2000 Washington State Population Survey

Data Report

THE 2000 WASHINGTON STATE POPULATION SURVEY (SPS) is the second in a series of surveys designed to provide, biennially, a detailed profile of Washington State residents. The survey provides information on topics such as employment, income, education, immigration, health, and health insurance and borrows its structure and many of its questions from the national Current Population Survey (CPS). The original survey was constructed in 1998, and minor changes were made for the 2000 survey.

The SPS is a valuable complement to other reports and data resources addressing Washington's population. For example, although the federal Census was also conducted in 2000, the SPS provides a consistent source of information at more frequent intervals than the decennial Census. Also, the March CPS measures income statewide on a year-to-year basis, but has a much smaller sample size and does not provide the regional income data found in the SPS.

A total of 6,726 households from the two separate samples completed the telephone interview in spring of 2000. The response rate was 43 percent for the general population sample and was 29 percent for the expanded sample. The average interview time was approximately 24 minutes. The interview questions were only translated into Spanish for the 2000 survey. OFM used the 2000 Census figures for population by county as control totals to convert the raw statistical results of the survey into a recognizable and useful portrayal of the population of Washington State.

Survey Design

Sample Design

POPULATION

The population for this survey consists of all households located within the geographic boundaries of Washington State. Because this was a telephone survey, only the households with telephones were potential subjects. The 1990 census shows that less than four percent of Washington households did not have telephones. Households on military bases and other group quarters - such as student dormitories, prisons, and nursing homes - were also excluded from this survey.[1] Since there is no universal list of all the households as defined above from which a random sample can be obtained, SESRC used the random digit dialing (RDD) approach to obtain the required sample. The RDD approach is most commonly used to ensure equal probability of selection for each household with an activated telephone line, listed or not. Survey Sampling, Inc. (SSI) prepared the RDD sampling frame SESRC used.

SAMPLING

General Population Sample and Expanded Sample. Two separate samples were drawn for this survey. One was a random sample of all Washington State households, or the general population. The targeted number of completed interviews for the general population was 6,000. The second was an expanded sample of households in which the respondent was African American, Asian, Hispanic, or Native American. This expanded sample of minority groups enables data users to make inferences about characteristics of all major population groups. When examining the entire state population, responses from the expanded sample will be weighted to represent the incidence of these groups in the general population. The target for each of the minority groups identified above was 400 interviews from the general population sample and the expanded sample combined.

To control survey costs, SESRC recommended that the expanded sample be drawn only from the census tract regions containing the highest concentrations of each minority group. Since the RDD sample was inclusive of all state regions, it already provided a fair representation for each minority group. SESRC's approach for expanding the sample of these populations was to identify the top five to ten census tracts for each minority group and to obtain a sufficient quantity of telephone numbers to ensure completion of the desired 400 completed interviews for each minority group.

Regional Stratification. The general population sample is stratified into eight geographic regions based on county of primary residence. The target completion for each region was 750. This regional grouping considered the similarities of economic and population characteristics among the 39 counties in Washington State. It was the result of consultation with legislative and other advisory groups for the State Population Survey.

[1] See discussion of limitations.

Western Washington counties were grouped into five regions as follows:

Region 1: Island, San Juan, Skagit, Whatcom
Region 2: Clallam, Cowlitz, Grays Harbor, Jefferson, Klickitat, Lewis, Mason, Pacific, Skamania, Wahkiakum
Region 3: King
Region 4: Kitsap, Pierce, Snohomish, Thurston
Region 5: Clark

Eastern Washington counties were grouped into three regions as follows:

Region 6: Adams, Asotin, Chelan, Columbia, Douglas, Ferry, Garfield, Grant, Kittitas, Lincoln, Okanogan, Pend Oreille, Stevens, Walla Walla, Whitman
Region 7: Spokane
Region 8: Benton, Franklin, Yakima

Questionnaire Design

The initial draft of the 1998 questionnaire was based on the March CPS questionnaire. In addition to the CPS questionnaire, the OFM State Population Survey group solicited additional questions for key subject areas. The draft was reviewed by a group of more than eighty individuals representing different organizations. Their comments were collected and reviewed by OFM, and many of their recommendations were incorporated into the final questionnaire. In 2000, OFM staff reviewed the 1998 questionnaire and made minor changes incorporating some detailed income questions that were missed in the 1998 survey.

