STATISTICAL PROCESS CONTROL IN DECENNIAL CENSUS INDUSTRY AND OCCUPATION CODING*

Michael Mersch, Phil Gbur, and Chad Russell, U.S. Bureau of the Census
Michael Mersch, Bureau of the Census, DSSD, Room 3771-3, Washington, DC 20233

Key Words: Industry and Occupation Coding, Statistical process control, variability, Three-way independent verifiation, outliers

I. INTRODUCTION

A. Overview

A continuous improvement process is a part of the broad focus of Total Quality Management. It means that an organization creates and sustains a positive and dynamic working environment that fosters teamwork, applies quantitative methods and analytical techniques, and draws upon the creativity and inventiveness of all the people. With this in mind, statistical process control is an endeavor to add value or quality to a product or service.

Statistical process control nourishes quality while the product is being produced, not afterwards. The goal of statistical process control is to generate the highest level of consistency through the collection, analysis, and interpretation of data.

The effects of a working statistical process control system are to continuously improve the process by reducing variability. This is accomplished by: 1) achieving consistency by implementing corrective action when problems are detected, 2) simplifying procedures, methods, and tools, 3) measuring the long-term performance level after the process has been brought under statistical control, and 4) providing information for better feedback and a permanent record of actual performance.

B. New Philosophy

As a result of experiences gained from the 1980 Decennial Census it was determined that a change in philosophy was required if we were to attain a high level of quality. In 1980 we used a quality control type system whereby we inspected and repaired "poor" quality work. This philosophy created a number of management problems of quality versus production.

Therefore, for 1990 we had to change this philosophy. After attending several seminars by or about Dr. W. Edwards Deming and the use of statistical methods, we decided to give it a try. So in planning for the 1990 Decennial Census we decided to use the "Do it right the first time" philosophy.

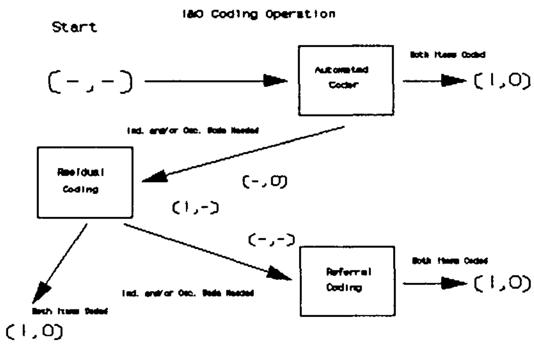
II. THE INDUSTRY AND OCCUPATION CODING PROCESS

A. I and O Production Coding

Industry and Occupation coding was accomplished through a combination of automated and computer assisted clerical coding. The automated coding was conducted at headquarters in Suitland, MD, and the clerical coding at our Kansas City Processing Office (KCPO). All write-in response data on the questionnaire were first keyed into a computer database for processing by the automated coding system. The automated coder calculated a "score" based on a number of factors including the number of matching words in an industry and occupation description, and the "weight," or importance, of the matching words. The score was thus a measure of the computer's confidence that the automated coder had assigned the correct code. If the automated coder derived a score greater than the targeted score, a code was assigned. If the automated coding system could not code the data with a minimum score, then the data were sent to the clerical coders.

Clerical coding operated on two levels - residual and referral coding. Cases first passed to the residual coding unit. If residual coders were not able to code an industry or occupation item, the case was referred to the referral unit by assigning a referral code. The referral coders assigned the final code. Figure 1 shows this coding system.

Figure 1



Clerical coding was computer assisted with each coder sitting at a computer terminal. Each industry and occupation case requiring a code was displayed on a coder's computer screen. The coder entered the necessary codes directly, using the keyboard.

B. <u>Workload</u>

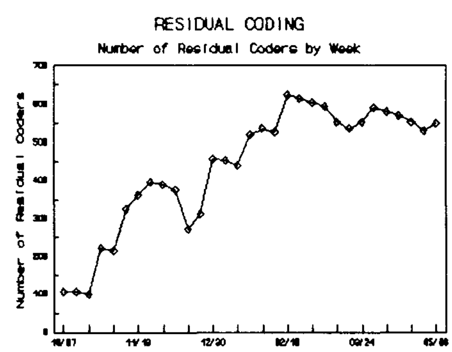
Table 1 shows that in 1990 a little more than 22 million persons required industry and occupation coding. The automated coder coded approximately 58 percent of all industry responses and about 37 percent of all occupation responses, or roughly 47 percent of all responses. Residual and referral coding assigned codes to the remaining 42 and 63 percent of the responses forwarded to clerical coding for industry and occupation items, respectively.

