Appendix:
Technical Documentation

The Survey Process

The Rural Manufacturing Survey (RMS) is a stratified random sample of all U.S. manufacturing establishments with 10 or more employees. The sample was drawn from a list purchased from a private vendor, Survey Sampling, Inc., of Westport, Connecticut. The sampling frame included all establishments in Standard Industrial Classification (SIC) Division D, with the exception of SIC 2711 (newspaper publishing). The Social and Economic Sciences Research Center at Washington State University (WSU) attempted to contact a sample of 8,800 establishments during the summer of 1996. The initial step was a verification interview, in which establishments were called to ascertain whether they were manufacturers with at least 10 employees and, if so, who in the business should be interviewed. These people were then asked to participate in the survey through a letter from the U.S. Department of Agriculture Under Secretary for Research, Education, and Economics.

In the initial screening call to the establishment, the interviewer stated that the questions asked in the study required knowledge of "the types of technology used in manufacturing, management practices, the education, training and pay of production workers, problems in hiring and problems in access to capital for expansion or modernization." The interviewer asked who, at that location, was most knowledgeable about this broad range of issues. The person named was the target respondent, and their phone number and address were taken. About two-thirds of target respondents were either a head of the organization or the general/plant manager (app. table 1). In branch plants, more than half of target respondents were heads of production, while in headquarter establishments the largest number of respondents were heads of the organization. Human resources directors and financial and administrative officers responded in a significant number of establishments.

The data were collected in a half-hour phone interview.[8] At least partial interviews were completed for 3,909 establishments (3,418 by phone; 491 by mail or fax). The 4,891 establishments not contacted included 2,235 determined to be ineligible and 115 eligible cases that refused to participate. Eligibility could not be determined for the remaining 2,656. WSU estimates that about 5,600 of the sampled cases were eligible for the study. The estimated response rate is 68 percent. Stratification of the RMS sample is based on metro-nonmetro location, nonmetro west-nonwest, and three employment size classes. The nine strata are shown in appendix table 2. The goal of the survey was to obtain reliable information on nonmetro establishments as well as a small sample of metro establishments for comparison. Nonmetro plants were more likely to be included in the sample than metro plants (about 7.5 percent vs. 0.7 percent). Also oversampled were large plants and nonmetro plants in the West. The sample includes nearly one-third of large plants in the nonmetro West.[9]

In their comments given at the end of the interview, many respondents said they were not able to answer all of the questions accurately, either because the range of questions was too broad for one person to answer or because they did not have records at hand to provide detailed information. Consequently, information is frequently missing for several variables, including costs, shipments, wages, and employees, especially values for 1992.

Sample Weights

Statistics obtained from a stratified sample do not reflect the population unless a weighting scheme is used to correct for the stratification in the sample. For example, the average number of technologies used computed from the unweighted sample will be affected by the disproportionate number of large establishments (which tend to use more technologies than others). Weights were calculated so that the weighted survey statistics will fairly represent all U.S. manufacturing establishments with 10 or more employees.[10] Let N_h

[8]Potential respondents not reached by phone or lacking time were sent a printed version of the questionnaire, which they could return by fax or mail. The interview was conducted with the most senior manager available at the location.

[9]Although the sample was drawn in 1996, the population numbers in appendix table 2 are from Bureau of the Census, *1993 County Business Patterns,* which was the most up-to-date information available at the time.

[10]For description of analysis of stratified survey data, see Levy and Lemeshow.

represent the population in stratum h, n_h the sample size in stratum h, q_h the sampling rate in stratum h. The sample weights are:

$$WEIGHT_h = \frac{N_h}{n_h} = \frac{1}{q_h},$$
$$where, \sum_{h=1}^{9} n_h WEIGHT_h = N$$

WEIGHT is equal to the inverse of the probability of being included in the sample, and the weights sum to the estimated population size, N.

Computation of Sample Statistics

Mean employment can be estimated by taking a weighted average of means across strata. The mean for a stratified random sample can be computed as,

$$\bar{x} = \sum_{h=1}^{L} \left[\frac{N_h}{N} \right] \bar{x}_h,$$

where $\bar{x}_h$ is the mean for stratum h, and L is the number of strata.[11] This is simply a weighted average of the stratum means and is equivalent to computing an overall mean using WEIGHT. The overall mean computed in this manner is 104 employees. The mean for metro strata is 103, and, for nonmetro strata, the mean is 110. The estimated total employment for the population represented by the sample is simply $N\bar{x}$. Appendix table 3 shows that the nonmetro establishments represent employment of about 4.2 million, and metro establishments in the sample represent about 15.4 million jobs.[12]

Reliability of the estimates can be judged using the standard errors, which are basically complex weighted averages of the standard errors of the various strata. The standard error of the mean $\bar{x}$ is computed by

$$\hat{se}(\bar{x}) = \left(\frac{1}{N} \right) \left[\sum_{h=1}^{L} N_h^2 \left(\frac{s_h^2}{n_h} \right) \left(\frac{N_h - n_h}{N_h} \right) \right]^{\frac{1}{2}}$$

where s_h^2 is the standard error for stratum h.