II. Survey Administration

The survey was administered by SESRC. Before the full-scale fielding, a pretest of 100 cases was conducted. The full-scale fielding started on February 27, 2000, and the interview phase was completed on June 11, 2000. The average interview time was approximately 24 minutes.

Advance Letter. To obtain full cooperation from the potential respondents, SESRC sent an advance letter to about 4,000 households to announce the survey and explain its purpose. Matching phone numbers with existing directories generated the addresses of these households.

Interview Languages. The interview script was only translated into Spanish for the 2000 survey. Russian, Korean, and Vietnamese translations were eliminated.

Response Rates. A total of 6,726 households completed the survey. Response rates were calculated separately for the general population sample and the expanded sample. The Council of American Survey Research Organizations recommends a calculation method that involves a total account of the sample dispositions and an estimation of eligible households from non-contact cases. According to this calculation method, the response rate for the general population sample is 43 percent and for the expanded sample is 29 percent. A forthcoming technical report will discuss in detail the sample disposition and calculation of the response rates for this survey.

III. The Analysis Data File

The analysis data file consists of 234 variables that were either extracted from the original survey data file or constructed at OFM. In the analysis file, the data are arranged so that each person's data are on an individual record. Thus, a household with five members has five records. The file contains 17,967 persons from 6,726 households.

The analysis data file is available in both SAS format and Excel format. It can be downloaded from the OFM Web-page for SPS. The URL address is http://www.ofm.wa.gov/ under Population, Economic and Labor Force Information/State Population Survey.

IV. Data Tabulations

The data tabulations are presented in two forms, one for categorical and one for continuous variables. The former uses frequency analysis and the latter means analysis.

Each record is weighted and the eight regional values are presented alongside the state value. Also, all tabulations include the variable name, variable label, and the universe for that variable. A subtitle indicates whether the variable is a person variable, a household or respondent-only variable, or a family variable.

A variable with a frequency analysis runs through at least two pages (indicated as Part 1 of 2 and Part 2 of 2). The first page lists the state value and the numbers for Regions 1 through 4. The second page lists the state total again and the numbers for Regions 5 through 8. The number of pages will increase by an increment of two depending on the number of data levels in a variable. Under the state and region headings, the weighted frequency counts and percentages for each data level are listed. Other information in the frequency tables includes a maximum margin of error at the 95 percent confidence level for the state and for each of the eight regions.

Margin of Error. Caution should be used in interpreting tabulations that contain small values with a relatively large margin of error. Take for example the question: *In which state did you [the respondent] live one year ago, if not in Washington?* The weighted tabulation shows that a state total of 7,144 people

lived in New York one year ago. They constituted about 4.7 percent of individuals who were reported to have moved to Washington from another state within the past year. However, the ±4.7 percent margin of error indicates that we are only reasonably confident that the true number of former New York residents is somewhere between zero and approximately 9.4 percent. A common practice to reduce the standard error in such situations is to combine the data levels with fewer categories. In this particular example, instead of individual states, regions can be created.

For each of the means analyses, the numbers for the eight regions and the state total are all listed on one page. This type of table includes the following statistics:

- Total non-missing observations
- Mean
- Minimum
- Maximum
- Median
- Total observations
- Total missing observations
- Sum of weights
- Lower limit of 95 percent confidence interval
- Upper limit of 95 percent confidence interval

It should be pointed out that because of extremely high values in some of the continuous variables, the mean tends to be skewed. In such cases, the median is a better measure of the central tendency.

V. Limitations

Since this survey was a telephone survey, households without telephones were excluded. This non-coverage is, however, quite small. Statewide, the percent of households without telephones was less than 4 percent according to the 1990 census. While there exists the risk of systematically missing some people in a telephone survey, most researchers do not consider it to be a serious problem.[2]

Another limitation common to all surveys is "non-responses." This term refers to households that refuse to participate in the survey. The response rate in this survey is 43 percent for the general population sample and 29 percent for the expanded sample. As in all surveys, there is a potential distortion in the results if the characteristics of the non-responding households are systematically different than those of the responding households. A common practice to partially compensate for the non-response error is to post-stratify the survey based on known population characteristics,[3] which was done in this project.

