MANUFACTURERS' SHIPMENTS, INVENTORIES, AND ORDERS SURVEY

Joanne M. Katz
Assistant Chief, Economic Indicator Programs
Industry Division
Bureau of the Census

Friday, October 9, 1987

Manufacturers' Shipments, Inventories, and Orders Survey

A. INTRODUCTION

The manufacturers' Shipments, Inventories, and Orders (M3) survey is one of the Census Bureau's most important monthly surveys. It is also a difficult survey to conduct. The survey has response and other problems that have no easy solutions. The purpose of this paper is to provide a current status report on the survey and to seek advice from the Census Advisory Committees about some of the issues we are considering.

The M3 survey covers the entire manufacturing sector of the economy and is intended to provide broad measures of month-to-month change in manufacturing. The data are widely used by Government and industry. The new orders data provide components of index of leading indicators; the shipments and inventory data are used in the gross national product calculations.

B. BACKGROUND

The Manufacturers' Shipments, Inventories, and Orders (M3) survey was started in 1939 by the Office of Business Economics (OBE) to provide monthly estimates of change in manufacturing. The monthly Industry Survey conducted by OBE collected data at the total company level from a relatively small group of companies. Initially, OBE published only index numbers but later published dollar aggregates by benchmarking the monthly series to Statistics of Income data from the Internal Revenue Service.

In 1957, responsibility for conducting the monthly survey was transferred to the Census Bureau. Improvements were needed in the data and the Census Bureau was able to develop more detailed monthly data by increasing the sample size, by introducing divisional reporting from large, diversified companies, and by benchmarking the survey to the annual survey of manufactures.

Data from the M3 survey are widely used by Government, business, and academic users. Among the most important uses of the M3 data are the Bureau of Economic Analysis (BEA) measurements of current investment in capital goods and estimates of the change in business inventories components of the gross national product. The data on new orders for durable goods provide one of the 12 series in the index of leading economic indicators. Another component of the index of leading indicators is net change in inventories which is developed by BEA using the M3 data for manufacturing and the monthly retail and wholesale survey data for trade inventories.

Each month, The Census Bureau publishes two M3 reports: an advance report and a preliminary report. The advance report contains data on shipments, new orders, and unfilled orders for the durable goods industries in total and for selected major industry groups. The preliminary report contains revised, more detailed shipments and orders data for the durable goods industries and the first estimates of shipments and orders for the nondurable

goods industries. The preliminary report also contains the first estimates of inventories for all manufacturing industries and final revised data for all series for the previous month.

The advance estimates are provided to meet the need for early economic indicator data. However, since data are not available from many companies early enough for use in the advance estimates, these data are based on a limited number of reports, primarily from very large companies.

The preliminary report is based on more complete reporting than the advance but with reports from some companies still missing. Data reported by companies that cannot respond in time for the preliminary estimates are included in the final estimates that are published as revised data the following month.

The monthly M3 estimates for shipments (sum of 12 months) and inventories (end of year) are benchmarked to totals from the annual survey of manufactures (ASM), except in economic census years when they are benchmarked to the census of manufactures. The latest benchmark was to the 1982 census and 1983-85 ASM and was released in March 1987.

1. The Survey Panel - The present M3 survey panel consists of approximately 4,100 reporting units. The original panel included virtually all companies with 1,000 or more employees and a sample of the smaller companies. Over time, we have deleted companies that refused to report from the mailing panel. At present, most of the companies with 2,500 or more employees, about half of those with 1,000-2,499 employees, and less than one-third of the sample selected for the 100-999 employee group are reporting in the survey. There is no representation of companies with fewer than 100 employees.

   Each year we prepare reconciliation listings, comparing the M3 data to the ASM data. We have used these listings as the basis for updating the M3 panel. The ASM is based on a well maintained sample panel which includes representation of new manufacturing establishments opened after selection of the sample as well as follow up on sampled establishments that say they have ceased operations.

   The reporting unit in the M3 survey typically comprises the entire operation of companies that operate in a single M3 industry category. Most of the large diversified companies file separate divisional reports, but the divisional structure of some companies does not provide adequate detail for the M3 industry categories. Thus, many of the reporting units include mixed industry activity even given the broad industry categories of the monthly survey.

   The survey methodology assumes that the month-to-month changes of the units reporting in each industry category represent effectively the month-to-month movements of the establishments in the Standard Industrial Classification (SIC) industries that make up that category.