The term $(N_h-n_h)/N_h$ is the finite population correction (FPC) factor and can be expressed more intuitively as $1-(n_h/N_h)$. When the sample is a large proportion of the population, this term becomes smaller and deflates the sample variance. The FPC takes on values ranging from 0.998 to 0.996 for the three metro strata, but is as low as 0.679 for nonmetro strata. The standard errors can also be computed for metro and nonmetro subgroups. The nonmetro mean employment per establishment is measured with much greater precision (standard deviation of 2) than the metro mean (standard deviation of 8). A 95-percent confidence interval for nonmetro employment per establishment runs from approximately 100 to 108. The standard error of total employment, $\hat{se}(x)$ *is* $N\hat{se}(\bar{x})$, is 90,800 for nonmetro total employment and 1.2 million for the metro total.

A 95-percent confidence interval for total nonmetro employment estimated from the RMS is from 4.01 to 4.36 million (appendix table 3). Note that the Census Bureau estimate from the Annual Survey of Manufacturers (4.04 million) for 1994 falls within this interval, although it is near the lower bound. The Bureau of Economic Analysis (BEA) estimate for the same year is slightly above the upper bound, at 4.4 million. A 95-percent confidence interval for metro employment estimated from the RMS runs from 13.1 to 17.7 million. The Census Bureau estimate for 1994 falls below the lower bound of the confidence interval, but the BEA estimate falls within the interval. We can be fairly confident about the representativeness of nonmetro results since the RMS estimate falls between two other estimates. The metro estimate is consistent with the BEA estimate, although it is higher than the Census number.

[11] See Levy and Lemeshow, chapter 6, for more details.

[12] The intent of the RMS was not to estimate the number of jobs in nonmetro manufacturing, since those estimates can be obtained from other sources. Comparisons with other estimates are made here to judge the representativeness of the sample. If the RMS data produce estimates comparable to those obtained from other sources, we can have greater confidence in the data.

Most of the questions on the RMS are yes-no questions, thus most analysis will be of proportions, $p_y = N_y/N$ (often converted to percent), where N_y is the number of establishments having the characteristic y. The estimate $\hat{p}_y$ is found as a weighted average of the stratum values, $\hat{p}_{hy} = n_{hy} / n_{h,}$ analogous to the formula for $\bar{x}$.

The standard error is computed by

$$\hat{se}(p_y) = \frac{1}{N}\left[\sum_{h=1}^{L} N_h^2 \left(\frac{\hat{p}_{hy}\left(1-\hat{p}_{hy}\right)}{n_h - 1}\right)\frac{N_h - n_h}{N_h}\right]^{\frac{1}{2}}$$

where $\hat{p}_{hy}$ is the proportion with characteristic y computed for stratum h.

Using these formulas, as an example, we computed the percentage of establishments reporting use of computer-assisted design or engineering (Q17 on the RMS questionnaire). As was the case with employment, the means are found by taking weighted averages of stratum means. The nonmetro mean is 44.8 percent, while the metro mean is 53.8 percent. Standard errors within strata are fairly large, due to the small number of observations within each stratum. The metro and nonmetro means are measured with more precision, however. A 95-percent confidence interval for the nonmetro percentage ranges from 42.8 to 46.8 percent. The t-value for a test of the difference between metro and nonmetro percentages is equal to 6.1, indicating a statistically significant difference.

Appendix table 1—People contacted as survey respondents

Position in plant	Branch plants	Others	Total
		Percent	
Head of:			
Organization (owner, CEO, VP)	43.2	6.2	30.7
Production (General or Plant Manager)	26.2	57.0	36.6
Human Resources (Personnel Director)	9.3	17.9	12.2
Finance/Administration (CFO, Office Administrator)	9.6	5.4	8.2
Department			
Production (Foreman, Engineer)	1.0	3.0	1.7
Human Resources	2.9	4.3	3.4
Finance (Bookkeeper, accountant)	1.3	0.5	1.0
Other or missing	6.3	5.9	6.2
Total	100	100	100
Number of cases	2,591	1,315	3,906

Source: ERS Rural Manufacturing Survey, 1996.

Appendix table 2—Number of establishments by strata, Rural Manufacturing Survey

Stratum: Geography	Stratum: Plant size	Establishments		Sampling rate
		Sample[1]	Population[2]	
Region	*Employment*	*Number*	*Number*	*Percent*
Metro	10-49	365	97,920	0.4
Metro	50-249	503	41,788	1.2
Metro	250 or more	197	10,215	1.9
Nonmetro-West	10-49	172	2,815	6.1
Nonmetro-West	50-249	135	978	13.8
Nonmetro-West	250 or more	63	196	32.1
Nonmetro-Nonwest	10-49	851	19,776	4.3
Nonmetro-Nonwest	50-249	997	10,613	9.4
Nonmetro-Nonwest	250 or more	626	3,654	17.1
Metro total	NA	1,065	149,923	0.7
Nonmetro total	NA	2,844	38,032	7.5
Overall total	NA	3,909	187,955	2.1

[1]Completed usable interviews. [2]Estimated from U.S. Department of Commerce, County Business Patterns.
Source: ERS Rural Manufacturing Survey, 1996.

Appendix table 3—Comparison of RMS total manufacturing employment with other estimates

Item	Nonmetro	Metro
	Million	*Million*
RMS estimate, 1995	4.2	15.4
95-percent confidence interval	(4.01, 4.36)	(13.06, 17.74)
BEA, 1995	4.44	14.79
Census, 1994	4.04	12.87

Source: ERS Rural Manufacturing Survey; ERS analysis of Bureau of Economic Analysis unpublished data;
U.S. Department of Commerce, Bureau of the Census, Annual Survey of Manufacturers, unpublished tabulation.