[2] Folz, D.H. (1996). *Survey Research for Public Administration*. Thousand Oaks, CA: Sage.
Frey, J.H. (1989). *Survey Research by Telephone*. Newbury Park, CA: Sage.
[3] Lavrakas, P.J. (1993). *Telephone Survey Methods: Sampling, Selection, and Supervision*. Newbury Park, CA: Sage.

An examination of the responses suggests that the degree of distortion due to non-responses is small. OFM examined frequencies, means, and medians of selected key variables in the data set and compared the results with alternative data sources. For example, wage data from the survey was compared with wage information from the state Unemployment Insurance System. In virtually all cases where survey data were compared with alternative data sources, the results were very similar. The issue of non-response and comparisons between survey results and alternative data sources for key variables will be discussed in a forthcoming technical report.

A third limitation in this survey is the difference between the design and the post-stratification with respect to group-quarters populations. While the design called for exclusion of group-quarters populations, in the post-stratification process, the group-quarters population could not be separated from the general population estimates. Thus, the survey data were weighted to the entire state population.

Acknowledgments

The following individuals and entities formed the extended network of the 2000 Washington State Population Survey whose input to the design and other aspects of the survey proved valuable.

Lourdes Alvarado-Ramos	WA Veterans Home	Jim Mayfield	WA ST Institute for Public Policy
Bernal Baca	Yakima Valley Comm. College	Jim Medina	OMWBE
Lori Barnett	Chelan County Planning Dept.	David Merrifield	WWU
Sandi Benbrook	CTED	Gary Moore	L&I
Rick Bender	WA State Labor Council	Mike Moore	Cowlitz-Wahkiakum
Debbie Bevier	Laird Norton Trust Co.	Victor Moore*	House
William Beyers	UW	Richard Morrill	UW
Mike Bigelow	SPI	Fred Nick	UW
Gary Bodeutsch	ESD	Tim Norris	ESD
Pamela Bracy	Seattle Public School District	Ellen O'Brien-Saunders	WTECB
Richard Brandon	UW	Christie O'Loughlin	DOA
Ann Bregent	WA State Library	Desmond O'Rourke	Impact Center
Don Brunell	Assoc. of WA Business	Jan Olmstead	Gov.'s Exec. Policy Office
Phillip Bussey	WA Roundtable	Glenn Olson	Clark County
Gregory Chan	CWU	Toby Olson	Comm. on Disability Issues
John Connors	Microsoft Inc.	Tony Orange	Comm. on African-Amer. Affairs
Richard Conway		Cedric Page	HECB
Annabel Kirshner-Cook	WSU	Dave Pavelchek	WTECB
Richard Davis	WA Research Council	Douglas Pedersen	
David Della	Comm. on API Affairs	Rick Peterson*	House
Tom DiDomemico	Benton-Franklin Regional Council	Polly Phipps	Sentencing Guidelines Comm.
Paul Dziedzic	DNR	Tina Preston	WA Fed of State Employees
Stuart Elway	Elway Research	Jennifer Priddy	SPI
Mark Endresen	Joint Council of Teamsters 28	Gary Ray	DOT
Geoffrey Eng	Community Colleges of Spokane	Patricia Richards	DSHS
Vicki Fabré*	LTC	Manuel Romero	Comm. on Hispanic Affairs
Chandler Felt	King County Gov.	Jane Rushford	DNR
Stan Finkelstein	Assoc. of WA Cities	Fred Saeger	WA Assoc. of County Officials
Tom Fitzsimmons	DOE	Eric Schnapper	UW
Judie Fortier	City of Tacoma	Jennifer Scott	Gov.'s Office of Indian Affairs
Jaime Gallardo	Seattle Office for Civil Rights	Bern Shanks	DFW
Marc Gaspard	HECB	Scott Sherman	
Bev Goodman	Dispute Resolution Center	David Schumacher*	Senate
John Griffiths	Boeing	Bob Sicko	Puget Sound Regional Council
Earl Hale	SBCTC	Art Siegal	
Hebert Research Inc.		Kriss Sjoblom	WA Research Council
Dawn Hitchens	HECB	R. Peggy Smith	DOC
Randy Hodgins*	Senate	Chang Mook Sohn	Office of the Forecast Council
Jim Jesernig	DOA	Paul Sommers	Northwest Policy Center
Terry Johnson	Battelle	Roy Standifer	DOP
Allen Jones	SPI	Fred Stephens	Office of the Governor
Susan Jordan	HRC	Stephen Stockett	The Bon Marché
Debora Juarez	Gov.'s Office of Indian Affairs	Ed Strozyk	SPI
Joe Kabel	DSHS	Tom Sykes	HECB
Dennis Karras	DOP	Lucky Tedrow	WWU
David Kaufman	CWU	Tim Triesch	Grays Harbor
Liz Kohlenberg	DSHS	Greg Weeks	
Roxanne Lieb	WA ST Institute for Public Policy	Mary Welsh	DOR
Douglas Ljungren	Port of Tacoma	William Whitlow	Safeco
Bob Longman*	House	Vicki Wilson	HCA
Gary Lowe	WA State Assoc. of Counties	Richard Yukubousky	Municipal Research Serv. Center