2. <u>Monthly Estimation Procedure</u> - The emphasis in the survey is on percent change from the previous month rather than on estimating levels for the current month. The monthly estimates of shipments, unfilled orders, and total inventories are ratio estimates, derived for each industry category by multiplying the industry estimate for the previous month by the percentage change from the previous month for companies reporting in the current month. New orders, however, are not calculated according to the standard ratio-estimate procedure. New orders are computed by adjusting the current month's shipments by the change in the backlog of orders. (New orders equal current month unfilled orders minus prior month unfilled orders plus current month shipments.) We do collect new orders in the survey from companies that can report it and use it for data editing.

3. <u>Response</u> - Historically, M3 response from large companies has been better than from smaller companies. The initial sample for the survey included companies of all sizes, but we decided several years ago to discontinue collecting data from companies with fewer than 100 employees. This decision was reached because of poor response from small companies combined with their small contribution to the results. Response from a supplemental sample of companies with 100-999 employees was about 60 percent when this sample was introduced in 1978, but is now only about 30 percent. Given our budgetary constraints and the importance of large companies in manufacturing, we have emphasized maintaining and improving response from the largest manufacturing companies.

   There are significant large companies that do not respond and others that respond but do not report adequate divisional detail. In addition, many companies respond too late to be included in our early estimates. The primary difference between our advance and preliminary estimates is caused by late reporting by a few large durable goods companies.

   Despite programs aimed at improving the M3 survey estimates, the proportion of value of shipments of M3 respondents, as compared to the manufacturing universe, has been declining. During the last 10 years, the percent of reported shipments data to the total has declined from 54 percent in 1975 to 49 percent in 1985, a decline in coverage of 9.3 percent.

4. <u>Benchmarking</u>

   a. <u>Shipments and Inventories</u> - We benchmark the estimates of annual shipments and end-of-year inventories from the monthly survey to the annual totals compiled in the census of manufactures and in the ASM as those data become available.

Our current procedure for benchmarking includes comparing M3 data to ASM data as reported by companies in each survey and attempting to resolve large differences by telephone contact with the company and research into publicly available sources, such as stockholders' reports.

The monthly M3 survey is based on company or divisional level data. Each reporting unit (entire company or division of a company) in the M3 survey is classified into 1 of 74 industry categories based on the major activity of the reporting unit. Five of the 74 industry categories are further classified into defense and nondefense components.

The ASM and census of manufactures are based on establishment reporting with each establishment classified into one of 449 4-digit SIC industries based on its product shipments. The 74 M3 industry categories represent combinations of the 4-digit SIC industries.

As indicated above, the M3 survey methodology assumes that the month-to-month changes for reporting units classified in each industry category effectively represent the month-to-month movements of the establishments which make up the category. The benchmarking adjusts for differences that result from collecting the monthly data on a divisional or company basis as compared to the establishment basis used in both the ASM and the census of manufactures. It also corrects for errors resulting from the relatively small monthly survey and for any reporting or processing errors that may not have been detected as part of the review of the monthly estimates. The benchmark procedures adjust the monthly estimates to the annual totals for shipments and to the December year-end figures for inventories by minimizing the revision to the month-to-month percent changes in the original data.

Reconciliation of the M3 data with the ASM for the years 1979-83 was impaired by the change to an establishment-based sample for the ASM in 1979. Although this reduced the size of the ASM sample and, thus, the cost of conducting the survey and the reporting burden, the establishment sample made reconciliation of the ASM with the M3 difficult, if not impossible. Prior to 1979, the ASM sample had been company-based and included all establishments of any company that had at least one establishment selected for the sample. With the company-based sample, reconciliation of the M3 data with the ASM data had been performed annually and differences between the surveys reviewed and resolved when possible. Because of the M3 need for annual company totals and because potential users of the Longitudinal Establishment Data file expressed a strong need for data for an entire firm for many kinds of studies, the ASM sample was changed in 1984. The new ASM sample for 1984, while still essentially an establishment sample, includes with certainty all

establishments of the largest companies (those with shipments of $500 million or more in 1982) and all other establishments with 250 or more employees regardless of company size; in addition, there is a probability sample of the remaining establishments. The ASM certainty companies represent about 60 percent of the total value of shipments of manufacturing establishments. For those very large companies, we once again have more comparable data to reconcile the M3 data to the ASM. This improved reconciliation of the data reported in the M3 and the ASM should result in smaller benchmark revisions.

b. Unfilled Orders - We have not had annual benchmark data for unfilled orders from a larger scale survey as for shipments and inventories. We have made periodic adjustments to the M3 unfilled orders data, but these adjustments have been based primarily on revised data reported in the monthly survey and supplemented by single-establishment company data from the ASM. We do not collect unfilled orders data from multiestablishment companies in the ASM because a 1976 recordkeeping practices survey indicated that many large companies do not maintain orders records at the establishment level.

