

## MEASUREMENT DIFFERENCES IN KEY ECONOMIC INDICATORS

Karen L. Goldenberg and Diane K. Willimack\*

Karen L. Goldenberg, BLS, 2 Massachusetts Avenue N.E., Room 4985, Washington, DC 20212, USA

**Key Words:** Measurement error; employment statistics; concept measurement

### 1.0 INTRODUCTION

Changes in employment, earnings, and work hours over time are standard indicators of economic health in the United States. These key concepts are measured through surveys of business establishments. In the decentralized United States statistical system, individual statistical agencies are allowed some discretion in defining and measuring survey concepts, leading to variations in understanding and interpretation of indicators which appear to be completely equivalent.

This paper has two objectives. The first is to compare the collection of economic indicators by two U.S. statistical agencies in terms of their concepts and presentation to respondents. We focus our efforts on establishment-level measures of "employment," "production worker employment," "payroll," and "work hours." The second goal is to determine whether questionnaire differences are reflected in published survey estimates.

To achieve our objectives, we first describe two self-administered government establishment surveys that collect the indicators and present an in-depth examination of their collection across the surveys. Then we compare published statistics for employment,

production worker employment, payroll, and work hours. In doing so, we consider whether respondent use of the data collection instrument contributes to any differences observed between the two surveys.

### 1.1 Background

The work presented here is an outgrowth of an earlier study investigating the measurement of work hours in establishment surveys conducted by either the U.S. Bureau of Labor Statistics (BLS) or the U.S. Census Bureau (Fisher et al. 2001). Since work hours are based on specified groups of employees, the concept review included measures of employment. Fisher et al. highlighted differences in the economic concepts of interest in terms of question wording and questionnaire format and layout. They also compared published survey estimates from a Census Bureau survey and a BLS survey. We build upon the earlier research by examining the economic indicator concepts in considerably more detail, and by adding establishment payroll to the list of indicators. We extend the earlier research with a detailed statistical analysis of the published data.

### 1.2 Surveys Reviewed

We focus on the Current Employment Statistics (CES) Survey at BLS, and the Annual Survey of Manufactures (ASM) at the Census Bureau. Both of these surveys publish data on total employment, production worker employment, production worker payroll, and production worker hours. ASM is a panel survey of establishments in manufacturing industries conducted annually by the Census Bureau. The survey consists of a mail-out/mail-back questionnaire sent to approximately 55,000 selected establishments, supplemented statistically using administrative data for businesses with fewer than five employees and for new businesses. While the ASM sample covers all establishments of companies with multiple units and single unit companies with five or more employees, estimates represent the entire manufacturing sector. The survey includes new units and adjusts for closed businesses.

CES is a monthly BLS survey with a sample size in 2002 of approximately 300,000 business establishments in all major industrial sectors. It is the source of data on month-to-month change in U.S. payroll employment, hours, and earnings, by detailed industry. Respondents report monthly by mail, fax, telephone interview, or touchtone telephone data entry, using industry-specific questionnaires to compile monthly information on six or seven variables. Our research examines the seven-variable CES Manufacturing questionnaire. Data used in this paper were based on a

---

\* Karen Goldenberg is a Survey Methodologist in the Office of Employment and Unemployment Statistics at the U.S. Bureau of Labor Statistics (BLS). Diane Willimack is Chief, Establishment Survey Methods Staff, Economic Statistical Methods and Programming Division, at the U.S. Census Bureau. The opinions expressed in this paper are those of the authors and do not reflect official policy of the Bureau of Labor Statistics or the U.S. Census Bureau. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

The authors extend their appreciation to Amy Anderson and Eileen O'Brien of the U.S. Census Bureau and Sylvia Fisher, formerly of the Bureau of Labor Statistics, for their contributions to earlier phases of this research; to Richard Sigman of the U.S. Census Bureau for his guidance with the statistical analysis; to the staffs of the Annual Survey of Manufactures and the Current Employment Statistics survey for their support; and to Martin David, Judy Dodds, John Gates, Patricia Getz, Nash Monsour, George Stamas, Jay Stewart, and Clyde Tucker, who provided helpful review comments on earlier versions of this paper.

sample of approximately 58,000 manufacturing establishments.

## 2.0 QUESTIONNAIRE REVIEW

### 2.1 Methodology

Building on the Fisher et al. (2001) review, our evaluation began with a detailed assessment of the concepts of "total employment", "production workers," "employee hours," and "payroll," as collected by the ASM and CES surveys. We discuss in detail both the measurement of each concept and the potential implications for measurement error. We address the question: To the extent that the measures differ, how might those differences affect the resulting data?

Our analysis looked separately at the data collection forms and at the survey instructions. We evaluated the four variables in terms of: (1) the survey reference period; (2) the strategy for asking questions on the survey instrument or data collection form; and (3) the strategy presented in the instructions for answering those questions. The result, a detailed compilation of definitions, differences, and similarities, appears in Table 1.

### 2.2 Results

***Total Employment and Production Worker Employment.*** Establishment surveys collect the number of workers engaged in a relationship with a business involving the exchange of labor for pay. Put another way, "total employment" (sometimes called "payroll employment") refers to the number of jobs that the establishment provides during the reference period, and is a count of employees on the payroll.

"Production workers" in manufacturing establishments are the employees who fabricate, assemble, or otherwise contribute to producing the establishment's goods. Some data are collected only for production workers.

**Survey reference period.** The reference period for establishment surveys conducted by or for U.S. government agencies is "the payroll period containing the 12th of the month" (U.S. Office of Management and Budget, 1978). As a monthly survey, all CES variables refer to the pay period including the 12th of the reference month (usually the current month). The ASM uses the pay period including the 12th of specific months, one from each calendar quarter.

