

Examining the Revisions in Monthly Retail and Wholesale Trade Surveys Under a Rotating Panel Design

Patrick J. Cantwell¹ and Carol V. Caldwell¹

Under the rotating-panel design used in the U.S. Census Bureau's monthly surveys of retail and wholesale trade, sample firms belong to one of three panels, each reporting only every third month, providing data simultaneously for the month just completed and the month before that. Compositing information obtained from the overlapping panels produces estimates of level with a smaller variance. Due to the timing of the data collection under this design, the Bureau releases a preliminary estimate one month and then revises it one month later. Each of the published estimates is a composite, incorporating data from the overlapping panels. Several factors can affect the size and direction of the revisions. A differential bias in the firms' responses for the two reported months can lead to consistent, positive revisions. But more important, if any of the three panels changes in volume relative to the other two, the result can be a predictable cycle of large revisions. We describe these problems in greater detail, investigating how persistent they are in the Census Bureau's monthly retail and wholesale trade surveys, how they affect the revisions to the estimates, and what can be done about them.

Key words: Panel imbalance; response bias; composite estimation; fixed-panel design.

1. Introduction

In the U.S. Census Bureau's monthly surveys of retail and wholesale trade, a large proportion of the sample firms rotate in and out of sample. Each of these firms belongs to one of three panels and reports every third month, giving sales or inventory figures for the current month (just completed) and the prior month. 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 final and the preliminary estimates is called the revision to the estimate.

Several factors can affect the size and direction of the revisions. As Waite (1974) observed, the respondents may report differently for current- and prior-month sales. Current-month estimates tend to be lower than prior-month estimates for the same data month. As we show, this condition can lead to a constant upward revision. But a more serious problem arises if the panels of sample units become unbalanced. That is, one panel

¹ U.S. Bureau of the Census, SVSD, Room 2653-3, Washington, DC 20233, U.S.A.

Acknowledgment: This article reports the general results of research undertaken by U.S. Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau. The authors thank Susan Iacaruso, James Burton, Michael Shimberg, and Nancy Piesto for providing data, Lynn Weidman for advice on analyzing the variance when modeling the unbiased estimators, and three referees and an Associate Editor for providing helpful comments.

becomes significantly larger or smaller in dollar volume than the others. When this happens, the revisions can become large in absolute value and follow a predictable three-month cycle.

We describe these problems in greater detail, investigating how persistent they are in the Census Bureau's monthly surveys, how they affect the revisions to the estimates, and what can be done about them. An important result suggested by recent data is that *panel imbalance tends to dominate any differential response effect*, at least for the individual kinds of business. This result continues to hold as the kinds of business are aggregated to the U.S. total in wholesale trade. But in retail trade the strong response effect is equally important at the U.S. total level.

In Sections 2 and 3, we describe the design of the monthly trade surveys and the system of estimation. The next section addresses the problems of panel imbalance and differential response bias. In Section 5, we model the unbiased estimates to see what happens to the revisions when these problems (effects) are present. For many important kinds of business, these effects are tested for statistical significance. To counter the consequences of these problems, we consider two alternatives in Section 6. The Census Bureau has decided to incorporate one of these options when the new samples are phased in in early 1997.

2. Design of the Monthly Retail and Wholesale Trade Surveys

The Bureau of the Census conducts several monthly trade surveys using rotating panels, including the Monthly Retail Trade Survey (MRTS) and the Monthly Wholesale Trade Survey (MWTS). The MRTS measures sales in the kinds of business designated by Standard Industrial Classification (SIC) codes 52 through 59. In the MWTS, the Census Bureau collects sales and inventory data from merchant wholesalers in SICs 50 and 51. Each survey provides estimates of level and trend for the nation.

The designs of the two surveys are similar in most aspects except the industries they cover. For each survey, new samples are selected every five years from an establishment list that is maintained and constantly updated by the Census Bureau. 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 (or subunits of 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. 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 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 assign each case to one of 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. For example, early in March sample units in Panel 2 report their "current month" sales for February and their "prior month" sales for January. One month later, Panel 3 units report sales for

Table 1. Source of monthly data from the three rotating panels

	Panel ¹		
	1	2	3
January	current	prior ²	
February		current ²	prior
March	prior		current
April	current	prior	
May		current	prior
June	prior		current
July	current	prior	
August		current	prior
September	prior		current
.	.	.	.
.	.	.	.
.	.	.	.