We are conducting a benchmark survey to collect unfilled orders data as of the end of 1986 from multiestablishment companies. We will use the results of this benchmark survey with the single establishment company data from the ASM to set the level of unfilled orders as of the end of 1986. The benchmarked levels of unfilled orders will be introduced at the same time that we benchmark to the 1986 ASM, in early 1988.

Since the data we publish for new orders are calculated from shipments and unfilled orders, no benchmark is needed for this item.

5. Importance of a Few Companies in the New Orders Data - There are very large fluctuations in new orders from month to month in some industries, notably complete aircraft, shipbuilding, and a few others. These fluctuations may be caused by one company or group of companies one month and another company or group the next month. The companies that report in time for the advance estimates may not be representative of the late reporters for any month. We have not been able to find any statistically acceptable method of estimating sporadic new orders without data from the companies that received those orders. A single company that had received new orders of only 100 million dollars 1 month might have more than 1 billion dollars (about 1 percent of the total new orders for all durable goods) the next month. We cannot tell which companies or how many companies might have unusually large or small orders in any given month. We follow announcements of large orders in the press but have found that they frequently are announced before the orders are in company accounts. We also follow the Department of Defense (DOD) contract awards

but find this information of limited use because of timing differences, differences between the amounts given in the DOD announcements and what is in the company books, and problems with reporting of multiyear awards. This is our most serious imputation problem and one for which we do not see any solution other than getting the reported data. There are certain companies from which it is essential that we have data before we can develop industry estimates of new orders.

6. Concurrent Seasonal Adjustment - We implemented concurrent seasonal adjustment effective with the February 1987 reports. Concurrent seasonal adjustment has been determined superior to the previous methodology which used forecasted factors from the X-11 program. With concurrent adjustment, all available data including the current month's data are used in the calculation of seasonal adjustment factors for the current month. This procedure in most cases results in smaller revisions to the historical seasonal factors as subsequent observations are added to the series.

With the calculation of a unique seasonal factor for a series for a given current month, all seasonal factors for prior observations for the series are also revised. Generally these revisions are not significant. In order to minimize the inconvenience to data users caused by revising all historical data points every month, we revise only the prior month and the current month year ago in each monthly publication. This procedure gives a more accurate current month-to-month change and year-to-year change. While concurrent seasonal adjustment means the historical data in this publication may be revised each month, a full set of revisions will be published only in the annual publications.

At the same time that we implemented concurrent seasonal adjustment, we discontinued the seasonal adjustment of series which do not have identifiable seasonal patterns. Tests using summary statistics from the X-11 ARIMA seasonal adjustment program and from sliding spans analysis were used in examining all M3 series for seasonality. For those series determined to have no identifiable seasonal pattern, the not seasonally adjusted data are included in both the adjusted and unadjusted industry totals.

C. MAJOR ISSUES

1. How to Measure the Quality of the M3 Survey - In measuring the quality of the M3 estimates, a primary issue is how important it is to have a probability sample. Some question the quality of the M3 survey because of the lack of a probability sample and related measures of sampling error. If we do not have a probability sample, what is the best way to measure mean square error? Some initial work on model-based variances is being done within Industry Division but results are still some time away.

In 1986, we completed an evaluation of the M3 survey. This evaluation was completed as part of an effort to meet the requirements of OMB's Statistical Policy Directive No. 3, "Compilation, Release, and Evaluation of Principal Federal Economic Indicators." We decided to judge the quality of the monthly estimates by measuring the magnitude of relative revisions. If revisions were small, quality was judged to be high; if revisions were large, quality was considered low. We wanted to examine the accuracy of the early estimates. We also wanted to examine the closeness of the total annual shipments and end-of-year inventories to the ASM and census benchmark totals. A description of the statistical tests we completed to meet OMB's requirement for an assessment of the survey, the criteria used and a summary of the results are given in the appendix of this paper.