**Questionnaire.** Total employment and production worker employment are treated separately on CES, but are intermingled on the ASM. The CES form defines All Employees directly on the questionnaire, and explicitly states that production workers are a subset of total employment. We refer to this subsetting approach as "top down." ASM respondents compute Total Employment from production worker employment and

All Other Employment, implying the subset from the structure and layout of the question through what we describe as a "bottom up" strategy. That is, the questionnaire first asks for the production worker counts for the pay periods that include the 12th of March, May, August, and November. It next directs the respondent to compute the average number of production workers across those four pay periods, and finally obtains Total Employment by summing the production worker average and the number of All Other Employees for the March 12 pay period. As a derived figure, the resulting employment figure may not represent an actual employment count for the establishment.

**Instructions.** Both ASM and CES instructions begin with the general concept of employment and provide guidance on who to count as an employee. Instructions for both surveys specify a set of employees to include and to exclude. One difference between the two forms is that CES specifically excludes employees of contractors. ASM implicitly excludes them through the general employment concept, which is defined as the number of employees reported on IRS Form 941, Employer's Quarterly Federal Tax Return. (Contractors are not included on the client firm's Form 941.)

While the content is generally similar, the instructions communicate the information in different ways. CES uses short lists to identify employees to include and to exclude. ASM presents the "includes" and "excludes" in paragraph form, with sentences alternating between "includes" and "excludes" rather than grouping each set together.

The survey instructions also describe production workers. CES uses short column "include" lists that show production work occupations, followed by a separate list of job functions to exclude. In contrast, the ASM instructions begin with an explanation of Employment, setting a context for subsequent instructions regarding production workers, even though the form collects quarterly production workers first. From this point, the instructions parallel the approach on the form, listing types of production work in paragraph form. There are no exclusions; rather, the production worker list is followed by a similarly detailed "Include" list for All Other Employees, which corresponds to the "excludes" from CES.

***Payroll.*** "Payroll" refers to the aggregate total wages paid to an establishment's employees for their services. The CES Manufacturing form collects payroll for production workers each month, using a limited concept of payroll that excludes irregular payments such as bonuses. ASM obtains an annual figure, which is a more inclusive payroll measure.

**Reference period.** The CES reference period for both payroll and work hours is the pay period including the 12th of the month. The comparable reference period for the ASM for both of these items is the entire

previous calendar year. The ASM reference periods are different from those used for total employment or production worker employment.

**Questionnaire.** The CES form has both a column heading for Production Worker Payroll and a definition of payroll on the actual form. The ASM form shows topic headings only.

**Instructions.** CES requests payroll for production workers for the pay period including the 12th of the reference month. The instructions specify that pay should be reported before all deductions and list earnings components to include and to exclude. CES payroll emphasizes regular payments and excludes occasional or irregular payments which might distort the monthly earnings data series.

The ASM collects *annual* payroll for all employees, first for production workers and then for all other employees, with instructions in paragraph form. Annual payroll includes pay for vacation and sick leave (paid holidays are not mentioned) and the cash equivalent of payments in kind. The latter item is specifically excluded on CES.

**Work Hours.** “Work hours” refers to the time component of the exchange of labor for pay. As a rule, work hours in establishment surveys come directly from employer records and refer to actual hours, as most production workers are paid by the hour. Both CES and ASM collect work hours only for production workers. The key difference is that ASM asks for *hours worked*, while CES obtains *hours paid*.

**Questionnaire.** CES asks for production worker hours paid, including overtime, as well as a separate report of overtime hours. ASM asks for “Plant hours worked by production workers (Annual),” and uses the word “Total” as part of a subheading.

**Instructions.** Once again, instructions provide considerably more information about what is to be reported than the questionnaire. For example, the ASM form refers to “hours worked.” A respondent looking only at the questionnaire could miss the fact that “hours worked” is not the same as “hours paid”—particularly since the hours item follows the payroll question, which *includes* earnings for paid leave. The instructions list components of hours worked in paragraph form, including the statement “Include all hours worked or paid for, except hours paid for vacations, holidays, or sick leave.” We found this phrasing to be confusing because time paid normally includes paid leave, which is specifically excluded.

CES requests hours paid and the instructions show various types of paid leave to include along with regular work hours. The instruction defines hours paid as the sum of hours worked, hours paid for “portal-to-portal, stand-by, or reporting time,” and hours of paid leave.

### 3.0 IMPLICATIONS FOR MEASUREMENT ERROR

At first glance, the underlying concepts of “employment,” “production worker employment,” and “payroll/earnings” appear to be operationalized similarly for both the ASM and the CES, while “work hours” differs in terms of hours worked versus hours paid. However, our questionnaire review shows that the surveys convey this information to business respondents in different ways. On the ASM, the presentation strategy also changes between the form and its instructions. Do the different question and instruction strategies affect the reported figures? In this section we compare the four economic indicator variables across the ASM and the CES. We present our expectations for measurement error in terms of potential effects on variance and bias. Some of these expectations draw on the results of a small laboratory-based study in which participants completed forms from both surveys (Goldenberg et al., 2002).

**Total Employment.** On the surface, the “top-down” and “bottom-up” approaches identified in the questionnaire review would seem to require different cognitive response strategies. However, we know from prior studies that response strategies depend to some extent upon the structure of an organization and its payroll and personnel records. We have observed actual ASM respondents from establishments with complex organizational structures using a “top-down” approach instead of the “bottom-up” approach directed by the questionnaire. Likewise, some services industry CES respondents took a “bottom-up” approach (Tucker et al., 2000). We view both of these tactics as evidence of respondents searching for and exploiting “shortcuts” to provide requested data, which may be a business survey version of “satisficing”<sup>1</sup> (Krosnick and Alwin, 1987).

The variation in response strategies, regardless of the way questions are asked, contributes to a wider variance associated with summary statistics. To the extent that this variance is associated with the size and complexity of the establishment, there may be a systematic component to the variance. That is, if larger, complex establishments are more likely to use alternative response strategies and “satisfice” as compared with smaller, simpler establishments, we would expect the response variance to increase with establishment size and complexity. Both CES and ASM may be subject to increased response variance due to “satisficing.”

