ELIMINATING ROTATING PANELS IN THE CENSUS BUREAU’S SURVEYS OF RETAIL AND WHOLESALE TRADE

Patrick J. Cantwell and Carol V. Caldwell, Bureau of the Census*

Patrick J. Cantwell, Bureau of the Census, Washington, DC 20233 pcantwell@ccmail.census.gov

Key words: Fixed panel, revisions, composite estimation, panel imbalance, response bias.

1. Introduction

In the Census Bureau's monthly surveys of retail and wholesale trade, we use a rotating-panel design. Larger firms are asked to report sales or inventories each month. All other firms belong to one of three rotating panels. They are contacted only every third month and report data for the two most recent months. Under this design we obtain eight months of data from the smaller firms through only four contacts per year, potentially reducing costs. By compositing data from the current and the prior month, we reduce the variance of estimates of monthly level significantly.

Because the data for a given month are collected during two separate periods, the Bureau first releases a preliminary estimate for monthly level and month-to-month trend. A month later we provide the final estimate, incorporating sample units that report later. The difference between the preliminary and the final estimates is called the revision to the estimate. Through statistical analysis and our estimation methods, we hope to keep the monthly revisions as small as possible. Yet they have been too large in many months, and for some kinds of business they tend to follow a clear cyclical pattern.

With the introduction of new samples in early 1997, the Census Bureau plans to change from rotating panels to a single fixed panel in its monthly surveys of retail and wholesale trade. All sample units would report each month for only the current month. In this paper we compare a fixed-panel design with the current rotating-panel design. We examine the issues of sampling variability, revisions, panel imbalance, response bias, cost, and data quality.

The chief drawback is the expected increased variance in estimates of monthly level. However, for month-to-month trend, the variance is not expected to change noticeably. Although we will obtain less information per sample contact, all units will report in the two months.

But the main reason for changing the design is to reduce the size of the revisions. Revisions will no longer be dependent on which panels report and how badly the panels are out of balance. Smaller factors, such as data corrections, will determine the size of the revisions. Differences due to the way respondents report their data will be nullified, and response bias on the estimates will be decreased.

Section 2 describes the design of our monthly surveys of retail and wholesale trade, while Section 3 details the system of estimation. The main problem with the current procedure--large revisions--is discussed in Section 4. Here, we portray the sources of the revisions--panel imbalance and differential response bias.

In Section 5, the fixed-panel design is introduced. Its effects on variances and bias are discussed in Section 6. Section 7 covers some of the effects on data quality. These are more difficult to assess, but offer arguments for changing to a fixed-panel sample.

2. Background on the Survey Designs

The Monthly Retail Trade Survey (MRTS), conducted by the Bureau of the Census, measures sales in the kinds of business designated by Standard Industrial Classification (SIC) codes 52 through 59. In the Monthly Wholesale Trade Survey (MWTS) the Bureau collects sales and inventory data from merchant wholesalers in SICs 50 and 51. The current designs of these surveys are similar in most aspects except the industries and the geographic areas they cover. Their monthly samples are selected every five years from the Standard Statistical Establishment List, a register of establishments that report quarterly payroll data to the Internal Revenue Service.

Before selecting the sample, we group together establishments belonging to the same company, and assign a major kind of business to the company according to its SIC. Within each major SIC, the largest companies are designated as “certainties,” that is, placed in sample with probability one. These companies report their sales every month shortly after the end of the month. In 1992, when the Census Bureau last selected new samples, about 3500 companies were selected with probability 1 in the MRTS and about 1800 in the MWTS.

The establishments in all remaining companies are then identified by their Employer Identification Number (EIN), and placed together with any other establishment in the same trade area and with the same EIN. Within major SIC and trade area, the EINs are stratified according to their projected total annual sales. We select a simple random sample from each stratum, and assign weights inversely proportional to the probabilities of selection.

To extract more information from fewer sample cases, we select three times as many noncertainty sampling units as the design calls for, and systematically divide these cases into three rotating panels. The firms in a given panel are contacted only every third month, and report their sales or inventories from the most recent two months (thus, a two-
level, three-panel design). For example, early in March sample units in Panel 2 report their "current month" sales for February and their "prior month" sales for January. These firms in Panel 2 are contacted again three months later to provide sales figures for May and April, and so on. Each panel has about 9000 sample units covering the various SICs in retail, and about 1700 in wholesale.