Table 1
1990 DECENNIAL CENSUS
INDUSTRY AND OCCUPATION DATA

WORKLOADS	
Questionnaires	18,144,371
Persons	22,175,992
Codes	44,351,984
AUTOMATED CODER	
(Codes Assigned)	
Industry	12,826,129 (58%)
Occupation	8,199,268 (37%)
Total	21,025,397 (47%)
Est. Error Rate	6.2% Ind., 11.8% Occ.
RESIDUAL CODING	
(Codes Assigned)	
Industry	9,352,181 (42%)
Occupation	14,094,888 (63%)
Sub-Total	23,447,069 (53%)
QA	3,031,157
Total	26,478,226
(Productivity)	
Codes/Hour	94
Est. Error Rate	8.2% Ind., 11.5% Occ.
Referral Rate	13% Ind., 9% Occ.
Schedule	10/1/90 - 5/06/91
Cost (MIS)	$4,853,048
REFERRAL CODING	
(Codes Assigned)	
Industry	1,608,512 (incl. QA)
Occupation	1,668,438 (incl. QA)
Total	3,276,950
(Productivity)	
Codes/Hour	61
Est. Error Rate	12.3% Ind., 12.7% Occ.
Schedule	11/19/90 - 5/17/91
Cost (MIS)	$663,330

Thus, as a result of the automated coder, it required roughly a peak staff of about 600 coders spread over day and night shifts to code the approximately 23.5 million combined industry and occupation responses in seven months. This was a significant decrease from 1980 in the number of coders and time required to code all industry and occupation responses. In 1980 roughly 1200 coders worked for 13 months to code about a million fewer responses.

Figure 2



C. **Training and Qualification**

Each clerk received two weeks of intensive training on coding concepts, with practice on the VAX computer terminals. The training was divided into two phases - one week of industry training and one week of occupation training. To go into actual production coding, the residual clerks had to qualify separately on **both** industry **and** occupation coding. After practice on a coding test deck, each coder had to pass at least one of three additional test decks for each item. The test deck consisted of cases similar to those the clerk could expect to encounter during production.

D. **Quality Assurance System**

Three-way independent coding was used to monitor the quality of both computer and clerical coding. Quality assurance samples were selected from: cases completely coded by the computer (computer sample), cases passed to and coded by the residual coding unit (residual sample), and cases passed to and coded by the referral coding unit (referral sample). Each sampled case was replicated twice, resulting in three "copies", or quality assurance (QA) cases. These three copies were distributed among three different work units assigned to different coders. After the work units containing corresponding quality assurance cases were completed, the assigned coders for the cases were compared.

The principal of "majority rule" was used to judge the codes assigned to the QA replicates. A code assigned by a clerical coder (residual or referral) was considered "in error" if it was the minority code in a minority/majority situation. For example, if the three independent coders coded a response X,X, and Y, the two coders who coded X were considered the majority coders and the coder that coded Y was considered the minority coder.

For purposes of making a decision of correct/incorrect, a referral code was considered an assigned code. The coder who assigned the minority code is called the "minority coder." The correct code was used as the production code when a clerical majority existed (except when the majority code was the referral code). Otherwise, the case was sent to referral coding. Three-way differences in the referral sample were not sent to the referral unit again, rather the "first" code determined by the Computer Assisted Clerical Coding (CACC) system was used as the production code.

E. **Monitoring**

A major improvement in the 1990 Census Industry and Occupation operation, as compared to previous censuses, was the ability to provide immediate feedback of results from the QA process. The computer system used to control this operation generated data that was used to assist the clerks and constantly improve their performance. The data the system generated included: production rates, majority/minority rates, three-way difference rates, and referral rates.

The Computer Assisted Clerical Coding (CACC) system tracked industry and occupation minority rates for each clerk and coding unit, as well as production, three-way difference, and referral rates. These statistics were reported to supervisors weekly in printed form. The Weekly Coder Performance Summary reported these statistics by coder for the current week and for the three weeks prior to the current week. The Weekly Unit Performance Summary did the same for each coding **unit**.