---

<sup>1</sup> “Satisficing” refers to the tendency of a respondent to do “enough” to answer a question, but not to put forth maximum effort (Krosnick and Alwin, 1987).

An additional source of error for the ASM comes from “Total Employment” being a derived, or artificial, figure. Although errors in respondents’ calculations to obtain the reported Total Employment figure, per the steps outlined on the form, can be corrected during automated batch editing of survey responses, errors in respondents’ judgments cannot. The Total Employment figure cannot be directly associated with actual employment figures in establishment records, such as end-of-year total employment; therefore, reporting errors may go undetected or considered unimportant by respondents. This may result in greater response variance in reported Total Employment figures for ASM as compared to CES.

**Production Worker Employment.** From the questionnaire review, we see that the two surveys count essentially the same categories of employees as production workers, although they use different formats for presenting that information. However, many employers do not summarize production workers in their payroll information (U.S. BLS 1983; Phipps 1989, Goldenberg 1994). Even if they do, the individuals labeled as production workers in records may not match the groups so defined by the survey. The questionnaire review highlights the need for respondents to make judgments about which employees to classify as production workers. Respondents must compare the functions or activities shown in the instructions with their knowledge of the individuals and the information in their records. However, respondents for the ASM or the CES are often accountants, bookkeepers, or payroll clerks, and may not be familiar with the activities of those workers. The result is that judgments identifying production workers may be based on respondent interpretations of the organizational structure or availability of records, rather than on job descriptions, known work activities, or the surveys’ include/exclude lists. Errors in include/exclude judgments lead to variation in the target data item. The ASM may be more prone to systematic error in this area because the include/exclude text in paragraph form is more difficult for respondents to process than similar information presented in list form.

These judgment errors have consequences for the resulting data. Respondents may base their calculations on departmental totals that are readily available in records (satisficing). If the totals include systematic errors, such as a department payroll total based on both production and nonproduction workers, the result is a bias. Similarly, if respondents make systematic errors in judgments of who to include or exclude as production workers (e.g., by incorrectly excluding one group more than they incorrectly include another), there is potential for bias in reporting. In both cases, the direction of bias depends on the number of employees in the incorrectly-handled categories.

Recall that CES obtains total employment and production worker employment, while for at least one quarter ASM respondents are asked for both production worker employment and “all other employees.” By asking for a complete accounting of all workers, we speculate that ASM calls attention to what does, and what does not, constitute a production worker. We think this specificity has the potential for more accurate reporting on the ASM, but only if the respondent considers individual employees and not entire departments.

**Production Worker Payroll and Hours.** Looking first at the payroll concept, we see from the questionnaire review that ASM payroll should be greater than or equal to CES payroll, because the ASM payroll definition is more comprehensive. The CES concept specifically excludes irregular payments such as bonuses. Turning to hours, we find that the concepts differ across the surveys. ASM requests hours worked, without paid leave, while CES collects hours paid, including paid leave. Thus, reported ASM hours should be lower than reported CES hours.

Judgments regarding payroll and hours should be relatively straightforward, because their components are typically well defined and consistently appear in business records. If we think of the employees as the rows of a spreadsheet, then the components of payroll (hours and various types of earnings) are the columns. To report employment and production worker counts, respondents determine which rows to include in the tally, working along only one dimension of the spreadsheet. Reporting payroll and hours is a two dimensional task, where the respondent first determines which rows (employees) to include, and then presumably holds these rows constant while selecting the appropriate columns (components) for reporting payroll and hours.

Clearly, if respondents define production workers incorrectly, the error will carry forward to their reporting of production worker payroll and hours. Reporting error could also occur if respondents do not hold the rows of the spreadsheet constant. Similarly, reporting error could occur if the content of the payroll and hours items (columns) differs from the data definitions requested on the survey form.

How does the spreadsheet analogy apply to payroll and hours on the two surveys? We believe that the potential exists for overreporting in CES, based on the exclusion of irregular payments. The payroll columns from which the monthly data are taken may include the additional payments, but not in a way that respondents can identify or remove them (Goldenberg and Stewart 1999).

Respondents could also select the wrong item (column) from their payroll records. This would result in incorrect reporting of the components of hours or payroll. Indeed, the lack of parallel construction in the payroll and hours items on the ASM questionnaire may

lead to systematic selection of the wrong column for hours data. ASM payroll includes pay for paid leave and vacation hours, whereas ASM production worker hours excludes these hours. Because payroll is asked before hours on the ASM form, the likely reporting error associated with this order effect is for “hours paid” to be reported rather than the requested “hours worked.” The result would be overreporting on ASM, and would cause ASM hours to approach CES hours paid.

#### 4.0 EVALUATION OF PUBLISHED DATA

Our questionnaire review has pointed to a number of expectations for the data. To the extent possible, we evaluate these expectations through an examination of published statistics from the two surveys. Because published aggregates cannot address our expectations or hypotheses regarding the variance component of potential measurement error, we limit our assessment to differences that may be indicative of potential bias.

