ICPSR Inter-university Consortium for Political and Social Research

Agricultural and Demographic Records for Rural Households in the North, 1860: [Instructional Materials]

Handout (PDF)

James W. Oberly

ICPSR 3463

Agricultural and Demographic Records for Rural Households in the North, 1860: [Instructional Materials]

(ICPSR 3463)

Principal Investigator

James W. Oberly University of Wisconsin, Eau-Claire

First ICPSR Version October 2002

Inter-university Consortium for Political and Social Research P.O. Box 1248
Ann Arbor, Michigan 48106

BIBLIOGRAPHIC CITATION

Publications based on ICPSR data collections should acknowledge those sources by means of bibliographic citations. To ensure that such source attributions are captured for social science bibliographic utilities, citations must appear in footnotes or in the reference section of publications. The bibliographic citation for this data collection is:

Oberly, James W. AGRICULTURAL AND DEMOGRAPHIC RECORDS FOR RURAL HOUSEHOLDS IN THE NORTH, 1860: [INSTRUCTIONAL MATERIALS] [Computer file]. Prepared by James W. Oberly, University of Wisconsin-Eau Claire, Dept. of History. ICPSR version. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002.

REQUEST FOR INFORMATION ON USE OF ICPSR RESOURCES

To provide funding agencies with essential information about use of archival resources and to facilitate the exchange of information about ICPSR participants' research activities, users of ICPSR data are requested to send to ICPSR bibliographic citations for each completed manuscript or thesis abstract. Please indicate in a cover letter which data were used.

DATA DISCLAIMER

The original collector of the data, ICPSR, and the relevant funding agency bear no responsibility for uses of this collection or for interpretations or inferences based upon such uses.

DATA COLLECTION DESCRIPTION

James W. Oberly

Agricultural and Demographic Records for Rural Households in the North, 1860: [Instructional Materials] (ICPSR 3463)

SUMMARY: These instructional materials were prepared for use with AGRICULTURAL AND DEMOGRAPHIC RECORDS FOR HOUSEHOLDS IN THE NORTH, 1860 (ICPSR 7420), compiled by Fred Bateman and James D. Foust. The data file (an SPSS portable file) and accompanying documentation are provided to assist educators in instructing students about the history of agriculture and rural life in the North, just prior to the Civil War. An instructor's handout has also been included. This handout contains the following sections, among others: (1) General quals for student analysis of quantitative datasets, (2) Specific qoals in studying this dataset, (3) Suggested appropriate courses for use of the dataset, (4) Tips for using the dataset, and (5) Related secondary source readings. Demographic, occupational, and economic information for over 21,000 rural households in the northern United States in 1860 are presented in the dataset. The data were obtained from the manuscript agricultural and population schedules of the 1860 United States Census and are provided for all households in a single township from each of the 102 randomlyselected counties in 16 northern states. Variables in the dataset include farm values, livestock, and crop production figures for the households that owned or operated farms (over half the households sampled), as well as value of real and personal estate, color, sex, age, literacy, school attendance, occupation, place of birth, and parents' nationality of all individuals residing in the sampled townships.

UNIVERSE: 956 counties in 20 northern states.

SAMPLING: The data (AGRICULTURAL AND DEMOGRAPHIC RECORDS FOR RURAL HOUSEHOLDS IN THE NORTH, 1860 [ICPSR 7420]) were collected using a two-stage variable-cluster sample design. A random sample of 102 counties in 16 northern states was drawn. From each of these counties one non-urban township was randomly selected. Census records from all households in the selected townships were included in the study.

NOTE: (1) This collection is part of an ICPSR initiative to promote the use of social science data in undergraduate education through the sharing of faculty-submitted ICPSR data-based instructional materials that have been developed for use in the classroom. (2) These materials were designed for use with the dataset AGRICULTURAL AND DEMOGRAPHIC RECORDS FOR RURAL HOUSEHOLDS IN THE NORTH, 1860

(ICPSR 7420). (3) The data are provided as an SPSS portable file, which was created from the ASCII data and SPSS data definition statements files available from ICPSR 7420. (4) The codebook and handout are provided by ICPSR as Portable Document Format (PDF) files. The PDF file format was developed by Adobe Systems Incorporated and can be accessed using PDF reader software, such as the Adobe Acrobat Reader. Information on how to obtain a copy of the Acrobat Reader is provided on the ICPSR Web site.

RESTRICTIONS: At the request of the principal investigator, publications that draw upon the data, AGRICULTURAL AND DEMOGRAPHIC RECORDS FOR RURAL HOUSEHOLDS IN THE NORTH, 1860 (ICPSR 7420), must observe the copyright held by Professors Bateman and Foust.