Under this design, each panel reports four times a year, giving us eight months of data through only four contacts, potentially reducing costs. Thus, for any specific month, we collect sales or inventory data from two of the three rotating panels (in two successive monthly data collections) in addition to the certainties, which report every month. For more information on the design of the Monthly Retail Trade Survey, see U.S. Bureau of the Census (1997). As we stated above, the design of the Monthly Wholesale Trade Survey is similar.

3. Estimation

To estimate total sales, we start by summing the weighted sales values. However, because we rotate three panels of noncertainty units in and out of sample, we might see considerable differences in consecutive monthly levels due merely to the different constitutions of the panels. To benefit from the rotating panel design, we apply a composite estimator—a linear combination of estimates using data from the current month and earlier months. This estimator, as applied in the MRTS and the MWTS, is described in Woodruff (1963) and Wolter (1979). They demonstrate how composite estimation reduces the variance of estimates of monthly level significantly compared to the usual weighted estimator.

Let \( U_{it} \) be the "unbiased" sample weighted estimator of sales from the certainty units and from the panel reporting for month \( t \), where \( i = 1 \) (current-month estimator) or \( 2 \) (prior-month estimator) and \( t = 1, 2, 3, \ldots \) (If we let \( \text{mod}_b(n) \) be the number \( n \) modulo base \( b \), the panel reporting is Panel \( \text{mod}_3(t+i+1)+1 \). The weight for any sample unit is the inverse of its probability of selection.

Shortly after month \( t \) ends, the units in the designated panel report (i) current-month sales for month \( t \) (yielding, along with responses from the certainties, \( U_{it} \)) and (ii) prior-month sales for month \( t-1 \) (yielding \( U_{i+1,t} \)). After the responses are processed, edited, and combined with data from the previous month, the Census Bureau releases a "preliminary" composite estimate for month \( t \), defined recursively as

\[
P_t = (1-\beta) U_{1,t} + \beta (U_{1,t} / U_{1,t+1}) P_{t-1},
\]

where \( \beta = .75 \) in MRTS and .65 in MWTS.

One month later, we collect prior-month data for month \( t \) from the next panel, yielding \( U_{2,t} \). (Additional data collected from this panel also produce a current-month estimate, \( U_{0+t+1} \), for data month \( t+1 \).) Combining these responses with those obtained earlier, we publish a "final" composite estimate for month \( t \):

\[
F_t = (1-\alpha) U_{1,t} + \alpha P_t,
\]

where \( \alpha = .80 \) in MRTS and .70 in MWTS. The demand for the data as soon as they are available makes it necessary to release the preliminary estimate before data from the second panel are processed. Because the certainties report every month for current-month sales, they typically do not report a prior-month figure unless there is a correction to make or a revised sales figure.

We call the change from \( P_t \) to \( F_t \), that is, \( F_t - P_t \), the "revision" in sales for monthly level.

4. Revisions to the Data

Too often in recent years our retail and wholesale data have been plagued by occasional large revisions from the preliminary to the final estimate. For example, since April 1992, when the Census Bureau started releasing estimates from a new sample, most of the revisions for the U.S. total retail sales have been upward, that is, the final estimate has been greater than the preliminary. More important, the revisions for some months have been as large as .3% or .4% of the total value. The average percent revision over this period (through April 1996) has been +.12% for sales (with a standard deviation of .15%).

In addition to an upward trend in these revisions, there appears to be a cyclical pattern. The largest revisions tend to be in cycle 2 (the months of February, May, August, and November), while most of the downward revisions are in cycle 3 (March, June, September, and December). Table 1 presents the mean percent revisions by cycle for U.S. Total Retail Sales from April 1992 to April 1996.