##### 4.1 Methodology

Our analysis of published industry-level statistics requires comparable industry classifications across surveys. However, BLS and the Census Bureau converted from the Standard Industrial Classification (SIC) system to the North American Industry Classification System (NAICS) on different timetables. We based our analysis on SIC manufacturing industries for 1996, the last year for which published data can be directly compared.<sup>2</sup>

Official agency publications have different presentations for the variables of interest. The ASM publishes annual aggregates for total employment, production worker employment, hours worked by production workers, and production worker payroll. CES publishes monthly and annual average aggregate total employment and production worker employment. However, CES publishes production worker hours as average weekly hours per production worker, and payroll as average weekly earnings and average hourly earnings per production worker, instead of annual aggregates. In order to relate published statistics from the two surveys, we manipulated CES hours and earnings data to develop annual aggregates analogous to the ASM data for production worker hours and payroll.<sup>3</sup>

Our goal was to determine whether ASM and CES data were measuring the same phenomena. We assessed comparability relative to sampling errors associated with the difference between ASM and CES statistics for our four variables. We addressed the question: Is the difference between the survey estimates significantly different from zero? That is, are differences larger than an amount that could be attributed to sampling errors associated with the two surveys? If not, the consequences are not meaningful or substantive regardless of the size or type of measurement error, since the difference falls within deviations expected due to sampling variability. However, if the difference between the two survey estimates differs significantly from zero, then comparisons of the statistics and their supporting surveys are meaningful and worthy of further examination.

We compared estimates for the two surveys within 2-digit SIC manufacturing industry groups. More specifically, for each industry we tested the null hypothesis  $H_0: CES - ASM = 0$ . We established a 95 percent confidence interval around the difference between the CES and ASM estimates for each industry based on the standard error of the difference between the two surveys’ published estimates for that industry.<sup>4</sup> If the confidence interval includes zero, then the difference is not statistically different from zero at the  $\alpha=0.05$  level of confidence; that is, the evidence does not support a conclusion that the ASM and CES estimates are different. Thus, to the extent that these 95% confidence intervals do not include the zero point, the surveys are not comparable across industries.

In addition to comparing survey estimates within industries, we looked for a consistent pattern of differences between ASM and CES across the 20 two-digit SIC industries. After expressing ASM published estimates as a proportion of their counterpart CES figures, we subtracted this ratio from one. We then used the sign test (Snedecor and Cochran 1980) to assess whether these differences were in a consistent direction across industries (i.e., to test a null hypothesis  $H_0$  that there is no consistent direction to the differences). If we reject  $H_0$ , we find that there is a consistent direction to the differences across industries, suggesting an overall measurement effect. We compared the direction of those differences to our predictions based on the survey concepts presented above. We use results from the sign test, taken together with the results of the

---

<sup>2</sup> ASM began publication on a NAICS basis in 1998, while CES first published on a NAICS basis in June 2003.

<sup>3</sup> All CES data used in this analysis are annual averages for 1996, not seasonally adjusted, and were downloaded from the BLS web site ([http:// stats.bls.gov/ceshome.htm](http://stats.bls.gov/ceshome.htm)). To compare total employment, we used Census Bureau data from the 1996 ASM (U.S. Bureau of the Census, 1998) that include employment from auxiliary establishments, because the CES data do not differentiate between auxiliary and other units.

*(continued)*

---

Information on data manipulation procedures is available from the authors.

<sup>4</sup> Many of the standard errors had to be computed or interpolated from published data. Details are available from the authors.

significance tests, to judge whether the two surveys appear to be measuring the same concept.

## 4.2 Results

Table 2 summarizes the results of our analyses for each of the four economic indicators. Figures 1 through 4 illustrate them graphically. In the figures, each point represents a value for a different 2-digit SIC manufacturing industry, and shows the intersection of the ASM value (X axis) and the value for the difference between CES and ASM (Y axis). The bars surrounding the plotted points are 95 percent confidence intervals calculated using standard errors of differences between the ASM and CES survey estimates for the variable of interest. When the bars include the zero point the values are not statistically different.

**Total Employment.** Based on our concept review, we expect total employment to be approximately equal on the two surveys. In fact, as Figure 1 shows, 95% confidence intervals surrounding differences between CES and ASM include zero for more than half of the 20 SIC industries. In addition, results of the sign test fail to suggest that estimates from one survey are consistently higher or lower than estimates from the other survey across industries. Thus, we conclude that both ASM and CES appear to be measuring the same concept of Total Employment.

**Production Worker Employment.** Based on the concepts, we also expect production worker employment to be approximately equal across the two surveys. Compared to total employment, however, we find more variation among the differences between the two surveys' production worker estimates across industries (see Figure 2). Confidence intervals surrounding the differences between CES and ASM estimates of production worker employment include the zero point for only 5 of the 20 SIC industries. In addition, the sign test is significant (at  $\alpha = 0.05$ ), suggesting that ASM estimates of the number of production workers are consistently lower than the CES estimates. Thus, the ASM and CES do not appear to provide comparable measurements of the number of production workers across industries.

**Production Worker Payroll.** The concept review indicates that ASM captures a more inclusive payroll concept than CES, and therefore aggregate ASM payroll should be higher than aggregate CES payroll. Because the two surveys publish on a different basis, we derived a comparable annual aggregate figure for CES. Figure 3 shows that 95 percent confidence intervals around the difference between derived CES estimates and ASM include zero for only 7 of the 20 industries. Thus, as expected, derived estimates of aggregate CES production worker payroll do not appear to be measuring the same concept as ASM across industries. However, a statistically significant

sign test suggests that derived CES figures for aggregate Production Worker Payroll are consistently *higher* than ASM estimates across industries, which is the reverse of the expected direction. One possibility is that the higher value reflects the larger number of reported CES production workers.

We controlled for the difference in production worker estimates by computing average annual payroll *per production worker*. The sign test remains statistically significant, indicating that CES per-production-worker payroll is consistently higher than the counterpart ASM figure across industries. Since CES payroll is supposed to be *lower* than ASM payroll, we speculate that the consistently higher CES earnings per-production-worker may be evidence of measurement error. While this could be either overreporting on CES or underreporting on ASM, we believe the former is more likely—particularly if payroll records contain some of the irregular payments excluded by CES, and respondents are not able or willing to exclude them. Another possible explanation is that our data manipulation led to the reversal. Since the 95% confidence intervals for 18 of the 20 SIC industries include zero, on a per-production-worker basis the differences between CES and ASM payroll estimates are not statistically significant. However, the direction of the difference remains opposite from our expectation.