¹Note: Certainty units report every month.

²Panel 2 rotating units provide these current- and prior-month data at the same time. Other panels respond analogously in the appropriate months.

March and February. Panel 1 is canvassed one month after this. Finally, firms in Panel 2 are contacted again in early June to provide sales figures for May and April. This sequence continues for five years – the life of the sample – before new samples are selected. This two-level, three-panel design is depicted in Table 1.

Under this design, each panel reports four times a year, giving us eight months of data through only four contacts, potentially reducing costs and respondent burden. 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. Composite Estimation and Revisions to the Estimates

To estimate total sales or inventories, we could simply sum the weighted responses. However, because we rotate three panels of noncertainty units in and out of sample, we might see considerable differences in the measures of 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, and estimates of month-to-month trend slightly, compared to the usual weighted estimator.

Let us define $U_{t,i}$ as 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 $t = 1, 2, 3, \dots$ (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. Thus, 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, $U_{t,1}$) and (ii) prior-month sales for month $t - 1$ ($U_{t-1,2}$). 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

$$P_t = (1 - \beta)U_{t,1} + \beta P_{t-1}\Delta_t, \quad \text{where } \Delta_t = U_{t,1}/U_{t-1,2} \quad (1)$$

with $\beta = .75$ in MRTS and $.65$ in MWTS. Note that Δ_t is a ratio of unbiased estimates for months t and $t - 1$ based on common reporters. One month later, we collect prior-month data for month t from the next panel, yielding $U_{t,2}$. Combining these month t responses with those obtained earlier, we publish a “final” composite estimate for month t :

$$F_t = (1 - \alpha)U_{t,2} + \alpha P_t \quad (2)$$

where $\alpha = .80$ in MRTS and $.70$ in MWTS. (At this time, we also tabulate $U_{t+1,1}$ and compute the preliminary estimate for month $t + 1$, P_{t+1} .) 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 current-month sales every month, they typically do not report a prior-month figure unless there is a correction to make or a revised sales figure. Thus, the certainties’ contribution to $U_{t,2}$ is generally about the same as it was for $U_{t,1}$. Note also that, while rotating cases account for around 50% of the U.S. total volume in retail sales, this proportion varies considerably from one SIC to another.

We call the change from P_t to F_t , that is, $F_t - P_t$, the “revision” in sales for monthly level. The final estimate uses data from more reporters – two panels as well as the certainties. Further, responses pertaining to the prior month (those in $U_{t,2}$) are more likely to be “book values” rather than early estimates, given the additional 30 days to report. It follows that the final estimate is statistically the better of the two composite estimates for measuring monthly level. Thus, we hope to produce a preliminary estimate that will be revised as little as possible.

Table 2 presents the mean revisions in the estimates for the U.S. total sales in retail and wholesale trade for 50 months, from April 1992 – when the Bureau started releasing estimates from a new sample – to May 1996. The mean revisions are also given by “cycle.” Cycle 1 represents the data months January, April, July, and October. As one sees from Table 1, for any month in Cycle 1, Panel 1 reports current-month sales while Panel 2 reports prior-month sales (one month later). Cycles 2 and 3 are defined similarly. As we will discuss in Section 6, in October 1993 the Census Bureau began adjusting the preliminary composite estimates in some wholesale SICs in an effort to reduce the size of the revisions. The first set of revisions for wholesale (*without* adjustment) represent the revisions that would have occurred had we made no adjustments. Revisions for SIC 501 – wholesalers of motor vehicles, parts, and supplies – are included for reference in Section 5.