EXTENT OF COLLECTION: 1 data file + machine-readable documentation (MSWord 9.0 and PDF)

EXTENT OF PROCESSING: REFORM.DOC

DATA FORMAT: SPSS portable file

Cases: 21,118
Variables: 390

ICPSR-SIMI Submission, May, 2002

Studying the Rural North in 1860

by James Oberly, Department of History, University of Wisconsin-Eau Claire

General Goals for Student Analysis of Quantitative Datasets: Student use of quantitative datasets in history accomplishes several pedagogical goals. First, the datasets are compilations of primary sources, taken from the manuscript census and other listings. Studying primary sources, even once removed as datasets, is important as a way to introduce students to primary sources in history. Moreover, the very nature of computer analysis requires students to ask questions of the dataset in order to generate any output, and then to analyze the results and make sense of what they find. In addition, asking students to generate findings from a dataset helps them improve their ability to make sense of numerical information. Combining numerical analysis with a written report on their findings makes this type of assignment complete, at least as far as the development of reading, analysis, and writing. These are important goals that are part of any liberal education. An additional outcome of using datasets for undergraduate instruction is that students are included in the community of scholars who use the same materials and methods to formulate original contributions to knowledge. Senior majors in history and social sciences such as economics, political science, and sociology can examine the datasets produced by the leading scholars in the field. At a minimum, this teaching method helps the students to understand better the findings summarized in the textbook, monograph, or journal article they are assigned. Quite often, students are able to pose new questions of their own and test them with the ICPSR databases.

Specific Goals in Studying the Rural North in 1860: This dataset gives students a snapshot look at the rural, especially farm population of the northern states just before the Civil War. This was a critical moment in the history of northern agriculture as four movements were simultaneously having an enormous effect on the countryside: first, large numbers of European immigrants began moving to the North in the 1840s and 1850s; second, there was an accelerated westward population movement among rural people, from Northeast to Midwest; and third, there was the beginning of the mechanization of agricultural production on northern farms, which allowed farmers to increase their output by extending the amount of acreage planted. Fourth and last, the rural North was just beginning its demographic transition to a lower fertility level by 1860. Farm families, for their own reasons in some areas, were deciding to have fewer children. All of these phenomena are available for study in the dataset. The states

included are Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Ohio, Wisconsin from the Midwest, and Connecticut, Maryland, New Hampshire, New Jersey, New York, Pennsylvania, Vermont from the Northeast. The basic unit of analysis is the household as summarized in the manuscript census of 1860. Almost 12,000 farms households are included, along with more than 9,000 non-farm, but rural households. Students can examine the entire rural North, or just one state, one county, or even one township as a subset of the whole.

Data Source: ICPSR Study No. 7420, "Agricultural and Demographic Records for Rural Households in the North, 1860." by Principal Investigators Fred Bateman and James D. Foust, dataset created in 1984.

Suggested Appropriate Courses for Use: In the History Curriculum--Nineteenth Century U.S. History; Civil War and Reconstruction; U.S. Economic History; U.S. Rural and Agricultural History. In the Economics Curriculum--Economics of Development; U.S. Economic History.

The Files, available on the ICPSR-SIMI Website:

- 1) Instructor's Handout (this document). File size=44k, as an MS-Word file.
- 2) Codebook, which lists the variables in the dataset and their values. This is important for the instructor and the students to have together to decide on the questions to ask of the dataset, and to make sense of the output. Please note that the Principal Investigators who compiled the dataset from the primary sources also devised the codebook. Professors Bateman and Foust included simple frequency counts within their codebook for each value of their set of variables. Thus, but looking only at the codebook, a student can learn that the sample of households includes 155 taken from Mt. Pleasant Township, in Scotland County, Missouri in 1860. Or, that 13,205 heads of household listed "farmer" as their occupation, but only three "students" did so. File size=113k, as an MS-Word file.
- 3) *Dataset*, as an SPSS portable file. You can use SPSS to open the dataset and then save it as a system file. Anther option in SPSS, after you have opened the file, is to export it as a spreadsheet file for application in Microsoft Excel. File size=23648k, as an SPSS portable file.

Special Software Needed: The instructor and students will need access to the software package SPSS. This package is readily available on most college campuses, either for networked systems or as a standalone application for the PC and Mac.

Other Sources, available on the Web:

- 1) ICPSR Web site, at www.icpsr.umich.edu, contains the description of the dataset written by the Principal Investigators. Type the ICPSR Study Number into the site search engine to reach the description of the dataset and a version of the codebook in .pdf format.
- 2) <u>SPSS User's Manual</u>, available from <u>www.spss.com</u>, which may be useful to more advanced technical operations.

Tips about using the Dataset:

- 1) The dataset includes more than 21,000 rural households from northern townships, drawn from the census of 1860. There are 390 variables in the file that the Principal Investigators deposited in the ICPSR archive, but we will only be using the first 85, concentrating on farm output statistics and the household. Because the dataset is so large, students are urged to think ahead when running procedures so as *not* to generate tables that are hundreds of pages in length. In other words, be very sparing in your use of the cross-tabs procedure. If you are working with an older PC or Mac, be prepared for lengthy processing times.
- 2) Focus of analysis—it makes sense to focus on the rural households from a Midwestern state, and then perhaps an eastern state for comparative analysis. Pick your state of choice and in SPSS run a SELECT CASES procedure (DATA/SELECT CASES) and paste in the variable for state (v3) and the values. If you wanted to select Wisconsin paste in code 8.
- 3) Missing Values—the principal investigators inserted the value of "9999999" when they did not have any data to report for a given household. That can be very distorting if you treat that as an actual result—think of the output if you count missing values as actual bushels of wheat! Therefore, you will want to use the SELECT CASES command to filter out the missing values. You may have to do this for each variable or set of variables you analyze.
- 4) Running FREQUENCIES—Assuming that you have selected a state (Wisconsin, for example) run a frequencies procedure (ANALYZE/DESCRIPTIVE STATISTICS/FREQUENCIES) on v3 and see how many cases you get. Next, run a frequencies procedure on place of birth (v71) and see which state sent the most people to Wisconsin. The great value of the Bateman-Faust sample is that it lets us examine farms in the new, frontier region of the Midwest and compare it to older, more established farm regions of the Northeast. The key variables to examine are farm size and extent of improved land (v11, v12, and v13), the value