This problem of large and somewhat predictable revisions is not confined to retail sales. In wholesale, the cyclical pattern is even more pronounced at the U.S. total for sales and inventories. The average revisions in Table 1 for U.S. Total Wholesale Sales are also given by cycle for April 1992 through April 1996. As we will discuss, with the October 1993 data month the Census Bureau started adjusting the preliminary composite estimates in some wholesale SICs to lessen the problem of panel imbalance. Table 1 summarizes the revisions for wholesale sales two ways: as if no adjustment had been done (averaging over all months) and applying the adjustment (averaging only the 31 months where the preliminary was adjusted).

Small revisions in the estimates are expected because of the rotating design of the surveys and the use of composite estimation. But the cyclical pattern (for retail) and the consistent upward direction of the revisions are due mainly to phenomena we call panel imbalance and differential response bias.
Table 1. Mean Percent Revisions in U.S. Total Sales by Cycle for April 1992 through April 1996

<table>
<thead>
<tr>
<th></th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Jan</td>
<td>+.11%</td>
</tr>
<tr>
<td>Apr</td>
<td>.12%</td>
</tr>
<tr>
<td>Jul</td>
<td>+1.54%</td>
</tr>
<tr>
<td>Oct</td>
<td>.35%</td>
</tr>
<tr>
<td>Mar</td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td>+.46%</td>
</tr>
<tr>
<td>Jul</td>
<td>.50%</td>
</tr>
</tbody>
</table>

1 Based only on data from October 1993 to April 1996.

Panel Imbalance. When we select the noncertainties within each SIC and size stratum for the retail and wholesale surveys, we draw a sample three times the designated size and assign the units systematically to the three panels. Before the first contact, each unit is re-examined to make sure that the early estimate of sales used to stratify the units is as accurate as possible. In this way, as the new sample is phased in, the three panels have essentially the same number of units and, we hope, about the same total volume of sales.

Unfortunately, several things can happen to upset this balance as measured by the volume of sales, either at the phase-in of the sample or during the subsequent five years. Even before our first contact with new sample units, the dollar volumes of the panels may differ due simply to random chance in assigning units to the three panels or to an inaccurate measure of size used to stratify and select the units. Then, during the five years the firms are asked to report, sample births and deaths can further upset the balance among the three panels. In assigning births to panels we try to balance the number of sample units across panels within sampling strata. There is no guarantee, however, that the dollar volume of sales is balanced as well.

The effects of panel balance on the consistency between the preliminary and final estimates—and thus on the revisions—were studied in Cantwell, Caldwell, Hogan, and Konschnik (1995). They showed that, when one panel is much larger or smaller than the others, and panel imbalance dominates other potential factors, the result can be large revisions occurring in specific cycles. For a brief mathematical derivation of this result, see the full-length version of the current paper.

As an example of what might occur, suppose that panel 2 is much larger than panel 3, which in turn is larger than panel 1. If panel imbalance is the dominating factor, we might expect to see large positive revisions in cycle 1 and large negative revisions in cycle 3. This is indeed what has happened in several SICs responsible for a significant part of total sales and inventories in wholesale. Because panel 2 is the largest panel in many SICs, it is by far the largest when aggregated to the U.S. total for wholesale. Thus the revisions follow the aforementioned pattern, as is seen in the middle of Table 1.

In October 1993, the Census Bureau began adjusting the preliminary estimates in the MWTS for sales and inventories in several SICs. By modeling the recent revisions as a time series with a three-month cycle, the adjustment predicts the value of the next revision, and tries to bring the preliminary more in line with the final estimate (as yet undetermined). Greater detail about this method can be found in Caldwell, Monsell, Piesto, and Shimberg (1994). Adjustments were made to additional SICs beginning in April 1995. These are reflected in the third section of Table 1.

Although adjusting the preliminary estimates has greatly reduced revisions in wholesale and removed much of their cyclical pattern, this solution cannot be counted on to resolve the problem in general. The problem is that this method requires many months of preliminary and final estimates based on the new sample to determine the pattern of revisions and to model the three-month time series. This is usually too long to wait, leading to many large revisions before the adjustment can be implemented.