Using these reports, supervisors could identify coders and coding units with unusually high minority or referral rates and those with low production rates. To assist them in identifying such "outliers," the CACC constructed a box plot with the weekly quality statistics.

Other reports include the Weekly Individual Performance Summary, the Daily Dependent Review, and the Three-way Difference Reports. The Weekly Individual Performance Summary was a more detailed version of the Weekly Coder Performance Summary, produced for each coder flagged as an outlier.

Coders with minority rates in the upper 25th percentile from the previous day would appear on the Daily Dependent Review report, which listed the industry and occupation write-ins and the codes assigned to them by each of the three independent coders. This report was probably the most useful in detecting systematic errors (misunderstandings of procedure).

Another report that the supervisors had access to was the Control and Tracking System for Industry and Occupation (CATSIO) Control File. This file showed the supervisor the status of each work unit. The data included: 1) to whom each work unit was assigned, 2) the date the work unit was completed, 3) if the work unit passed the quality assurance criteria, 4) the number of industry and occupation items checked and in error in the work unit. Also, this file allowed the supervisor to view the status of clusters of batches. These batches contained the replicated quality assurance sample cases. Thus, all work units in a batch were required to be completed in order to compare the assigned codes and thus determine the quality of each work unit. This file allowed the supervisor to determine if any work unit had been assigned and had not been completed within a certain number of days. This was useful when coders were sick for an extended period of time or had resigned without completing their assigned work. In these situations the quality assurance data for the work units within a batch could not be computed. Therefore, the supervisor could reassign any uncompleted work units to another coder.

F. Quality Circles

A quality concept, Quality Circles, was introduced to I&O clerical coding in the 1988 Dress Rehearsal and was used in 1990. Comprised of weekly meetings of each coding unit, these Quality Circles provided a systematic forum for staff to ask questions and offer suggestions for improvement. Coders met with a representative for their unit. The unit representatives then met as a group with one of the two headquarters coding specialists who had been assigned to the KCPO for the entire coding operation. As necessary, suggestions and questions were sent to headquarters for action or reply within two weeks time of the quality circle meeting.

The Quality Circles became an excellent tool for communicating between coders and management. Many coders looked forward to the meetings, which gave them a sense of involvement in the process; they often came prepared with written lists of problems and suggestions.

G. Quality Incentive Bonuses

Another quality concept was a formal group quality incentive program. The goal of the I&O Quality Compensation Program was to recognize coding units who displayed a superior level of performance when compared to other coding units on their shift. Superior performance was defined in terms of production rate and level of quality, as reflected in weekly unit summaries. An I&O coding unit could achieve one of two group bonuses. The bonuses were items such as coffee mugs, desk clocks, pen and pencil sets, and teddy bears, which all had an I&O logo on them.

III. LIMITATIONS

The minority/error rate is a better estimate of the true error rate when there is a unique "true" code for each write-in. Unfortunately it is possible for all three codes in a three-way difference to be "true." Further, while the minority rate for an individual coder lies in the interval [0,1], the overall error rate based on these definitions is at most one third, since two other coders must agree against the minority coder for an error to occur.

For this report, both residual and referral coders were used to evaluate the computer coded universe. Computer coded items (not cases) were present in both the residual and referral samples and were verified by **either** residual or referral coders, depending on which sample (residual or referral) the case was in. The minority rates presented in this report for different code sources are only comparable to the extent that the two measuring systems (verification by residual coders and verification by referral coders) are the same.

Standard errors of estimated error rates are computed using a simple binomial model. This is probably not appropriate for error rates where portions of the QA sample are excluded (see Appendix).

IV. RESULTS

A. Summary of Coding Results

The probability that a particular item (industry or occupation) was coded correctly is shown on Table 2. This probability or "success rate" is one minus the estimated error rate. Success rates are estimated for each code source.

Table 2

Estimated Accuracy of Coding

	Automated Coder	Residual Coding	Referral Coding	Overall
Industry				
Percent Coded	57.8000	36.0000	6.200	
P (correct)	0.9000	0.8700	0.860	0.8900
Standard Error	0.0005	0.0006	0.001	0.0004
Occupation				
Percent Coded	36.8000	56.8000	6.400	
P (correct)	0.8700	0.8600	0.870	0.8700
Standard Error	0.0006	0.0005	0.001	0.0004

B. Effect of the Quality Assurance System

The purpose of the QA system was not only to measure the quality of coding but to positively influence the operation in some measurable way. To determine the extent to which this occurred, the average level of some quality measure Y as a function of the number of weeks that a coder had been coding, X, was examined.