*Denotes Legislative Advisory Group

2000 Washington State Population Survey

Technical Report 1: Sample Disposition and Response Rates

THIS REPORT is one of several technical appendices prepared by the Office of Financial Management to supplement the 2000 Washington State Population Survey (SPS). The survey was funded by the Legislature to provide social, demographic, and economic information about Washington residents biennially. A background report on the SPS titled, 2000 *Washington State Population Survey Data Report*, is available electronically at http://www.ofm.wa.gov/, or by calling OFM at (360) 902-0599.

Responses to the survey were obtained from telephone interviews with 6,726 households that represent the state population as a whole. The survey was designed by OFM and conducted by the Washington State University Social and Economic Sciences Research Center (SESRC) in the Spring 2000.

Sample Disposition and Response Rates

There are two base samples for this survey – a general population sample and an expanded sample of racial minorities. The response rates are calculated separately for the two samples.

The general population sample was drawn using the random digit dialing (RDD) technique. With RDD methodology, all households in the state of Washington with an activated telephone line, either listed or unlisted, have an equal probability of being selected. A target of 6,000 households was planned for the general population sample, the main sample of the SPS.

The expanded sample was drawn from census tracts that contained the highest concentration (40 percent or more) of minority groups. This additional sample allowed researchers to obtain sufficient information on racial minority groups for statistical analyses – information that would have been missed using only the RDD method. The expanded population groups include: African Americans, Native Americans, Asian/Pacific Islanders, and Hispanics. The completion target for each of these four groups was a total of 400 households from the general population and the expanded samples combined. For instance, if 250 households of the Hispanic group completed interviews in the general population sample, then only 150 completed interviews of this group would be needed from the expanded sample. A household is defined as a minority household if the respondent is a minority member. The respondent is the household head or the person most knowledgeable about the household's financial situation.

TABLE 1

Complete Sample Disposition for the General Population Sample and Expanded Sample Cases

Sample disposition	OF00 (General Population Sample)	EF00 (Expanded Sample *Before* Adjustment)	EF00Adj (Expanded Sample *After* Adjustment)
CM: Completed Interview	6,025	701	701
PC: Partial complete	517	125	125
PB: Partial complete	8	5	5
CP: Pretest completion	26	0	0
R1: Refusal, hang up	699	874	318
R2: Soft refusal	72	3	1
R3: Hostile refusal	755	223	81
R5: Refusal on conversion attempt	3,771	1,105	402
RN: Refusal for respondent not available	144	55	20
RP: Refusal by other person	3	3	1
CB: Specific callback	107	77	28
GB: General callback	54	43	16
BZ: Busy signal	177	118	43
NA: No answer	1,504	795	290
AM: Answering machine	503	1,001	365
LM: Answering machine, left message	190	0	0
DF: Deaf respondent	81	36	13
HC: Handicapped respondent	61	33	12
LG: Language problem	201	502	183
DD: Respondent deceased	6	4	4
IE: Ineligible, respondent too young	83	35	35
I2: Ineligible, respondent not minority	0	1,451	1,451
IEE: Ineligible estimate respondent not minority	0	0	3,097
BG: Business or government	2,124	616	616
DS: Disconnected	4,601	2,946	2,946
ED: Electronic device	1,034	396	396
BC: Blocked call	272	243	243
NL: No listing	1	0	0
OT: Other misc.	100	63	63
TR: Terminated by Interviewer	5	3	1
UP: Unpublished number	0	0	0
SA: Spanish no answer	4	1	1
SB: Spanish answering machine	6	1	1
SG: Spanish general callback	4	1	1
SM: Spanish answering machine, left message	1	7	7
SZ: Spanish busy signal	5	0	0
PN: Purged nonworking numbers	2,992	1,405	1,405
Total Sample	**26,136**	**12,871**	**12,871**