- of farm output (v 14 and v15, and v24 to v26), livestock (v16 through v23), land-extensive grain crops (v27, v28, v29, v30, v38 and v39), more specialized crops (v34 to v37, v40 to v42, v48-v62) and crops associated with women's agricultural output (v43 and v44). In addition, we can follow the differences in historical demography by examining v6 through v10. Go ahead and run Frequency tabs on some, or all, of these variables for Wisconsin or Minnesota.
- Solution (STABS) and the SELECT CASES Procedures—CROSSTABS (ANALYZE/DESCRIPTIVE STATISTICS/CROSSTABS) is a procedure that works best when the values of the variables are limited in number. The power of the cross-tabulation is in the ability to see the connections between two variables. The SELECT CASES procedure can be used to reduce the number of cases where there is missing or incomplete data. For example, an examination of gender (v67) shows that there are some cases coded other than male or female. It makes sense to use the SELECT CASES to limit the analysis to only codes 1 and 2 for that variable.
- 6) Computing New Variables—The COMPUTE (TRANSFORM/COMPUTE on the EDIT MENU) procedure in SPSS allows the student to make a new variable based on a combination of existing ones, through an arithmetic calculation. For example, agricultural productivity may be calculated in the most basic way as products per farm. A frequency tabulation on wheat output (v27) gives this statistic. A more refined types of output statistic is to calculate output per improved acre and output per farm worker. In the case of wheat production, a student would use the Compute Variable to make a new target variable, call it "whtimpr" and simply divide v27 by v12, the variable that measures improved acres. The frequency output in the new variable "whtimpr" can be studied on its own, or compared against other variables in additional cross-tabulations, or in a regression analysis.
- Running Regressions—Bi-variate and multiple regression analysis are powerful tools that go beyond cross-tabulations in establishing the direction and extent of a statistical relationship between two or more variables. They may be the most useful SPSS procedures for testing hypotheses that the students have about the lives of the people listed in the dataset. For example, a bi-variate regression (ANALYZE/CORRELATION/BI-VARIATE) between total value of all property held by the head of household (v75) and the age of the head of household (v68) shows a statistically significant relationship. The instructor and students should take care that if they run bi-variate and multiple regressions on the dataset that they are actually testing meaningful relationships. It may be a good research question to ask if wheat production was greater in some states than others. However, testing the influence on the code for states, counties, and townships (v3 through v5) against wheat production (v27) will not produce a usable output, since the first set of variables is not interval in nature. Instead, students will need to make use of the RECODE procedure if they want to test a geographic place against an agricultural output variable.

Research Questions—

- 1) What is the farm commodities output of the unit you are examining (all of the North; region; state; county; or township)? Note: you can use the FREQUENCIES command for this question
- 2) What was the level of farm productivity in the unit you are examining? Note: you will need to use the COMPUTE procedure to determine productivity, either for output per acre, or for output per worker.
- 3) How did rural northerners in 1860 use land? How much was "improved" and how much "unimproved"? What was the value of land in the unit you are examining? Note: you can use the FREQUENCIES command for these questions.
- 4) What relationship, if any, do you find between the size of a household (v7) and its location and land use? Note: you might use the Crosstabs procedure, or even the Regression procedure.
- 5) What did rural northerners do if they did not work on a farm? What sort of occupations do you find among the non-farm rural population that might have been especially valuable to farm families? Was there any difference in wealth-holding between farm households and non-farm households? What about household size? Note: you can use the COMPARE MEANS (ANALYZE/COMPARE MEANS/MEANS) procedure taking care to place the agricultural status of the household (v6) as the independent variable and wealtholding (v75) or household size (v7) as the dependent variables.

Related Secondary Source Readings—

- 1) Jeremy Atack and Fred Bateman, *To Their Own Soil: Agriculture in the Antebellum North*. Ames: Iowa State University Press, 1987.
- 2) Lee A. Craig, *To Sow One More Acre: Childbearing and Farm Productivity in the Antebellum North.* Baltimore: Johns Hopkins University Press, 1993.

Prepared by Jim Oberly, Department of History, University of Wisconsin-Eau Claire Question & Comments?: Email me or telephone me at 715-836-4599. Surface mail: Department of History/UW-Eau Claire Eau Claire, WI 54701 Updated: May 20, 2002