Our examination of the accuracy of the early estimates showed a statistically significant understatement of the advance estimates of new orders for total durable goods and for defense capital goods. The tests did not identify any serious problems in the advance estimates for shipments or unfilled orders. The tests on the preliminary estimates also did not show any problems that we considered significant.

The tests to measure the effect of benchmark revisions across all industries for the years 1977 through 1982 indicated that we have had statistically significant revisions in census years. For the ASM years (other than 1978, which was benchmarked at the same time as the 1977 Census and, thus was revised proportionally to the 1977 revision), the revisions were not statistically significant.

We also examined each M3 industry category and higher level aggregate for average annual benchmark revisions. The 2-digit SIC group with the largest revisions was Group 30, "Rubber and Plastics Products." In addition, since this is a very large industry group in terms of number of companies, we decided to select a probability sample of companies classified in this group for evaluation of several questions regarding probability sampling for the M3 survey. Selection of the sample is underway and we plan to mail it and use it as a separate panel beginning with the January 1988 data month. Among the issues we plan to study are problems in getting response from small companies, alternative kinds of follow-up procedures, and the difference in estimates between the current panel methods and the probability sample caused by nonresponse.

2. Coverage and Response - A second major issue is coverage and response and how they contribute to the quality of the estimates. Therefore, in addition to examining the revisions to the M3 estimates, we wanted to examine the adequacy of the M3 panel and of the reporting panel in representing the manufacturing universe. To do this, we used estimates developed from the ASM. We matched the M3 panel to the ASM panel for the year 1984 (the most recent ASM data available at the time). In this

way, we identified the ASM establishments of companies in the M3 panel.
These establishments were further classified as M3 reporters or nonre-
porters. Then, using establishment level data, we estimated the 1982
to 1984 rates of change for the M3 panel, the M3 reporting panel, and
the manufacturing universe as measured by the ASM panel. We found that
14 of the 74 M3 industry categories had panel problems or both panel and
reporting problems. In addition, there were 22 industry categories for
which the panel was adequate but that portion of the panel reporting in
the M3 survey was not representative. (See appendix item 2b for further
explanation of this test and criteria used to define problem panels.)

Our conclusion from the evaluation studies was that the major need for
improving the M3 survey is for an improvement in response. We believe
that the most cost-effective way to increase response in the M3 survey
would be to make the survey mandatory. In 1983 we developed a proposal
requesting mandatory reporting authority. The OMB rejected this request,
expressing doubts that mandatory reporting would achieve the desired
results and suggesting that we try other approaches to increasing
response.

For several years, we have been trying to improve response through
company visits, correspondence, and increased telephone follow-up.
In 1978-82 we contacted nonrespondent companies with 1,000 or more
employees through telephone calls, letters, and visits to encourage
response. This was discontinued in 1982 when we were not achieving
results.

In 1983, the M3 staff began an evaluation of coverage in each of the
industry categories tabulated in the M3 survey. We identified for every
category large nonreporting companies, companies not providing comparable
data between the M3 survey and its ASM benchmark, and industries which
were comprised for the most part by small companies.

Beginning in August 1984, we started another company visit program with
emphasis on the largest nonresponding companies and companies with
reporting problems. We also have increased our telephone follow-up and
correspondence with companies.

Our concerns about nonresponse have been one of the major factors in
our decision to continue without a probability sample for the M3 survey.
Without good response from most small companies in the panel, the use of
a probability sample raises questions on how to impute for these cases
and how to deal with highly volatile reporting from a few small reporters
with large weights. We are not convinced that using a probability sample
will necessarily improve the quality of the M3 estimates.

3. Divisional Reporting - In addition to the problem of nonresponse, there
   are companies that report in the survey but do not report adequate
   industry detail. As stated above, the OMB evaluation study showed that

22 of the 74 M3 industry categories had inadequate reporting panels.
Part of the inadequate reporting panel problem is due to inadequate
industry detail. The M3 reports are generally based on data that
responding companies compile monthly for their internal purposes and,
thus, usually are for an entire company or divisions of a company. Many
reporting units include activities that belong in another industry cate-
gory. Some companies are, therefore, overrepresented in some industry
categories and not represented in others for which they have activities.
At the all manufacturing level, companies with value of shipments repre-
senting 10.7 percent of the total shipments reported their M3 shipments
in a different industry category than they should have, based on their
SIC classification in the 1984 ASM.