If (i) the differences in the composition of the three panels were due only to sampling variability, and (ii) the estimates based on current month ($U_{t,1}$) and prior month ($U_{t,2}$) reports were both unbiased, then we would expect the mean per cent revisions to be

Table 2. Mean per cent revisions in sales by cycle

Mean per cent revision (standard error in parentheses)	Data months used in analysis	By cycle			Over all months, per cent
		1 Jan, Apr, Jul, Oct, per cent	2 Feb, May, Aug, Nov, per cent	3 Mar, Jun, Sep, Dec, per cent	
Retail, U.S. total	4/92 to 5/96	+ .11 (.12)	+ .22 (.13)	+ .02 (.14)	+ .12 (.15)
Wholesale, U.S. total w/o adjustment	4/92 to 5/96	+1.54 (.35)	-.26 (.35)	-.65 (.41)	+ .23 (1.03)
Wholesale, U.S. total with adjustment	10/93 to 5/96	+ .46 (.50)	+ .21 (.40)	+ .24 (.29)	+ .31 (.41)
SIC 501 w/o adjustment	4/92 to 5/96	+3.81 (.98)	-.93 (.94)	-2.70 (1.21)	+ .12 (2.96)

somewhat balanced around 0 with a standard error of about .23% in MRTS and .63% in MWTS (based on formulae in Wolter (1979)). For U.S. total in retail, the revisions have been upward, that is, $P_t < F_t$, in 40 of the 50 months – 31 of the 34 months in Cycles 1 and 2 – and some have been as large as .3% or .4% of the total value. In wholesale trade, one sees the large cyclical revisions – large positive in Cycle 1 and large negative in Cycle 3. In Cycle 1, 14 of the 17 monthly revisions are more than two standard errors ($2 \times .63\%$) away from 0. In the next two sections we study two problems that can produce these troublesome revisions.

4. Panel Imbalance and Response Bias

When we select the noncertainty sample units 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 weighted 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.

What is the effect of panel imbalance on the estimates and on the revision from preliminary estimate to final? Recall the definitions of P_t and F_t in Expressions (1) and

(2). In addition to data collected in earlier months, $U_{t,1}$ contributes to the preliminary estimate for data month t , while $U_{t,2}$ – collected from a different panel of respondents (as well as from the certainty units) – contributes to the final estimate. If one panel is much larger than the other in terms of weighted volume of sales, we are likely to see a substantial revision. Because of the three-month reporting pattern of the panels, the set of revisions tends to follow a three-month cycle. We examine this phenomenon mathematically in Section 5.

A different problem can arise if the noncertainty sample units report their sales figures differently for the current month and the prior month – what we call *differential response bias*. 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 possibility 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 Census Bureau’s imputation procedures.

Waite (1974, p. 604) investigated the biases in responses to the MRTS. Based on data collected in 1973, he observed that “This bias [due to early reporting] does seem to differ for the two reporting periods. . . . The current month’s sales seems to be underestimated to a greater degree than the previous month’s sales.” As Waite showed – and we will study in the next section – a differential bias in the responses can cause undesirably large revisions from the preliminary to the final estimates.

For various SICs we extracted from data files the unbiased estimates, $U_{t,1}$ and $U_{t,2}$, for the two reporting panels in each month. The MRTS (MWTS) series supplied 36 (30) months of sales data from April 1992 – the first month of the current sample – through March 1995 (September 1994). From these estimates we removed the volume due to the certainty sample units, leaving only the weighted sum from rotating-panel reporters. As we mentioned in Section 3, among the certainty units – who report every month – most do not make corrections to their sales for the prior month (reported one month earlier), but give us *only current-month* sales. Consequently the contribution to $U_{t,2}(i = 2)$ from certainties is primarily a “current-month” report ($i = 1$) in terms of response type or bias. Thus we decided to remove all certainty units from $U_{t,2}$ to ensure that these prior-month estimates contain only prior-month reports, that is, those from the noncertainties. To compare $U_{t,1}$ and $U_{t,2}$ on an equal basis, we also removed the certainties from $U_{t,1}$.

The mean per cent revisions to the preliminary composite estimate, tabulating *only* the weighted responses from *rotating sample units*, are found in Table 3. Note that fewer months of data are used in Table 3 than in Table 2. Revisions for SIC 501 are included; they will be studied more closely in the next section.