Calculating response rates for a survey requires first a complete account of the disposition of all sample cases regarding the outcome of the interview attempt. Table 1 presents the complete sample disposition of the two samples in this survey. The first column is a description of the dispositions. The second column (OF00) contains the disposition figures for the general population sample. The third column (EF00) contains the disposition figures for the expanded sample *before* adjusting for the

non-minority status of the non-completes. The last column (EF00Adj) lists the dispositions for the expanded sample *after* adjusting for the non-minority status of non-completes. The adjustment for non-minority status of the non-completes for the expanded sample is discussed below.

Adjusting Non-completes for Non-Minority Status

The calculation of response rates requires that sample units that are not eligible be excluded. For example, because the SPS is a household telephone survey, if a business phone number is encountered, the interview is terminated. The case is declared ineligible and removed from the sample. Determining eligibility can be difficult when attempts to contact fail, as in attempts that result in busy signals, hang-ups, or answering machine pick ups. The Council of American Survey Research Organizations (CASRO) has proposed adjusting response rates to reflect the obtained proportion of ineligible households of all contacted and identified households. In such cases, an estimate is obtained for the proportion of those households that would have been ineligible if the contacts had been successful.

In the expanded sample of the current survey, such estimation is more significant, because of the high proportion of households that are ineligible due to the selection criterion of race/ethnicity. For this sample, if the respondent was not a Hispanic, African American, Asian, or American Indian, the household was considered ineligible. Because of the large proportion of non-minority households and a large number of refusal cases in this sample, the estimation of ineligibles is important to determine the response rate. The last column of Table 1 reflects the adjustment of the disposition of the expanded sample for non-minority households among those that refused or that were not successfully contacted (R1 through LG plus TR).

An estimated proportion of non-minority household was derived for the non-completes in the expanded sample. This proportion was calculated by dividing the identified non-minority households (I2) by all identified households (CM through PB plus I2). The proportion was 0.64. In other words, about 64 percent of the identified cases in the expanded sample were non-minority households. Therefore, only about 36 percent of the identified cases were considered eligible. This latter rate was applied to rows R1 through LG and TR in the third column to obtain an estimated number of eligible cases among the no-contact cases, resulting in the reduced numbers in the corresponding cells in the last column. The rows that were affected by this adjustment are R1 through LG plus TR.

Sample Disposition Summary

Table 2 contains a summary of the sample disposition details. It shows a total of 26,136 cases were released for the general population sample (OF00). For the expanded sample, the total was 12,871 (EF00Adj).

Of the 26,136 cases released for the general population sample, 6,025 resulted in completion (CM), 5,995 in refusal (RF1 and RF2), 2,903 in no contact (NA1, NA2, and SNA), and 11,213 in ineligible (IEH, IEE, and IEO). The 12,871 expanded sample cases were disposed, after adjustment for non-minority status among the non-completes, into the following: 701 completes, 954 refusals, 960 no contacts, and 10,256 ineligibles.

TABLE 2
Sample Disposition Summary

Sample disposition Summary	OF00 (General Population Sample)	EF00 (Expanded Sample *Before* Adjustment)	EF00Adj (Expanded Sample *After* Adjustment)
CM = (CM,CI,BI)	6,025	701	701
RF1 = (PC,PB,CP)	551	130	130
RF2 = (R1,R2,R3,R5,RN,RP)	5,444	2,263	824
NA1 = (CB,GB,BZ,AM,LM,DF,HC,LG,TR)	1,379	1,813	660
NA2 = (NA)	1,504	795	290
SNA = (SA,SB,SG,SM,SZ)	20	10	10
IEH = (DD,IE)	89	39	39
IEE = Nonminority HH (I2,IEE)	0	1,451	4,548
IEO = (BG,DS,ED,FX,NL,OT,UP,PN)	11,124	5,669	5,669
Total	**26,136**	**12,871**	**12,871**

Note 1: Disposition codes in parentheses refer to those in Table 1.