Minority or error rates measure the level of agreement/ consistency between coders. Figure 3 shows the average industry and occupation minority rates as a function of coding experience measured in weeks. The minority rate for occupation was consistently higher than that for industry items.

Figure 3

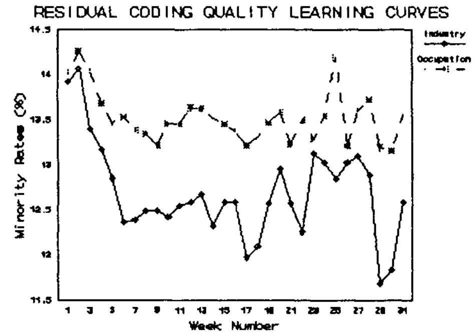


Figure 4 shows the average production rate as a function of coding experience. As expected, production rates increased steadily as a coder gained experience - rapidly at first, then more slowly. With minority rates holding steady during the same period, this suggests that coders learned to code faster with the same level of quality.

Figure 4

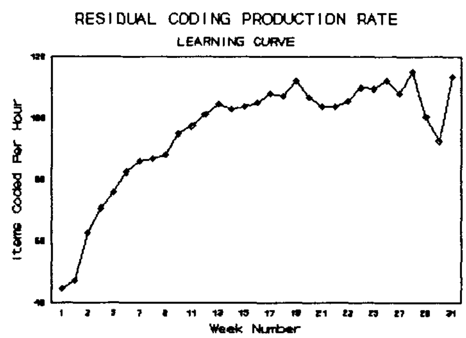
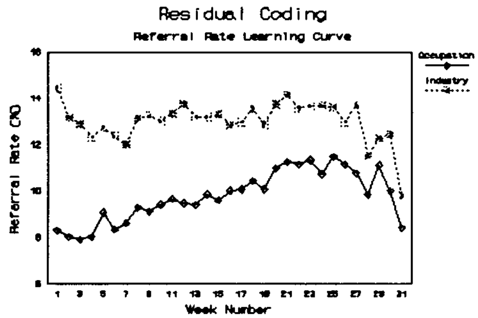


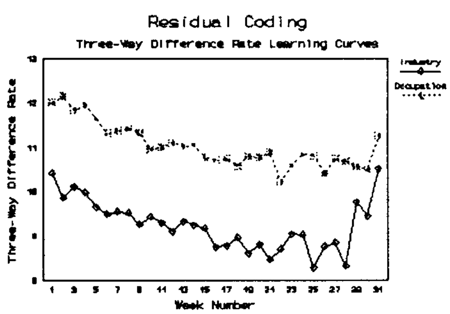
Figure 5 shows the referral rates for industry and occupation items as a function of coder experience. The referral rate is the proportion of items which were assigned to the referral coding unit. These items were considered by the residual coders to be too difficult to code. The referral rate for industry items was higher than that for occupation items. A slight upward trend is apparent among the occupation referral rates in Figure 5, while the mean industry referral rate remained stable.

Figure 5



A three-way difference occurs when the three independent coders each assign a different code to an item. Thus, we believe a higher three-way difference rate indicates a higher "level of confusion" among the coders. Figure 6 shows the three-way difference rates as a function of coder experience. Since the three- way difference rates remained relatively unchanged over time, it does not appear that our process lowered the "level of confusion."

Figure 6



C. <u>Comparison with 1980 Census</u>

The 1990 I&O coding process was largely automated, while in 1980, the I&O coding process was totally clerical. The automated coder coded about 47 percent of all items, nine (9) percent more than estimated in the planning phase. This reduced the workload going into clerical coding. The CACC, with its on-line references and automatic data collection features, made the coding process less cumbersome and easier to monitor than the paper driven process used in 1980. Table 3 compares the 1980 and 1990 I&O Coding operations on a few key points.

