from offsprings' to parents' religiosity rather than the other way around.

Whereas the public at large identifies survey research more with practices of opinion polling (univariate analysis) than with theory testing (bivariate/multivariate analysis), it is important to underscore that survey research is in fact better equipped for the latter than for the former. This is because while percentages found are heavily dependent on exact question wording, relationships between variables are quite resistant to that. One can for instance without much difficulty produce a long series of statements (e.g., the so-called 'Likert items' of the 'agree strongly' to 'disagree strongly' type) that all tap into the same attitude, opinion or belief, with some of these items suggesting the latter's virtual non-existence and others its virtual omnipresence. 'Belief in the God within', that is, the belief that the sacred ('God') does not so much exist 'out there' but is rather immanently present in the deeper layers of the self (a belief associated nowadays with the category that self-identifies as 'not religious but spiritual') can for instance be measured with items like (1) 'I am absolutely sure that God is something within each person', (2) 'I think that God is something within each person', (3) 'I do not believe that God is something within each person', (4) 'God is something within each person', or (5) 'I believe that God is something within each person'. Responses to these five items will surely be strongly correlated, indicating that they all tap into the same belief, so that differences in question wording do not affect their usefulness for measuring the belief at stake.

Yet, it is quite clear that item 2 ('I think . . .') produces higher percentages of 'belief in the God within' than item 1 ('I am absolutely sure . . .'). Even though changes in question wording do not affect their usefulness for measuring the belief at stake, then, such changes can have major consequences for the frequencies found. Move from 'A is not a nice man' to 'A is a jerk' and watch the number of negative evaluations of 'A' decline. Because questions for the measurement of any attitude, opinion, or belief are moreover selected from a pool of potential and interchangeable questions that is virtually infinite, survey-based claims about how widespread an attitude, opinion, or belief 'really' is should not be taken very seriously. This is of course less of a problem for questions into institutional affiliations and participation (e.g., 'Are you a member of church X?'; 'If so, how often do you attend?') but precisely such questions have become less relevant now that religious meaning and discourse have progressively escaped their traditional institutional moorings.

While relationships between variables are quite resistant to differences in question wording, then, research aimed at hypothesis testing has its own types of problems and shortcomings, as renowned specialist in sociological survey research James A. Davis observes. All too often, he points out, studies produce merely relationships between variables without clear implications for the tenability of sociological theories. This is why he argues for the need to take theory testing to a higher level by also addressing how observed relationships between variables can be explained in the first place. Indeed, all sociological theories worth their salt make suggestions about (a) the types of social contexts that bring about particular relationships ('contextual variables'), (b) the types of persons that bring about the latter ('moderating variables'), and/or (c) the types of mechanisms at the individual level that do so ('mediating variables').

The explanatory role of contextual variables is typically studied by combining sets of survey data (e.g., from different countries) that do as such contain many individual respondents (typically tens of thousands), all of them situated in just one of a much smaller number of social contexts (typically a few dozen) defined at a supra-individual level (e.g., countries). Such data are nowadays typically analyzed by means of 'multi-level analysis', which allows for the simultaneous estimation of effects of independent variables at the two levels of analysis separately and the testing of hypotheses pertaining to the so-called cross-level interaction effects. These are hypotheses that predict a relationship between two variables at the individual level to be stronger or weaker due to a variable defined at the contextual level. Ernst Troeltsch's classical distinction between (Catholic-style, collectivist) church religion and (Protestant-style, individualist) sect religion, for instance, informs the hypothesis that the individual-level relationship between religious belief and church attendance will be stronger in countries with a predominantly Protestant religious heritage than in those with a largely Catholic one.

The role of 'moderating variables' can be studied by assessing whether a relationship between two variables differs between groups of respondents (this is hence not the type of cross-level interaction effect just discussed but rather an interaction at the individual level). An example would be the hypothesis that the relation between religious belief and church attendance is stronger for Protestant than for Catholic persons (so rather than in Protestant and Catholics contexts). The role of 'mediating variables', finally, can be studied by testing whether the initial relationship between two variables declines (or even completely disappears) if a particular independent variable is added. If such a decline occurs, the additional independent variable apparently operates as an explanatory mechanism that brings about the initial bivariate relationship. In other words, the effect of the initial independent variable then apparently runs 'indirectly' through the variable added in the second step. This logic can, for instance, be used to find out whether lower levels of Christian religiosity among younger birth cohorts stem from stronger confidence in science or from stronger moral individualism, as different renditions of secularization theory maintain. The art of survey research, to sum up Davis's argument, thus requires not only methodological and statistical skills but also a keen eye for theoretical puzzles in need of empirical resolution.

Dick Houtman

See also Culture; Ethnography; Lived Religion; Longitudinal Study; Qualitative Research

Further Readings

Davis, J. A. (1994). What's wrong with sociology? *Sociological Forum, 9*(2), 179–197.

Houtman, D., & Achterberg, P. (2016). Quantitative analysis in cultural sociology: Why it should be done, how it can be done. In D. Inglis & A.-M. Almila (Eds.), *Sage handbook of cultural sociology* (pp. 225–236). London, UK: Sage.

Houtman, D., Heelas, P., & Achterberg, P. (2012). Counting spirituality? Survey methodology after the spiritual turn. In L. Berzano & O. Riis (Eds.), *Annual review of the sociology of religion, volume 3: New methods in the sociology of religion* (pp. 25–44). Leiden, the Netherlands: Brill.

QUIMBANDA

See Umbanda