Response Rates

Table 3 contains four different response rates. The reason for inclusion of different response rates is that different organizations may have varying needs for presenting information and some response rates are more appropriate than others. A total of four response rates are calculated. These rates are based on definitions of response rates set by the CASRO.

The first two rows of Table 3 are two adjustment factors. The first adjustment factor (ADJ1) represents the percent of contacted households that are ineligible for the survey because of death, age, or minority status (for the expanded sample). This is 1.3 percent for the general population sample (OF00) and 64 percent for the expanded sample (EF00). Note that this is only 4.5 percent for the adjusted column (EF00(Adj)). Since the minority adjustment has already been factored in to the sample disposition numbers, the only remaining ineligible households are those excluded due to age or death. The second adjustment factor (ADJ2) represents the percent of telephone numbers that are excluded from the sample because they are not residential households.

Table 3
Response Rate Calculations

Response Rate Calculations	OF00 (General Population Sample)	EF00 (Expanded Sample *Before* Adjustment)	EF00Adj (Expanded Sample *After* Adjustment)
ADJ1 = %INELIGIBLE= (IEH+IEE)/(CM +RF1+IEH+IEE)	1.3%	64.2%	4.5%*
ADJ2 = %OUT OF SAMPLE = IEO/TOTAL	42.6%	44.0%	44.0%
RR1 = CM/CM+RF1+RF2	50.1%	22.7%	42.4%
RR2 = CM/CM+RF1+(1-ADJ1)*RF2	**50.4%**	42.7%	**43.3%**
RR3 = CM/CM+RF1+RF2+NA1+SNA+NA2	40.4%	12.3%	26.8%
RR4 = CM/(CM+RF1+ (1-ADJ1)*(RF2+NA1+SNA)+ (1-ADJ2)*NA2)	**42.5%**	25.6%	**29.0%**

* IEE removed from both the numerator and the denominator.

Note 1: Disposition codes on right-hand side of the equation refer to those in Table 2.

The four response rates displayed in the next four rows differ only in what is included in the denominator. To avoid confusion, the following discussion will not involve the EF00 column. The rates under this column are considered to be a less accurate representation of the expanded sample, because they do not exclude the ineligibles from refusal and no-contact cases.

The first rate (RR1) is a crude rate. It is the ratio of completes over the sum of completes (CM) and refusals (RF1 and RF2). RF1 in the denominator refers to refusals after the identifying information is collected. RF2 refers to refusals before the identifying information is collected. This rate can be considered as a crude cooperation rate, i.e. the proportion of all successfully contacted cases that completed the interview. For the general population sample, this rate is 50.1 percent and for the expanded sample (EF00(ADJ)) is 42.4 percent.

The second rate (RR2) is very similar to RR1 except that in RR2, RF2 is adjusted for death and young age of the intended respondents (ADJ1). This rate is a refined cooperation rate. For the general population sample, it increases slightly from 50.1 to 50.4 percent. For the expanded sample, it increases by about one percentage point from 42.4 to 43.3 percent.

The third rate (RR3) takes into account those households that are never successfully contacted during the survey period. They are included in the denominator of the ratio. The denominator for this ratio includes: completed and partially completed interviews (CM), refusals (RF1 and RF2), call-backs (NA1), no-answers (NA2), and Spanish call-backs (SNA). By including the no-contact households without adjusting for ineligibles, this rate is the most conservative of the four rates presented in Table 3. For the general population sample, this rate is about 40.4 percent. For the expanded sample, the estimated rate is about 26.8 percent.

The fourth rate (RR4) is similar to the RR3, but it takes into account adjustment for death and young age among the Spanish call-back cases (SNA), other call-back cases (NA1), and contacted cases from which no identifying information was collected (RF2). In addition, this rate is adjusted for ineligibles (ADJ2) among the no-answer cases (NA2). For the general population sample this leads to an adjusted rate of 42.5 percent. For the expanded sample, the estimated rate is 29.0 percent.

RR2 and RR4 are the two most often quoted rates with the former being generally referred to as the cooperation rate and latter referred to as the response rate.