**Production Worker Hours.** By definition, the two surveys differ in their collection of production worker hours. CES includes hours of paid leave, and ASM excludes paid leave. Thus, reported ASM hours should be lower than hours reported on CES. Our derived annual aggregate for reported CES data appears to support our expectation for hours. The 95% confidence intervals around the differences between CES and ASM include zero in only 3 of the 20 SIC industries, suggesting, as expected, that the surveys are measuring different phenomena. The sign test provides further corroborating evidence of consistent differences in the correct direction, with reported ASM hours (worked) being less than reported CES hours (paid).

However, the differences in the hours concepts may not be the only reason for the observed discrepancy. To control for the conceptual differences, we adjusted the CES hours paid downward to derive an estimate of hours worked analogous to that collected by ASM.<sup>5</sup> Thus, the intent of this adjustment is to equalize the two survey measures. Figure 4 shows the differences between the ASM and the CES derived hours worked data, along with their 95% confidence intervals. Results of the sign test—that there is no

---

<sup>5</sup> We used 1996 ratios of hours worked to hours paid from the BLS Hours at Work Survey (U.S. BLS 2001, Table 1) to derive CES annual aggregate hours worked.

consistent direction to the differences—suggest that the direction of the differences between ASM and CES hours estimates has been removed by adjusting CES hours paid downward to reflect hours worked. However, confidence intervals around the differences between estimates for only 3 SIC industries include zero, indicating that differences remain in the hours measurements for many industries.

Since these industry differences may be related to the differences in Production Worker counts discovered earlier, we attempted to control for this effect by calculating annual hours worked per production worker. Table 2 shows that CES values for only 2 of the industries fell within the 95% confidence intervals around the differences between CES and ASM, a strong indication of survey differences in the concepts being measured. A statistically significant sign test tells us that derived hours worked per-production worker figures for ASM are consistently *higher* than their CES counterparts. Since our calculations were intended to adjust CES hours downward by excluding paid leave, the resulting consistently higher ASM figures may suggest that ASM respondents incorrectly included paid leave hours. If this were the case, however, we would have expected to observe it in the aggregate analysis shown in Figure 4. A more plausible explanation, paralleling that suggested earlier regarding our analysis of the payroll data, is a disproportionate effect of the adjustment for production worker counts.

**Summary.** Our evaluation of the published statistics leads us to conclude that for Total Employment, ASM and CES are both measuring the same phenomenon. However, results for the other three variables—Production Worker Employment, Payroll and Hours Worked/Paid—are less definitive. Any true reporting errors for payroll or hours data may be exacerbated by our manipulation of CES data to derive annual aggregates corresponding to published ASM figures, a calculation that required the use of production worker employment statistics. Our attempt to control for differences in the production worker counts on payroll and hours figures by expressing them per production worker only strengthens our suspicion that the production worker counts are plausible explanations for observed differences. Thus, we speculate that any over-arching measurement error may be associated with reporting of production worker employment.

## 5.0 DISCUSSION

In this paper, we compare four key economic indicators collected in two different establishment surveys, the U.S. Census Bureau's Annual Survey of Manufactures and the BLS Current Employment Statistics Survey. We examine Total Employment, Production Worker Employment, Production Worker Payroll and Production Worker Hours Worked/Paid

from the perspective of a questionnaire review delineating ways the measures differ, and by a statistical analysis of the differences in published survey data.

We discuss reasons for the observed differences in the published data from these two surveys, focusing on potential measurement error and speculating about the error sources and consequences. We recognize that some portion of the observed differences may be attributed to differences in survey procedures, especially sample design, estimation methods, coverage, non-response, use of benchmarking, or other survey features, as well as agency differences in sampling frames and industry classification rules. We also recognize that some of the differences we identify here could be artifacts of our data manipulation activities.

Nevertheless, there remain some discrepancies that may be attributed to measurement error. We speculate about how the findings might result from the differences we identified in the questionnaires and how those questionnaires might influence respondent strategies for answering the questions. In our opinion, the statistical analysis suggests that the most likely contributor to potential measurement error among these four variables is the collection of production worker employment. We speculate that reporting error is associated with a combination of factors. Respondents' handling of the definition of production workers communicated to them via instructions to include or exclude certain types of workers is probably one element. The nature of establishment records and their identification or compilation of production workers and associated data is another.

A comparison of published statistical data can only point out differences, not identify the source of those differences. A more informative study of measurement error would require a matched sample of respondents who participate in both ASM and CES. Such a study would allow us to examine reported data on both surveys for the same establishments, and to estimate both the variance and bias components of measurement error. We could then undertake qualitative research to investigate the reasons for these errors, and use the findings both to inform our knowledge about the quality of the data and to help both BLS and the Census Bureau to improve data collection forms and survey procedures. Recently-enacted legislation (the Confidential Information Protection and Statistical Efficiency Act of 2002) may enable studies of this type in ways that have not previously been possible. In the meantime, our methods—questionnaire review and statistical analyses comparing published data from the two surveys—are at least suggestive of potential measurement errors and sources that may warrant further investigation.

We wish to be clear that our purpose in conducting this research is not to pass judgment on

either survey, but rather to look at the surveys from the perspective of the respondent. Respondents from large businesses have observed that multiple government agencies ask for what seems to be the same information, but in different ways (Nichols and Willimack, 2001). Our questionnaire review provides examples of such differences.

More broadly, we feel that there is value in examining key economic variables across surveys, whether those surveys are conducted by the same organization or by different organizations. Nuances in the definitions of economic variables may or may not lead to meaningful differences in published statistics, but they do have consequences. Slight variations in definitions may translate into substantial response burden for diligent business respondents, and there may be a real cost associated with efforts to provide information for a survey. Alternatively, survey respondents might overlook the specifics of a definition, with a consequent effect on data quality.