While we have had some success in negotiating improved divisional
reporting in the company visits, this is a very costly and inefficient
way to proceed. Mandatory reporting would provide more and better data
at lower cost. This also would increase the reporting burden but we
believe that most companies have monthly data for broad categories of
products and just are not willing to put much effort into a voluntary
survey.

Our research plans include investigating alternatives for allocating
the monthly reports that cover more than one industry group. We plan an
initial investigation using the probability sample for the "Rubber and
Plastics" industry and the most recent ASM data. Allocation of M3
reports based on the ASM is complicated by the fact that the surveys are
collected on different bases: the monthly data are generally net sales,
excluding interplant transfers; the annual data are establishment ship-
ments, including interplant transfers. There can be a very significant
difference between net sales and plant shipments for companies with large
amounts of interplant transfers. We have a considerable amount of
research to do before we can determine if suitable methodology can be
developed for allocation of M3 reports to more detailed categories.

Another divisional reporting issue requiring research is how to develop
a frame for probability sampling based on divisional reporting. We have
the files needed for selection of samples at the establishment level or
at the company level but not at the divisional level. The M3 divisional
reports have been negotiated individually with the large companies, to
fit data available from company records as closely as possible to the
M3 industry categories. Much work remains to be done, including exami-
nation of the structure of large, complex companies in developing a
divisional or segment reporting system. This is related to our review
of the M3 levels of detail, which is discussed in 4b below.

4. Other Issues

    a. Keep or Discontinue the Advance Report - We are considering
       discontinuing publication of the advance report. The arguments in
       favor of discontinuing are that it is released only 5 working days

before the regular report, is subject to occasional significant
revisions when companies with unusually large orders in a month do
not report in time for inclusion in the advance estimates, and there
would be some cost savings. While we have had some success in
obtaining earlier data from companies, it is too soon to determine
if the bias in the estimates we identified in the OMB evaluation
study has been removed. The arguments against discontinuing are
related to user needs for early data on the manufacturing sector of
the economy including use of the data in the index of leading
indicators.

b. Levels of Detail - At present, we tabulate data in the M3 survey for
74 industry categories, with 5 of these further classified into
defense and nondefense components. The current M3 categories have
not been changed for many years. Many of the existing M3 categories
are distinct sectors of the manufacturing universe for which we
believe users need separate monthly data. Some of the existing
categories are relatively small and are tabulated separately pri-
marily to permit aggregation of the M3 data into market categories.

We plan a review of the existing M3 industry categories with the
aim of reducing the number of separate categories. This is part of
a larger project for development of a reporting system for the
collection of economic data based on segments of business. We will
continue visiting large, complex companies in our M3 company visit
program, gathering information on data available monthly from com-
pany records. We also are developing plans for a comprehensive
recordkeeping survey of large companies before the 1992 Census of
Manufactures to determine what data are available at what levels in
these companies.

c. Inventory Data - The inventory data collected in the M3 survey prior
to 1987 were defined as book value inventories. This resulted in a
mix of valuation methods, first-in-first-out (FIFO), last-in-
first-out (LIFO), average cost, market, etc., in the inventory
estimates. In 1987, we started collecting and publishing the M3
inventory data on a pre-LIFO basis, consistent with the data collected
in the ASM.

In the benchmark report issued in March 1987, the monthly M3
inventory data, which were collected using the book value definition,
were redefined and benchmarked to the census and ASM data collected
on a current cost or pre-LIFO basis. This change in definition
resulted in larger than usual revisions for 1982-86.

5. Summary - All of the issues discussed above are intertwined. If we need
a probability sample to measure quality of the estimates, then we need
improved response. To obtain sufficiently improved response to go to a
full probability sample, we need mandatory reporting. A major part of

the need for improved response is for better divisional reporting from
many large companies, which we believe we can obtain only if the survey
is mandatory. We also need additional funding to introduce, implement,
and maintain a probability sample.

D. ISSUES PROPOSED FOR COMMITTEE DISCUSSION

1. Assessment of Quality - We would like to have comments on the approach
    we used in the OMB evaluation of the survey and suggestions for other
    approaches to measuring quality. How useful are the M3 estimates?

2. Probability Sample - How important is it to have a probability sample
    for this survey? Significant additional funding, probably a 40-50 percent
    increase in the present budget, would be required for implementation,
    coverage maintenance, and follow-up of a probability sample. Our past
    experience has been that small companies either do not report or report
    poor quality data.

