Appendix B:
Notes to the Instructor

DATA

The data for this module are available as IBM SPSS portable files. There is one file containing only the 2010 sample data (called gss_10.por) and another file containing the variables in all time periods (called gss_10_04_02_96_89_82_75.por). These files can be downloaded from the Social Sciences Teaching Resources Depository website located at http://www.ssric.org.

SPSS FOR WINDOWS

There is an introduction to the basics of IBM SPSS available for your use, *IBM SPSS for Windows Version 19: A Basic Tutorial* (Linda Fiddler, Laura Hecht, Edward Nelson, Elizabeth Ness Nelson, James Ross). This can be ordered from McGraw-Hill (ISBN 0-07-804183-3). It is a basic introduction to SPSS and can be used as a supplement for a class or as a tutorial to learn SPSS by oneself. The version 19 book is on the web at http://ssric.org/tr/onlinetextbooks.

CODEBOOK

The codebook is included in this module as Appendix A. A list of all the variables for student use is at the beginning of the codebook. A variable (YEAR) was created specifying the year of the survey. If you are using the file containing all time periods, remember that you must either select out the cases for one of the years or use YEAR as a control variable in a crosstabulation to be able to distinguish one year from the other year.

Several of the variables have been recoded to make them easier to use. AGE was recoded into AGE1 and AGE2 to reduce the number of categories. EDUC was also recoded into EDUC1 for the same reason.

There are two occupational prestige variables (PRESTIGE and PRESTG80). PRESTIGE is based on a code introduced in 1970 and used in the 1975 and 1982 surveys. PRESTG80 is based on a code introduced in 1980 and used in the 1989, 1996, 2002, 2004, and 2010 surveys. These variables were recoded to make it possible to combine them into a single variable (PRESTG1).

Total family income has six variables. INCOME was used in the 1975 survey. INCOME82 is based on a code introduced in 1982 and was

used in the 1982 survey. INCOME86 is based on a code introduced in 1986 and was used in the 1989 survey. INCOME91 is based on a code introduced in 1991 and was used in the 1996 survey. INCOME98 is based on a code introduced in 1998 and was used in the 2002 and 2004 surveys. INCOME06 is based on a code introduced in 2006 and was used in the 2010 survey. We didn't attempt to combine these variables into a single variable. Respondent's income also has six variables (RINCOME, RINCOM82, RINCOM86, RINCOM91, RINCOM98, RINCOM06) and follows the same pattern as total family income.

There is a socioeconomic status index (SEI) in the GSS that we recoded to make it easier to use (SEI1).

Race in the 1975, 1982, 1989, and 1996 surveys consisted of three categories (White, Black, Other). In 2002, 2004, and 2010 the GSS included three race variables (RACECEN1, RACECEN2, RACECEN3). These variables are the respondent's first three answers to the question "What race do you consider yourself to be?"

We also created two new variables to measure religiosity (RELIGOSA, RELIGOSB) and five new variables to measure tolerance (TOLATH, TOLCOM, TOLHOMO, TOLMIL, TOLRAC). See the codebook in Appendix A for a description of these new variables.

WEIGHTING

There is a weight variable (WEIGHT) in the data set that should be used. The population for the survey was adults (18+). If there was only one eligible person in the household that person was interviewed. But if there were two or more eligible people, then one of them was randomly selected. If there were two eligible people, then each person had a 1 out of 2 chance of being selected. If there were three eligible people, then each eligible person had a 1 out of 3 chance and so on. In other words, the more eligible people in the household, the smaller the chance of selection for any one of them. We can correct for this by weighting each case by the number of eligible people in the household. There is a variable called ADULTS which is the number of people 18 years of age or older in the household and this is, of course, also the number of eligible people in the household. The weighted number of cases is just the number of eligible adults multiplied by the number of cases. This means that each case with two eligible adults had a weight twice that of each case with one eligible adult, each case with three eligible adults has a weight three times that of each case with one eligible, and so on. The problem with this is that it also inflates the total number of cases. So we have to adjust the weights so the number of cases stays approximately constant.

In 2004 and subsequent years, the General Social Survey modified their sampling design and included two weight variables that they hadn't included in previous surveys. For 2004 and 2010, we used one of their weight variables in the data set. For a more complete description of the weight variables in 2004, see Appendix A in the

General Social Survey 1972-2004 Cumulative Codebook (James A. Davis, Tom W. Smith and Peter V. Marsden, available through the Roper Center for Public Opinion).

The weight variables are not turned on in the data sets you will download. It is recommended that you use the weights for both data sets. You will have to turn them on in SPSS by going to Data/Weight Cases.

CUSTOMIZING THE MODULE

The module can be integrated into your classes in many ways. One approach is to assign the introductory chapters to be read by the students and work through the simpler exercises in class. Ideally students should be given the opportunity to use the computer themselves and not just watch you do it. Later, exercises can be assigned as homework and then reviewed in class. Chapters 4 and 6 contain a culminating exercise in which students choose a problem, produce the appropriate tables and statistics, interpret the tables, and write a brief research report.

A second approach is for faculty to develop their own exercises to accompany the module. There is no reason that you should limit yourself to the exercises we developed. If you do develop your own exercises, be sure to test them before assigning them.

A third approach is for faculty to use the data set accompanying the module but not use the module itself. You would have to develop your own introductory material to make it relevant to your particular class.

The only statistics used in the module are percentages, chi square, Cramer's V, and Gamma. You could skip over all these statistics except percentages or you could introduce other statistical techniques.

Two other appendices are included for your possible use. Appendix C contains supplemental instructional materials. Appendix D contains a description of how to compute the measures of association used in this module--Cramer's V and Gamma.