Compared to the revisions in Table 2, where the certainty units were included in the estimates, one sees that in Table 3 (i) the differences in the per cent revisions by cycle are more pronounced, and (ii) the mean revisions for total wholesale are now positive in each cycle. These results follow because the certainty units – reporting mostly the same in $U_{t,1}$ and $U_{t,2}$ – tend to dampen the relative panel imbalance and the differential bias in the responses.

Table 3. Mean per cent revisions in sales by cycle – noncertainty units only

Mean per cent revision (standard error in parentheses)	Data months used in analysis	By cycle			Over all months, per cent
		1 Jan, Apr, Jul, Oct, per cent	2 Feb, May, Aug, Nov, per cent	3 Mar, Jun, Sep, Dec, per cent	
Retail, U.S. total	4/92 to 3/95	+ .49 (.49)	+ .65 (.24)	+ .20 (.27)	+ .45 (.40)
Wholesale, U.S. total w/o adjustment	4/92 to 9/94	+4.23 (1.16)	+ .85 (.72)	+ .14 (1.31)	+1.74 (2.10)
SIC 501 w/o adjustment	4/92 to 9/94	+8.68 (1.71)	–1.04 (1.88)	–4.67 (1.23)	+ .99 (5.95)

5. Modeling and Analyzing the Revisions

To understand the sources of problems with rotating-panel designs, it is insightful to express the unbiased estimate $U_{t,i}$ as a product of several components. In what follows, $U_{t,i}$ will be written as $U_{t,j,i}$ to stress its dependence on the reporting panel, indexed by j . (As noted in Section 3, j is uniquely determined by t and i : $j = \text{mod}_3(t + i + 1) + 1$.) For now, when referring to a specific panel reporting in its appropriate month, we also implicitly include the certainty sample units, who report every month; later in this section, we will remove the certainty units from $U_{t,j,i}$ to help analyze the effects of panel imbalance and differential response bias. For a specific kind of business (or for the total across all SICs), let

$$U_{t,j,i} = m_t \times p_{j(t,i)} \times r_i \times e_{t,i}, \quad \text{where} \quad (3)$$

- ▶ m_t is the true, unknown sum of sales (or inventories) over all population units in month t ;
- ▶ $p_{j(t,i)}$ is the “effect of panel j ,” for $j = 1, 2, \text{ or } 3$ (i.e., $j = \text{mod}_3(t + i + 1) + 1$): the average weighted total of only those sample units reporting in panel j divided by the average sum over all population units; for simplicity, we drop the implicit t and i , and label this effect p_j ;
- ▶ r_i is the “response effect,” for $i = 1, 2$: the average weighted total of current-month ($i = 1$) or prior-month ($i = 2$) sample reporters divided by the average sum over all population units; and
- ▶ $e_{t,i}$ is the residual in the model for $U_{t,j,i}$, with expected value 1.

The weighted total of the sample units from a specific panel (or from only current- or prior-month reporters) varies over time. Therefore, we considered the average over all months covered by our analyses – April 1992 through September 1994 for wholesale SICs (or total); April 1992 through March 1995 for retail. Later in this section an interaction term between panel and response effects will be added. A multiplicative model appears to be more realistic than an additive one because panel and response differences

tend to be proportional to the size of the monthly level of sales or inventories. The residual $e_{t,i}$ includes sources of variation beyond those introduced here.

To see the effect of *panel imbalance* on the estimates and the revisions, consider Expression (3) for $U_{t,j,i}$ and make two simple assumptions. Suppose

- (a) $r_i = 1$ for $i = 1, 2$; i.e., the response effects are negligible, and
- (b) the residuals $e_{t,i}$ are small relative to the panel effects p_j , so that $U_{t,j,i} \approx m_t \times p_j$.

In theory, the second assumption is reasonable when the panel effects are substantial because the large sample size induces a relatively small variance for the residuals $e_{t,i}$. In practice, these two assumptions will not hold exactly, but they help demonstrate what happens to the estimates and revisions in this ideal situation.