We do not address the broader issue of redundancy in data collection. Clearly there are public policy reasons for this duality, given that CES is a closely-watched monthly indicator and the ASM provides an annual snapshot with a high level of industry detail. Indeed, David (2001) discusses the value of redundancy through the opportunities provided for improving data quality.

By their very nature, economic indicator variables have implications for policymakers, analysts, and other data users. If analysts expect the data items in both surveys to be the same, and treat variables as identical when in fact they are not, they may use and interpret the data in inappropriate ways. Considering the importance of these economic variables for monitoring and managing the economy, as well as the costs to businesses associated with statistical reporting, we believe there is merit in this kind of scrutiny.

## 6.0 REFERENCES

David, M.H. 2001. "Why Evaluate Business Data Collections?" *Proceedings of the Federal Committee on Statistical Methodology Research Conference*, Washington D.C.: U.S. Office of Management and Budget.

Fisher, S.K., Goldenberg, K.L., O'Brien, E., Tucker, C. and Willimack, D. 2001. "Measuring Employee Hours in Government Surveys." Paper prepared for presentation to the Federal Economic Statistics Advisory Committee, Washington, D.C., June.

Goldenberg, K.L. 1994. "Answering Questions, Questioning Answers: Evaluating Data Quality in an Establishment Survey." *Proceedings of the Section on Survey Research Methods, American Statistical Association*.

Goldenberg, K.L. and Stewart, J. 1999. "Earnings Concepts and Data Availability for the Current Employment Statistics Survey: Findings from cognitive interviews." *Proceedings of the Section on Survey Research Methods, American Statistical Association*.

Goldenberg, K.L., Willimack, D.K., Fisher, S.K., and Anderson, A.E. 2002. "Measuring Key Economic Indicators in U.S. Government Establishment Surveys." Presented at the International Conference on Improving Surveys, Copenhagen, August.

Krosnick, J.A. and Alwin, D. 1987. "An evaluation of a cognitive theory of response-order effects in survey measurement." *Public Opinion Quarterly* 51: 201-219.

Nichols, E. and Willimack, D.K. 2001. "Balancing Confidentiality and Burden Concerns in Censuses and Surveys of Large Businesses." *Proceedings of the Federal Committee on Statistical Methodology Research Conference*, Washington D.C.: U.S. Office of Management and Budget.

Phipps, P. 1989. "Non-Wage Cash Supplement Response Analysis Survey." Internal Memorandum, Bureau of Labor Statistics, November 1989.

Snedecor, G.W. and Cochran, W.G. 1980. *Statistical Methods*, 7<sup>th</sup> Edition, Ames: The Iowa State University Press, p.138.

Tucker, C., Fisher, S., Goldenberg, K., O'Brien, E., Rosen, R., Stinson, L., and Willimack, D. 2000. "Using Cognitive Methods to Improve Questionnaires for Establishment Surveys: A Demonstration." Paper prepared for presentation to the Federal Economic Statistics Advisory Committee, Washington, D.C., December.

U.S. Bureau of the Census. 1998. *1996 Annual Survey of Manufactures. M96(AS)-1. Statistics for Industry Groups and Industries*. Washington, D.C.

U.S. Bureau of Labor Statistics. 1983. *Employer Records Analysis Survey of 1981*. Unpublished report.

U.S. Bureau of Labor Statistics, Office of Productivity and Technology. 2001. "1999 Hours at Work Survey." Unpublished report.

U.S. Office of Management and Budget. 1978. *Statistical Policy Handbook*, Directive No. 12.

**Table 1. Summary of Differences Identified in Questionnaire Review**

<b>Concept</b>	<b>Annual Survey of Manufactures (ASM) Census Bureau</b>	<b>Current Establishment Statistics (CES) Survey (Manufacturing) - BLS</b>
<b>Total Employment</b>	Derived figure	Record-based figure
• Reference period	- Production workers (PW): Pay period including the 12th of March, May, August, November of the previous year - Other employees: pay period 12th of March	Monthly: pay period including the 12th of current or reference month
• Questionnaire	<b>Strategy: Collect employment subgroups and sum them</b> (“Bottom up”) - Asks for number of PW in a designated pay period each quarter - Directs respondent to sum quarterly PW figures and divide by 4 for annual average - Asks for All Other Employees for pay period including March 12 - Sum PW average and All Other Employees to get total	<b>Strategy: Start with total</b> (“Top down”) - Asks for total employment for pay period including 12th of reference month
• Instructions	<b>Separate instruction booklet</b> - Paragraph describes who to count (Employees reported on IRS form 941, Employer’s Quarterly Federal Tax Form Return)	<b>Instructions on back of form</b> - Bullet list of employees to include - Bullet list of employees to exclude
<b>Production Workers</b>	Derived figure (average)	Record-based figure
• Reference period	See Total Employment	Monthly: pay period including the 12th of the reference month
• Questionnaire	Part of total employment (“Bottom up”)	Form asks for number of employees who are production workers (“Top down”)
• Instructions	<b>Separate instruction booklet</b> (“Top down”; “Bottom up”) - Paragraph listing workers to include based on work functions; also specific “excludes” - Directs respondent to compute average number of PWs - Paragraph listing All Other Employees to Include	<b>Instructions on back of form</b> - Bullet list of occupations to include for all manufacturing industries - Bullet list showing managers and nonproduction occupations to exclude
<b>PW Payroll</b>	Sum of PW payroll and All Other Employee payroll	Record based figure
• Reference period	Annual payroll: Entire previous year	Monthly: pay period including the 12th of the reference month
• Questionnaire	- Form asks for PW wages, All Other salaries and wages - Total Payroll is sum of the two components	Form asks for Total production worker payroll, including overtime and excluding Lump Sum Payments
• Instructions	<b>Separate instruction booklet</b> - Begins with exclusion of fringe benefits cost - Report the payroll which was included on IRS form 941 - Paragraph listing elements of earnings to include - Include such deductions as...(list) - Paragraphs of employees whose earnings should/should not be included	<b>Instructions on back of form</b> - Enter total amount of pay earned during the reference pay period for all PW in column 3. - Report pay before deductions - Bullet list of earnings components to include - Bullet list of earnings components to exclude (irregular payments such as lump sums, bonuses not paid each pay period)
<b>PW Hours</b>	Record-based figure	Record-based figure
• Reference period	Annual, previous year	Monthly: pay period including the 12th
• Hours paid or hours worked	Hours <b>worked</b>	Hours <b>paid</b>
• Questionnaire	Plant hours worked by PW (Annual)	Total PW hours paid, including overtime, for the pay period that includes the 12th of the month
• Instructions	<b>Separate instruction booklet</b> - Includes all PW hours worked of paid for <i>except paid leave</i> - Includes overtime hours	<b>Instructions on back of form</b> - Hours Paid is the sum of: Hours worked, including overtime; Hours paid for portal-to-portal, standby, or reporting time; Hours of paid leave