Differential Response Bias. A different problem can arise if, for the same data month, the rotating panels yield different estimates (after adjusting for panel differences) based on the current month ($U_t$) and the prior month ($U_{t-1}$). Reasons for differences in the reporting practices of sample firms have been proposed and studied for many years. Perhaps the brief period given to determine the sales figure after the data month ends allows some respondents only enough time to provide a rough estimate. But for the prior month, these same respondents have had plenty of time to complete their accounts and give us a good "book value." How prevalent this phenomenon is might depend on the size
of the company, the kind of business, the recent level of price changes, and the availability of computerized accounting systems. Part of the difference might also be due to the imputation procedures we use.

Waite (1974) investigated the bias due to early reporting in responses to the MRTS. Based on data collected in 1973, he observed that (p. 604) "This bias does seem to differ for the two reporting periods. ... The current month's sales seem to be underestimated to a greater degree than the previous month's sales." Although several theories have been offered, it is not known why early estimates tend to be lower than book values.

When current-month responses tend to be biased downward more so than prior-month responses and this response bias is the dominant effect, it is easy to show that the preliminary and final composites--our published response bias is the dominant effect, it is easy to show that the revisions from preliminary to final tend to be positive. Details can be found in Waite (1974) or the full-length version of this paper.

We've seen what can happen if panel imbalance or differential response bias alone are present in the design or the data collection. In reality, these effects often occur together. Then, the revisions are generally driven by the stronger factor which--data have shown--is usually panel imbalance (Cantwell et al. 1995). Thus, in many SICs where the response effect is strongly significant, the more powerful influence of panel imbalance dominates the three-month cycle of revisions. Across the three cycles, we see both upward and downward revisions. Still, in this situation the average revision over all months tends to be positive, as Table 1 demonstrates for retail sales.

5. A Fixed-Panel Design; Revisions

Rotating units in and out of sample can reduce variances. But other factors such as panel imbalance and reporting bias can induce large or cyclical revisions from the preliminary to the final estimate. An alternative is a fixed-panel design. Here, all units would report only current-month sales every month for the life of the sample--five years. As this design can eliminate large revisions, it is worth considering despite some increases in variance.

Under the new design for the monthly retail and wholesale surveys, large companies (certainties) would again be selected with probability 1 in the various kinds of business (SICs). Then one panel of noncertainties would be selected from the remainder of the frame, again using simple random sampling within strata determined by SIC and measure of size (projected sales). Our current plans call for selecting about the same number of certainty companies as under the old design. Similarly, the number of noncertainties reporting each month will be about the same as had been in each of the three rotating panels.

From the data for month t in the fixed panel, we derive an unbiased weighted estimate, denoted here by $U_t$. The sample weights are simply the inverse of the probabilities of selection. With sample rotation eliminated, composite estimation no longer reduces variances and thus is not considered. One month after its release, the published estimate $U_t$ would be revised only to reflect data corrections or revised sales figures, births and deaths being tabulated, and perhaps other minor changes. Our research has shown that, while it is not unusual for reporting units to correct their sales or inventories given 30 additional days, this should not cause major revisions in the published totals.

The key result is that the revisions will not be derived from panel imbalance--with only one panel, there will be no panel imbalance; or differential response bias--we'll receive only current month data.

6. Effects on Variance (CVs) and Bias

Along with cost, one of the most important features of any design is the level of the resulting sample variances or coefficients of variation (CVs). As we indicated in the last section, we retain the same sample size in the fixed panel (deriving $U_t$) as we had in the rotating design ($P_t$, equation (1)). But the variance of $P_t$ is smaller than that of $U_t$ due to the use of composite estimation. Further, after the next panel reports for month t in the rotating design, additional independent observations (noncertainties) are available, giving a final estimate $F_t$, whose variance is smaller still.

To compare CVs under the two designs, other conditions are kept the same wherever possible. Formulae for the CVs of the composite estimators are found in Wolter (1979). We insert $\beta = .75$ and $\alpha = .80$, coefficients used in the composite estimates in the MRTS. For this example, the correlations between unbiased estimates from the same panel 1, 3, 6, 9, and 12 months apart are assumed to be .90, .80, .75, .70, and .80, respectively. These numbers are used only for this demonstration; the actual correlations vary according to SIC.