If these conditions are true, it is easily shown through the recursive relationship of P_t on the $U_{t,j,i}$'s that for Cycle 1 (data months 1, 4, 7, 10, ...)

$$P_t \approx m_t(p_1 + \beta p_3 + \beta^2 p_2)/(1 + \beta + \beta^2) \quad (4)$$

and the revision is approximately

$$F_t - P_t = (1 - \alpha)(U_{t,j,2} - P_t) \approx m_t(1 - \alpha)\{(1 + \beta)p_2 - (p_1 + \beta p_3)\}/(1 + \beta + \beta^2) \quad (5)$$

From (5), it is clear that the size and direction of the revisions in Cycle 1 are a function of the relative sizes of the panel effects. Revisions for the other cycles can be expressed analogously in terms of the monthly and panel effects: Cycle 2 (data months 2, 5, 8, 11, ...) revisions are affected by the magnitude of p_3 relative to p_2 and p_1 ; Cycle 3 (data months 3, 6, 9, 12, ...) revisions pit p_1 against p_3 and p_2 . For example, suppose that p_2 is much larger than p_3 , which in turn is larger than p_1 . If the assumptions given above are roughly true, 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 that are responsible for a significant volume of total sales and inventories in the MWTS, including SIC 501 – motor vehicles, parts, and supplies.

To see the effect of differential response bias, we return to the expression for $U_{t,j,i}$ in (3). The assumptions we make now are similar to those above, but focus on the response effects. Suppose

- (a') $p_j = 1$ for $j = 1, 2, 3$; i.e., the panel effects are negligible, and
- (b') the residuals $e_{t,i}$ are small relative to the response effects r_i , so that $U_{t,j,i} \approx m_t \times r_i$.

If these conditions hold and $\beta r_1/r_2 < 1$, it is easily shown that

$$P_t \approx m_t(1 - \beta)r_1/(1 - \beta r_1/r_2) \quad (6)$$

and the revision is approximately

$$F_t - P_t \approx m_t\{(1 - \alpha)(r_2 - r_1)/(1 - \beta r_1/r_2)\} \quad (7)$$

This result is an extension of a derivation in Waite (1974). Note that the revision in (7), under the assumption of a dominant response effect, *applies to every month*. In particular, when $r_1 < r_2$, as appears to be the case in most retail kinds of business, the expression is always greater than 0. This differs from the previous result involving only a strong panel effect, where the revision exhibits a three-month cycle of positive and negative terms.

We have seen what can happen theoretically if panel or response effects *alone* are present in the design or the data collection. In reality, these effects often occur together. In that case, the revisions are usually driven by the stronger factor which – as the data will show – is often the panel effect.

Let us denote the weighted tabulations of the noncertainty units for data month t as $U_{t,j,i}^{NC}$, where j indicates the reporting panel and i represents the current ($i = 1$) or prior ($i = 2$) month estimate. On adding an interaction term $(p \times r)_{ji}$ to the model in (3) and taking natural logs, the model for $U_{t,j,i}^{NC}$ becomes an additive one:

$$\ln U_{t,j,i}^{NC} = \ln m_t + \ln p_j + \ln r_i + \ln(p \times r)_{ji} + \ln e_{t,i} \quad (8)$$

To estimate parameters in the model, we proceed as follows.

- ▶ $\widehat{\ln m_t} = (1/2)(\ln U_{t,j,1}^{NC} + \ln U_{t,j,2}^{NC})$;
- ▶ $\widehat{\ln p_1} = \Sigma_t(\ln U_{t,j,i}^{NC} - \widehat{\ln m_t})/M_1$, where the sum is over only those months for which Panel 1 reported, and M_1 is the number of such months – 20 (24) in our analyses on wholesale (retail) data; for some of these terms, $i = 1$; for others, $i = 2$; $\ln p_2$ and $\ln p_3$ are estimated similarly; and
- ▶ $\widehat{\ln r_1} = \Sigma_t(\ln U_{t,j,1}^{NC} - \widehat{\ln m_t})/M_0$, where the sum is over all months, but only current-month estimates $\ln U_{t,j,1}^{NC}$ are included; and M_0 is the number of months – 30 (36) for wholesale (retail); $\ln r_2$ is estimated similarly.