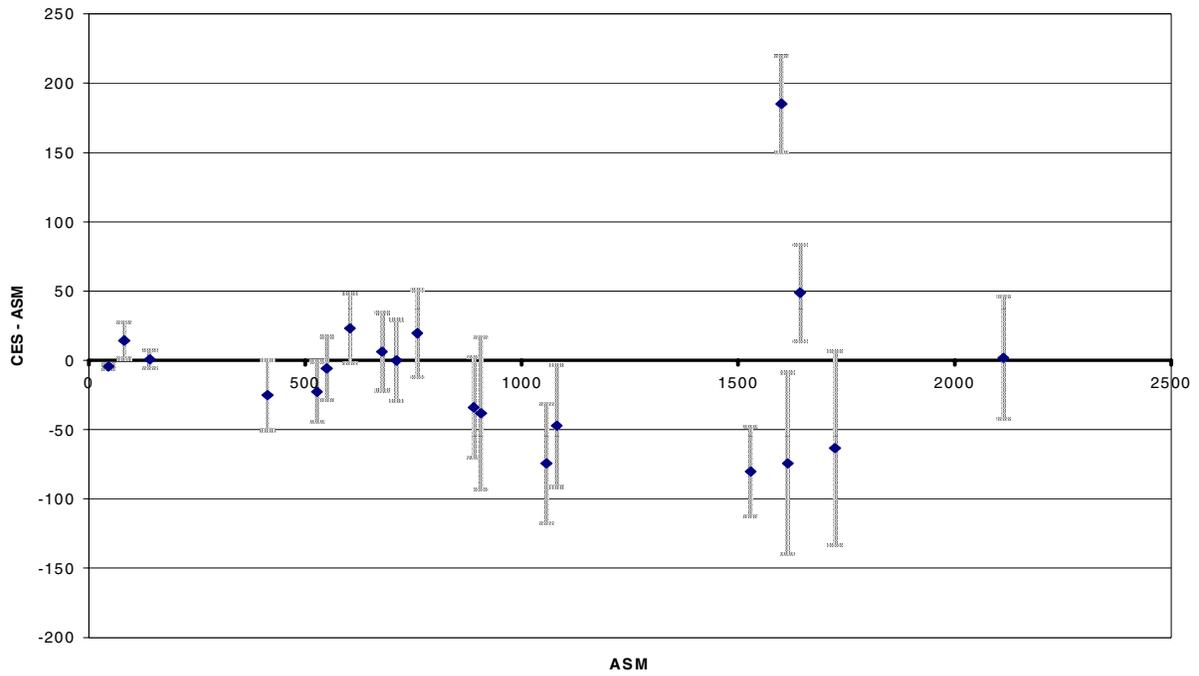
**Table 2. Comparison of Published Data: Summary of Results**

Concept/ Analysis	Expectation based on stated concept	Industries with 95% CES-ASM CI including zero point <sup>a</sup>	Sign test <sup>b</sup>	Direction of difference	Conclusion/Comments
Total employment	ASM = CES	11	Fail to reject $H_0$	- -	Probably same concept.
Production Worker (PW) employment	ASM = CES	5	Reject $H_0$ $\alpha = .05$	ASM < CES	Possibly not same concept.
Production Worker Payroll					
Aggregate PW payroll	ASM $\geq$ CES	7	Reject $H_0$ $\alpha = .05$	ASM < CES	Possibly not same concept; direction is counter to expectation.
Annual payroll per-PW	ASM $\geq$ CES	18	Reject $H_0$ $\alpha = .05$	ASM < CES	Controlling for the effect of PW count appears to eliminate the differences between the two surveys, but direction of results is counter to expectation. Could be due to measurement error in earnings data or could be a disproportionate effect of survey differences in PW counts.
Production Worker Hours					
Aggregate PW Hours using reported ASM hours worked and CES hours paid	ASM < CES	3	Reject $H_0$ $\alpha = .05$	ASM < CES	Consistent with expectations comparing Hours Paid to Hours Worked.
Aggregate PW Hours using "derived" CES hours worked	ASM = CES	3	Fail to reject $H_0$	- -	Adjusting CES hours to reflect hours worked controls the overall conceptual differences, but industry effects on the measurements remain.
Hours Per PW using "derived" CES hours worked	ASM = CES	2	Reject $\alpha = .05$	ASM > CES	Attempting to control for the effect of PW counts appears to exacerbate differences. Consistently higher ASM hours could be evidence of hypothesized question order effect. More likely due to disproportionate effect of differences in PW counts.