Table 3

Comparison with 1980 Census

	1980	1990
Method of Coding	Clerical	Automated & Clerical
Est. Error Rate - Ind	13.0% ± 0.5%	9.0% ± 0.06%
Est. Error Rate-Occ	18.9% ± 0.6%	11.0% ± 0.05%
Operation Time	13 months	7 months
# of Items Processed	43.2 million	44.3 million
# Processing Sites	3	1
# Coders (peak prod)	1200	600+

Estimated error rates for the 1990 operation in Table 2 were computed without including indeterminate cases caused by referrals. This is thought by some to be a better measure of outgoing data quality. The quality or accuracy rates given in Table 1 include the referral cases. The success rates in Table 1 can be thought of as estimating the probability that a coder **acts** correctly, which includes referring difficult cases. Error rates for 1980 and 1990 were computed by different methods, and should not be compared based on standard error alone.

V. CONCLUSIONS

The 1990 Industry and Occupation Coding operation coded approximately 44.3 million responses (industry **and** occupation).

The automated coder was more successful at coding industry items than occupation items. Fifty-eight percent of industry items were coded by the automated coder compared with 37.0 percent of occupation items. The clerical portion of the operation lasted seven (7) months and employed between six and seven hundred coders at peak.

The estimated overall success rate (one minus the error rate) is 0.89 for industry items (computer and clerically coded), with a standard error of 0.0004. The estimated overall success rate for occupation items (computer and clerically coded) is 0.87, with a standard error of 0.0004.

The overall production rate (clerical) was 82.8 items coded per hour.

The referral rate (the relative number of items assigned a referral code) was 13.1 percent for industry items and 9.2 percent for occupation items.

The estimated three-way difference rate was 9.41 percent with a standard error of 0.05 percent for industry items and 11.28 percent with a standard error of 0.04 percent for occupation items.

The only measure of quality which increased as a function of coder experience was the production rate. Coders got faster as they gained experience, with no significant change in minority rates.

In terms of processing time and convenience, the I&O Coding operation was much better in 1990 than in 1980. The operation owes a great deal to the success of the automated coder and the Computer Assisted Clerical Coding (CACC) system. The automated coder greatly reduced the workload of clerical coders. The CACC made clerical coding more convenient and easier to monitor than the paper driven process used in 1980. Automatic monitoring and report generation enabled managers to detect and correct problems on a much more timely basis.

Of all the quality measures tracked by the CACC, the only one to show improvement was the production rate. Overall minority rates remained stable. Why? Changes in minority rates may be hidden by limitations in the measurement system. There are no data on how or how often supervisors used the information in the CACC reports. Neither is there any data on the content/quality of feedback given to coders, or how timely such feedback was. Supervisors may have been overwhelmed with too much or the wrong type of information.

It also may be the situation that we have improved the coding system as much as can be done without changing the system itself. Another possibility is that we use the level of confusion as a measure of quality rather than the concept of incorrect coding.

VI. RECOMMENDATIONS

While 1990 was better than 1980, there are still areas where improvements can be made. In future operations of this type, it might prove useful to monitor the feedback that is given in terms of frequency, timeliness, and content. Also of interest would be how often a particular type of statistic (a unit minority rate, an individual referral rate, etc.) leads to the detection of a problem. With such data it would be easier to determine how well the monitoring/feedback systems worked, and to determine which reports/statistics were most useful in detecting problems.

Another possibility is to teach the coders how to monitor themselves with the information provided by the CACC (or other) system. The system might report such statistics to the coder's screen at logon time. This would eliminate the middle person and assure that coder's received consistent feedback as soon as possible. However, computers can give only information, not advice. Perhaps the coders could not make use of such information without supervisory guidance.

Advances in computing may lead to better automated coding algorithms. It is much easier to control the quality of an automated process than to control a clerical operation involving hundreds of individuals. Likewise, improved technology will hopefully increase the speed and efficiency of clerical coding systems like the CACC.

We should not forget that achieving the quality standards requires a preventative philosophy, not a reactive one. Technical skill should be applied at the earliest stages of the design of a system and include statistical process control capabilities. It also should be remembered that statistical process control involves everyone in the process and aims for continuous improvement.

References

Russell, Chad E.; DSSD 1990 Decennial Census Memorandum Series #N-92; December 14, 1992; Titled: "Quality Assurance Results for the 1990 Stateside Industry and Occupation Coding Operation."

Mersch, Michael L.; 1980 Preliminary Evaluation Results Memorandum #68; January 17, 1984; Titled: "Results of Processing Center Coding Performance Evaluation."

Scopp, Thomas S. and Tornell, Steven W.; Paper presented at the Southern Demographic Association annual meeting; October 11, 1991; Titled: "The 1990 Census Experience with Industry and Occupation Coding."

*This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.