Before estimating the interaction terms, the three relevant components are combined and estimated as one. Define $\ln(p \oplus r)_{ji}$ as $\ln p_j + \ln r_i + \ln(p \times r)_{ji}$, for $j = 1, 2, 3$; $i = 1, 2$.

- ▶ $\widehat{\ln(p \oplus r)_{11}} = \Sigma_t(\ln U_{t,j,1}^{NC} - \widehat{\ln m_t})/M_{11}$, where the sum is over only those months for which Panel 1 reported *and* $i = 1$ (that is, current-month estimates only), while M_{11} is the number of such months – 10 (12) for wholesale (retail); the other $\ln(p \oplus r)_{ji}$ terms are estimated similarly; from these estimates,
- ▶ $\widehat{\ln(p \times r)_{ji}} = \widehat{\ln(p \oplus r)_{ji}} - \widehat{\ln p_j} - \widehat{\ln r_i}$, for all j and i ; finally
- ▶ $\widehat{\ln e_{t,i}} = \ln U_{t,j,i}^{NC} - (\widehat{\ln m_t} + \widehat{\ln p_j} + \widehat{\ln r_i} + \widehat{\ln(p \times r)_{ji}})$

Note that the true, unknown values of the effects $\ln p_j$ and $\ln r_i$, being defined relative to the average of the real values of $\ln m_t$, need not sum to 0. However, the *estimates* we computed are so constrained, because the m_t 's are not known but must also be estimated. Thus, each of $\Sigma_j \widehat{\ln p_j}$, $\Sigma_i \widehat{\ln r_i}$, $\Sigma_j \widehat{\ln(p \times r)_{ji}}$, $\Sigma_i \widehat{\ln(p \times r)_{ji}}$, $\Sigma_t \widehat{\ln e_{t,i}}$, and $\Sigma_i \widehat{\ln e_{t,i}}$ is constrained to be 0.

Using these estimated parameters and residuals, we performed an analysis of variance (ANOVA) on the (log) estimates $\ln U_{t,j,i}^{NC}$ under a balanced incomplete design. To determine whether the (log) panel component, $\ln p_j$, for example, is statistically significant, we computed the sum of squares of the combined (log) panel-response terms, $\ln(p \oplus r)_{ji}$, and subtracted the part due to (log) response types, $\ln r_i$. In this way, the remaining sum of squares – for *the panel effect after the response effect is removed* – represents the real contribution of the panels to the overall variation in the estimates. We computed the response-after-panels component of the sum of squares similarly. The ANOVA for the panel effect and the response effect (with the other component removed) is seen in Table 4 for SIC 501.

The high values of the F statistics indicate that both effects are strongly statistically significant. An examination of the levels of the panels themselves shows that Panel 2 is

Table 4. Analysis of variance for log model components in SIC 501: Motor vehicles, parts, and supplies (wholesale)

Component	Sum of squares	df	Mean square	F statistic
Months	.6368	29	.0220	7.07
Panel and response (combined)	.9194	3	.3065	98.62
Response	.0023	1	.0023	
Panel after response	.9171	2	.4585	147.56**
Panel	.6878	2	.3439	
Response after panel	.2316	1	.2316	74.53**
Residuals	.0839	27	.0031	
Total	1.6401	59		

**Significant at $\alpha = .01$

much larger than the other two, while Panel 1 is the smallest of the three. As a standardized measure of the difference in the sizes of the panels, we divided each of $\ln p_j$ by the square root of the mean squared error from the ANOVA table, the latter being an estimate of the standard error of $\ln e_{t,i}$. The measures are -2.25 , 2.45 , and $-.20$, for $j = 1, 2$, and 3 , respectively. Thus we expect to see large positive revisions in Cycle 1 and large negative revisions in Cycle 3. This is confirmed in Tables 2 and 3.