3. Mandatory Reporting - Would the committees support mandatory reporting?
    As discussed above, we believe that we may obtain poorer results than
    we have now if we go to a full probability sample without mandatory
    reporting.

4. Divisional Reporting - How much industry detail is needed on a monthly
    basis? Would 2-digit SIC level data be sufficient? Are the market
    category data needed? Do the committees have any comments or suggestions
    on the use of divisional or company level data to estimate data at the
    establishment level? Should we consider again a different annual bench-
    mark for the M3 survey, one based on divisional level data? This, also,
    would require funding and has been rejected in the past. If we were to
    establish an annual benchmark based on divisional reporting, how would
    we handle classification problems resulting from company reorganizations,
    mergers, etc.?

5. Publications - Do the committees have any suggestions about the data
    published in the current M3 publications? Given the volatility of some
    of the new orders series, should we consider publishing moving averages?
    Now that we are using concurrent seasonal adjustment, how far back
    should we revise in the monthly publications? Is our current practice of
    revising the prior month and year-ago-current month acceptable?

6. Other - Are there any other issues the committees would like to discuss?

APPENDIX

OMB Evaluation Study

The study examined revisions caused by late response and benchmarking. In all computations, we used the data before seasonal adjustment. We excluded revisions due to seasonal adjustment because we are using the widely accepted X-11 method for seasonal adjustment and any problems related to this method would not be unique to the M3 survey.

1. Revisions Caused by Late Response - Revisions from the advance to preliminary to final estimates for any 1 month are caused primarily by late response from companies, although occasionally companies report estimates for the advance report and then later revise those estimates when actual data are available.

   To test for significant revisions in the progressive estimates of the monthly data, we assembled information from January 1981 through June 1985 for all M3 industry categories and aggregates for each publication cycle. Although we selected this time period primarily based on data availability, we believe it is sufficient for determining the effects of late response on the estimates. This time period is both recent and provides enough obser- vations to make meaningful statistical tests.

   We computed the percent revision and the dollar revision between the advance or preliminary and final estimates for each tabulation category and aggre- gate for each month. We computed the mean revision and used a t-statistic to determine if that mean revision differed significantly from 0. In addition we computed the mean ratio of change from the previous month, the mean revision to that ratio of change, and tested the significance of the dif- ference between the means.

   All of these computations were made for each tabulation and publication level over the months January 1981 through June 1985. All variables were tested, i.e., shipments, new orders, unfilled orders, and inventories for all industry categories both for predominance of underestimation or over- estimation and for significant revisions that were not consistently revised in the same direction.

   a. Tests for Systematic Underestimation or Overestimation

      (1) Advance Report - These tests indicated that the amount of error in the new orders estimates for total durable goods was significant at the 95 percent confidence level. This error was predominantly an underestimation.

          In examining the results of the tests for the specific M3 industry categories, we determined that the capital goods series are the primary sources of the error, specifically the aircraft, aircraft parts, and communication equipment series. For these industries

new orders are quite irregular, frequently very large and highly unpredictable. Companies reporting in the advance report are not always indicative of companies that report later; historical data also are not reliable as estimators.

For the shipments and unfilled orders data, the error level was not significant in any of the advance publication levels, using the same criteria as for new orders.

(2) Preliminary Report - Revisions caused by late reporting to shipments, new orders, and unfilled orders published in the preliminary report generally were not significant, although there were a few series with significant error, using the same criteria as for the advance report.

We found a small but statistically significant error in inventories at the all manufacturing and durable goods total levels for month-to-month dollar change. The error was predominantly an overestimation of the preliminary data. Again, only a few specific industries were involved.

b. Tests for Magnitude of Revisions - We developed a test of absolute revisions. For this test, significant error is indicated by a t-value of +2.0 or greater. We tested the advance publication levels against an average revision rate of 1/2 the average month-to-month change and of 1/4 the average month-to-month change. We believe that the 1/4 test may be more appropriate for aggregate series (2-digit SIC and higher) while the 1/2 test is appropriate for the M3 industry category levels.

(1) Advance Report - The advance publication levels showed no significant error when tested against an average revision rate of one-half the average month-to-month change.

(2) Preliminary Report - The publication levels in the preliminary reports showed no significant revisions when tested against an average revision rate of one-half the average month-to-month change.

2. Revisions Caused by Benchmarking - Revisions from the final estimates for any month to the benchmark estimates can result from any or all of the following:

Sample deficiencies - The sample may not be representative of the manufacturing universe.