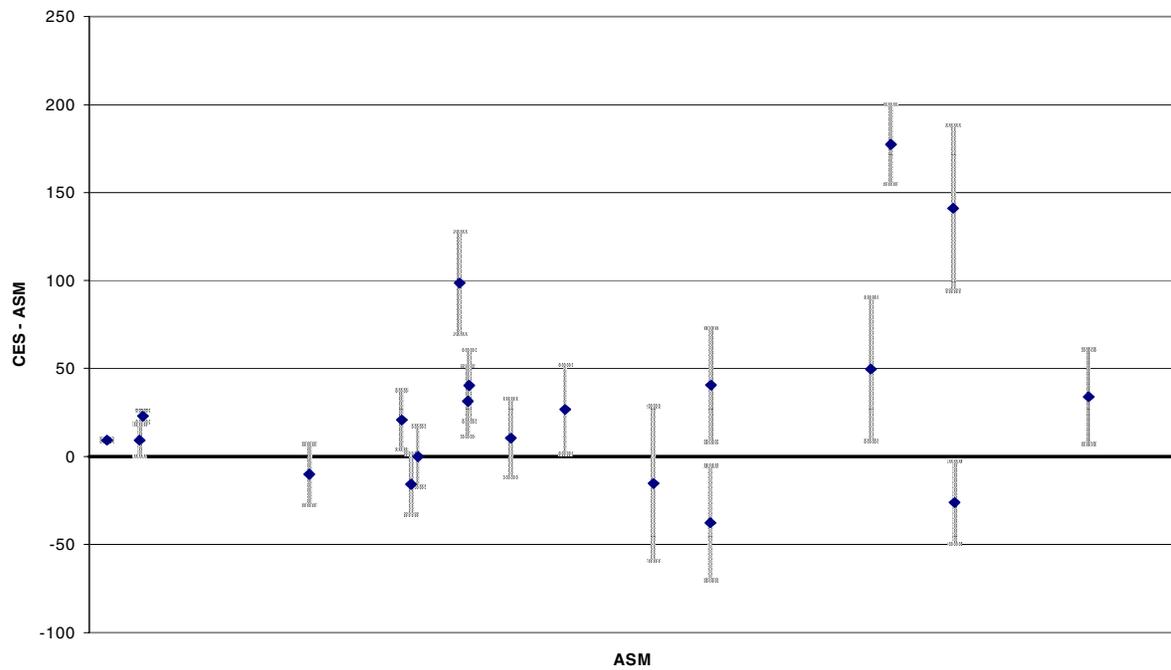
<sup>a</sup> Testing  $H_0$ : ASM = CES based on standard errors of differences between ASM and CES estimates.

<sup>b</sup>  $H_0$ : There is no consistent direction to the differences between the survey estimates. Reject  $H_0$  at  $\alpha = .05$  if the smaller of the two values is 5 or less.

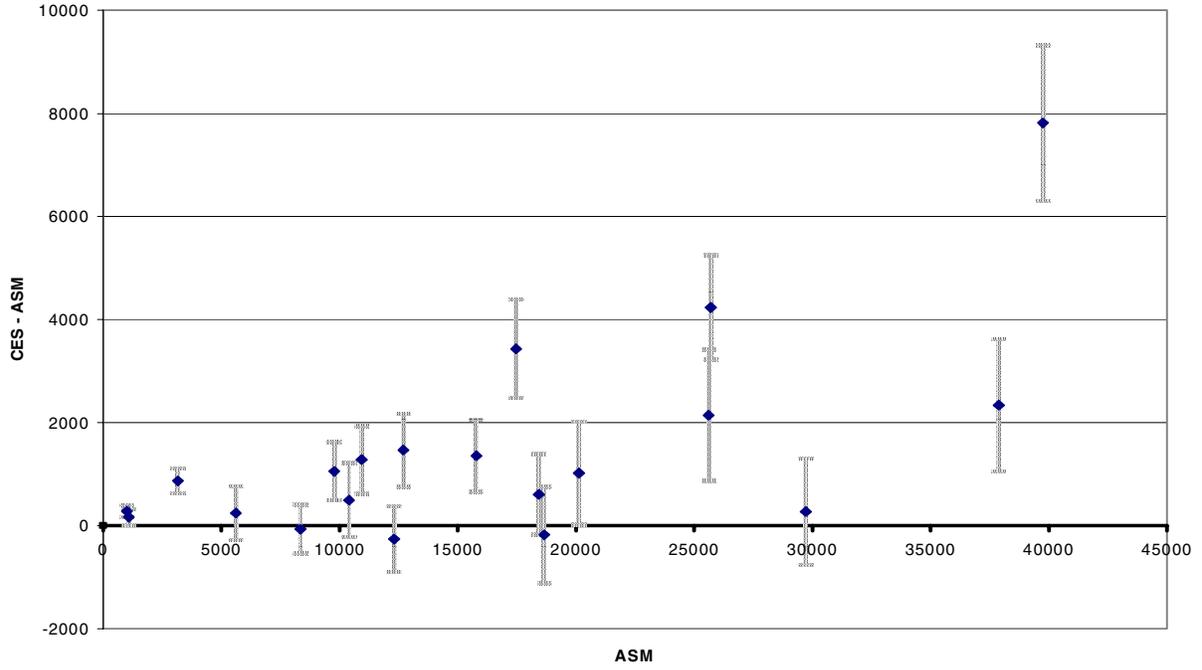
**Figure 1. Total Employment (thousands) for 2-Digit SIC Manufacturing Industries  
(Differences between CES and ASM, along with their 95% Confidence Intervals )**



**Figure 2. Production Worker Employment (thousands)  
for 2-Digit SIC Manufacturing Industries  
(Differences between CES and ASM, along with their 95% confidence intervals)**



**Figure 3. Aggregate Annual Production Worker Wages (millions)  
for 2-Digit SIC Manufacturing Industries  
(Differences between CES and ASM, along with their 95% confidence intervals )**



**Figure 4. Aggregate Annual Production Worker Hours Worked (millions)  
for 2-Digit SIC Manufacturing Industries  
(Differences between CES and ASM, along with their 95% confidence intervals )**