We assume further that the one-month correlation remains the same (.90, here) if we implement a fixed panel. Whether this actually holds is uncertain. With rotating panels, the one-month correlation is measured on $U_{t+1}$ and $U_{t+1.2}$. For rotating units, the pertinent sales figures are reported at the same time; this may induce an artificially high correlation between the estimates. It has also been suggested that the estimate of this correlation is increased because of our imputation procedure. With a fixed panel, the sales figures in $U_t$ and $U_{t+1}$ are reported a month apart and likely would yield a slightly smaller (but more realistic) correlation. Still, imputation for missing months may keep the estimated correlation higher than the true value.

With the parameters as given for estimating retail sales, we expect $CV(U_j)$ to be about 24.8% greater than $CV(P_j)$, and about 40.1% greater than $CV(F_j)$. One can see the greater precision realized by rotating cases in and out of
sample, asking them to report two months of data at a time, and applying composite estimation. However, $CV(U/U_t)$ should be only about 1.7% greater than $CV(P/F_t)$. ($P/F_t$ is considered to be the best measure of trend under the rotation design.) The last result follows because all respondents in the fixed-panel design are asked to report in consecutive months.

It should be noted that (1) the results given depend on the strengths of the correlations of the estimates across months, which vary among the kinds of business; and (2) other trends are also estimated and published each month: current month to the month one year ago, current quarter to previous quarter, current quarter to the quarter one year ago, etc.; here we focus on the two simplest and most important estimates.

When the Census Bureau began rotating panels in and out of sample, a greater emphasis was placed on estimates of monthly level than on estimates of month-to-month trend (Woodruff 1963, p. 455). Since that time, however, the Census Bureau has instituted a system by which the estimates of monthly level are benchmarked to the annual surveys, which are in turn benchmarked to the Economic Census (taken every five years). Because the benchmarking operations take advantage of the greater sample sizes and mandatory reporting in the annual surveys and the Economic Census, the importance of estimates of monthly level has diminished relative to that of estimates of month-to-month trend.

Note also that the CVs for monthly levels as computed here are based on the estimates before benchmarking to the annual surveys and Economic Census. If benchmarking is considered in determining the CVs for monthly level, we believe the CVs will decrease but leave a similar difference in CVs between the rotating and fixed-panel designs. The CVs for trend are not affected by the benchmarking.

Sampling via a fixed panel cannot eliminate all adverse effects. For example, the single panel itself may well over- or underrepresent the frame. That is, if all sample units in the panel reported their true value of sales, the weighted sum of these units might still show a difference somewhat above (or below) the true frame total for the SIC. This is a smaller problem under a fixed-panel design. When measuring month-to-month trend, the excess (or deficit) essentially cancels. This does not happen when estimating monthly level. But if the level is benchmarked effectively, it is brought more in line with the frame total.

Bias. As mentioned in Section 4, data have shown that in retail sales current-month estimates appear to be biased downward much more so than prior-month estimates. With a fixed panel, all sample units would report only the biased current-month sales. But with rotating panels, all noncertainty units report for current- and prior-month sales, the latter being incorporated into the final composite estimate ($F_t$). The question then arises: If only the current-month estimate is biased, which design is to be preferred?

To simplify the answer, we ignore the effects of panel imbalance. Suppose (1) the current-month estimate is biased downward, that is, $E(U_{t1}) = r \times (true \, monthly \, total)$, where $r<1$; but (2) the prior-month estimate is unbiased, that is, $E(U_{t2}) = (true \, monthly \, total)$. It is then easy to show that the preliminary and the final composites (under the rotating-panel design) tend to exhibit a much greater downward bias than the current-month "unbiased" estimator (under the fixed-panel design). That is, typically

$$P_t < F_t < U_t < \text{true monthly total}$$

This result can be explained by recalling the definitions of $P_t$ and $F_t$ in (1) and (2). Under these conditions, $U_{t1}$ and $U_{t2}/U_{t12}$ are biased downward; this forces the preliminary, $P_t$, downward. Further, in the formula for $F_t$ the unbiased estimate from prior-month reporters ($U_{t2}$) has a small coefficient, $1-\alpha$, while $P_t$ is more heavily weighted ($\alpha$ is 80% in MRTS, 70% in MWTS). Although the fixed-panel "unbiased" estimate is biased downward, its bias is much smaller--about 25% that of $P_t$ and about 33% that of $F_t$.