Nonresponse - The companies in the sample that do not respond in the survey may be needed to make the monthly estimates representative of the manufacturing universe.

Very late response or revisions to previously reported data - Data that we receive after publication of the final report for the month.

Inadequate divisional reporting - The M3 reports are generally based on data that responding companies compile monthly for their internal purposes and, thus, usually are for an entire company or divisions of a company. Many reporting units include industry activities in the M3 industry category in which they are reporting when, in fact, some of those activities belong in another category.

Accounting differences between the monthly company or divisional records and the annual establishment records:

>Monthly Sales versus Annual Shipments - The monthly company or divisional M3 reports generally contain net sales data (excluding interplant transfers); the annual ASM reports are for establishment shipments (which include interplant transfers). There are significant differences in the data levels between the ASM and M3 reports. In some industries, there is a large amount of captive production that may be reported as shipments from one plant of a company to another in the ASM but never shows up as sales in the M3.

>Inventories - The monthly M3 reports may include sales branch inventories that are not in the ASM. In addition, there are differences in inventory valuation methods between the M3 reports from company or division level and ASM reports from establishments.

Revisions based on the M3/ASM reconciliation (that is, comparing the sum of the monthly M3 data to the annual totals from the ASM by company) - In some cases, companies have revised their monthly M3 data in response to questions about the differences between the M3 and ASM reports.

The extent to which the ASM underestimates the manufacturing universe

We developed two studies to examine the revisions caused by benchmarking. The first of these involved statistical tests on the benchmark revisions. The second was a study examining sample deficiencies and nonresponse.

a. Tests on Benchmark Revisions - To test for significant revisions to the estimates from the monthly survey as compared to the benchmark data, we performed additional statistical tests.

Data used were shipments and inventories for all industry categories for the years 1977 through 1982. Although we had not published a benchmark of the M3 to the 1982 census data at the time of the evaluation, the census data were available and were included in the study. The 1983-85 ASM data were not included.

We conducted t-tests on percent revisions for benchmark data across all of the 74 M3 industry groups. These tests indicated significant problems for the years 1977, 1978, and 1982. Since the M3 survey was benchmarked to the 1978 ASM at the same time as to the 1977 Census of Manufactures, the M3 understatement of the 1977 census totals also was reflected in

the originally published M3 estimates of levels for 1978. That is why the 1978 benchmark revisions were also significant. Except for 1978, the M3 estimates were not revised significantly when benchmarked annually to the ASM.

These tests show that the most serious M3 problems are in the years when the survey is benchmarked to the census of manufactures. Therefore, part of the M3 estimation problem is related to the ASM sample deterioriation and under statement of the universe over this time period.

b. <u>Tests to Measure Sample Deficiencies and Nonresponse</u> - The next part of our study used the ASM to examine the adequacy of the M3 panel for the measurement of movements in the manufacturing sector. We examined both the M3 sample panel and that part of the panel that reported in the M3 during 1984. We matched the 1984 ASM panel to the existing M3 panel and classified all establishments in the ASM panel into one of the following groups:

(1) Matched to reporting M3 companies.

(2) Matched to companies in the M3 sample but not reporting.

(3) Not matched to the M3 panel.

Individual ASM establishment data were aggregated to the appropriate M3 industry categories. In each category, we developed several estimates of the 2-year percent change from 1982 to 1984. (This time period was selected since a new ASM sample was introduced for 1984 based on the 1982 Census of Manufactures.) One estimate was developed for the M3 panel; it included establishments classified in the first two groups listed above. Another estimate was calculated for the reporting portion of the M3 panel using establishments in Group (1). Finally, we estimated percent change for the complete universe of companies using the entire ASM.

To evaluate the representativeness of the M3 panel, we statistically compared the estimates of percent change for the M3 panel and the universe. Within an industry category, we considered the panel to be inadequate if the difference in the percent changes exceeded 5 percent. Of the 74 M3 industry categories, 14 were identified as having inadequate panel coverage.

Similarly, we compared the percent changes for the reporting M3 companies and the universe. We identified 22 categories with inadequate reporting. The M3 panels in these categories were not found to be inadequate; however, the companies that actually reported in the M3 were not representative of the universe.

Similar criteria were established to evaluate the industry categories aggregated to the 2-digit SIC level. A difference of 2 percent at this level was considered inadequate. At the 2-digit SIC level, 8 of the 20 industry groups had panel or panel and reporting problems.