### 7. Effects on Data Quality and Other Issues

Operating with a fixed panel will make analysis of microdata and aggregates much simpler. First, we will have twelve months of data per year (assuming complete response) rather than eight months. This will allow analysts to better monitor the microdata series and check for unusual or suspect responses. Second, with no composite estimator, we can more easily measure an individual firm's effect on the total estimate. Similarly, the effect of births and deaths will be easier to gauge. Removing the confounding effects of panel imbalance and response bias will simplify the analysis of estimates by kind of business.

Third, reconciliation with the annual surveys will be easier. As a check on data quality, at the end of the year the estimates of the monthly retail (MRTS) and wholesale (MWTS) surveys--summed over the twelve months--are compared by kind of business to those from the annual surveys. The differences are reconciled where possible. For individual units, we collect eights months of data from the monthly surveys and project a year's estimate for comparison with the reported value from the annual survey. Any difference arises from a combination of sampling error--having only eight months out of twelve--and reporting error. With a fixed panel, there will be reports from all twelve months. Any differences can be attributed strictly to reporting error.

Other operations are affected by the change to a fixed panel. Procedures for introducing births, eliminating deaths, and addressing firms that have grown unusually large are discussed in the full-length version of this paper.

Finally, we address three important areas. In all three, no firm conclusions between the two designs can be made
Cost. Since we do not plan to change the size of the monthly samples as we move to a fixed-panel design, the cost should decrease just slightly. The price for mailing out questionnaires, processing data as they arrive, following nonresponding cases, and analyzing data should be about the same.

The main difference may be in the start-up costs, that is, the additional cost of first introducing a unit into sample. For some units, there is no additional cost, because they respond as we intend immediately. For others, there is some initial correspondence—usually by telephone—to give the respondent information or instructions, to get him to subsample his establishments properly, to make alternate reporting arrangements, etc. These are usually measured in staff time. Although the two designs have the same number of sample units in any month, the fixed panel has only one-third as many noncertainty units. (Most certainty cases continue when the sample is reselected; their start-up costs then are relatively small.)

Unfortunately, attempts to measure start-up costs—even relative to the monthly cost per sample unit—have been unsuccessful. Overall costs are not easily broken out into the various components. Although the fixed-panel design will likely cost less than the current rotating-panel design, the difference is expected to be small.

Response burden. Response burden can be considered in two ways: as it applies to the individual respondent, and the total burden for the survey. For the former, we ignore certainty companies, since they report as before. But for noncertainty units, it is difficult to anticipate how respondents will react to the new design. Will they prefer to report once every three months, giving us their current sales value and looking up the value for the prior month? Or is it easier to report once each month, never going back to prior months? The answer probably depends on who is responding to the survey, whether he or she prefers to pull the sales figure the same time each month, or what type of accounting system he has. Without canvassing many respondents, we cannot pretend to know.

The overall survey burden, however, will be smaller with a fixed panel. Each month about 9000 noncertainty reporters in retail and about 1700 in wholesale will provide only current-month sales, as opposed to both current- and prior-month sales under the current design.

Response rates. What will happen to response rates when all respondents are asked to report every month? Obviously, the response rate is related to the response burden. How much so is undetermined.

We have studied the response rates of certain cases. These are firms that originally reported every third month. But because their sales had grown much larger than what is typical for their sampling stratum, they were later asked to report every month. (Their weights were adjusted appropriately.) For these cases, there is no significant difference in the response rate reporting every third month or every month. It is difficult to make any conclusions, however, because these cases have experienced unusual growth and are not typical of other respondents. Another approach might be to compare the response rates of the smallest certainty units in the MRTS or the MWTS with those of the largest noncertainty (rotating) units. Currently we are investigating several strategies.

Acknowledgments

The authors thank James Burton, Michael Shimberg, and Nancy Piesto for providing data; and Nash Monsour and Cristina Ibanez for reviewing the paper and providing useful comments.

References


* 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.