In fact, the second panel is larger than the other two for many SICs in wholesale, and indeed is larger when we aggregate to the U.S. total for sales and inventories. In the ANOVA applied to unbiased estimates for the U.S. total in sales (noncertainty units only), the F statistic for the panel effect (after the response effect is removed) is 174.76; the standardized measures of $\ln p_j$ from the ANOVA are -2.22 , 2.80 , and $-.59$. As anticipated, the result is a series of large positive revisions in Cycle 1. Although the per cent revisions for U.S. total in wholesale sales in Tables 2 and 3 are smaller than corresponding values for SIC 501, they represent larger dollar volume.

When modeling the (log) unbiased estimates for retail sales, we based the results on the data months April 1992 through March 1995, again using only the rotating-panel reporters. An analysis of the variance revealed statistically significant (log) panel and (log) response effects in many important SICs. For total U.S. retail sales, both sets of effects are significant. From these results and other analyses, we make the following observations.

- There is a strong correlation between the significance of panel and response type. When the response type is statistically significant, the panel effect usually is as well.
- The F statistic for the panel effect is generally larger than that for the response effect. In fact, *in volume of dollars*, the estimated effects for panels is usually much larger than that for response types. Thus, in many SICs where the response effect is strongly significant, the more powerful influence of the 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 (when $r_1 < r_2$).
- Serious panel imbalances in the detail-level SICs can partially cancel when aggregating to higher-level totals if different panels are larger in different SICs. This

occurs somewhat in retail, although the larger size of Panel 3 at the U.S. total level still drives up the revisions a bit in Cycle 2. In the wholesale sample, Panel 2 is larger than the others in many large SICs, leaving the U.S. total with a serious panel imbalance and a strong cycle of large revisions.

- For most of the retail SICs we studied, the response effect between current- and prior-month reporting is statistically significant. *In every such case*, $\widehat{\ln r_1} < \widehat{\ln r_2}$. This implies that current-month reports tend to be lower on average than prior-month reports.
- With the response effect consistently pointing in the same direction ($r_1 < r_2$), this effect does not cancel when aggregating to higher levels. For example, it is significant in the U.S. total retail sales. The result is a series of mostly upward revisions. (See Table 2.)
- Although the interaction effects $\ln(p \times r)_{ji}$ appear to be of moderate size in some kinds of business, we could find no particular pattern here.

6. Adjust the Preliminary or Change the Design?

Because of problems with severe panel imbalance, the Census Bureau has adjusted preliminary estimates several times in recent decades. Preliminary estimates for sales and inventories in the Monthly Wholesale Trade Survey have been adjusted in several SICs starting with the October 1993 data month, 18 months after the introduction of a new sample. The SICs affected – including 501 – were chosen according to the size and consistency of the revisions observed. Adjustments were made in several additional SICs in wholesale starting in April 1995.

A seasonal adjustment model has been used to express the recent revisions as a time series with a three-month period, and then predict the value of the next revision. This method produces a factor that, when applied to the preliminary estimate, brings the preliminary more in line with the final estimate to be computed one month later. Greater detail about the adjustment can be found in Caldwell, Monsell, Piesto, and Shimberg (1994). By adjusting the preliminary estimates in several problem SICs, the revisions at the U.S. total level have also been consistently reduced. From October 1993 through April 1996, the average revisions in wholesale sales for the U.S. total have dropped for Cycles 1, 2, and 3 to +.46%, +.21%, and +.24%. (See Table 2.)

This solution, however, cannot be counted on to resolve the problem of panel imbalance in general. The 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. An alternative is to use a fixed-panel design, where all units would report only current-month sales every month for the life of the sample. Births would be added on a regular schedule, and a new sample would be selected after five years. For month t , a fixed panel produces an unbiased weighted estimate, denoted here by U_t , with sample weights based on the probabilities of selection. Composite estimation is no longer beneficial and thus is not considered. The published estimate U_t would be revised one month after its release only to reflect data corrections or revised sales figures.

The variance of P_t is smaller than that of U_t because P_t combines data from all reporters up to time t . Further, after the next panel reports for month t , additional observations

(noncertainties) are available in the rotating design, giving a final estimate F_t whose variance is smaller still. To compare coefficients of variation (CVs) under the two designs, the monthly sample sizes (certainties and noncertainties), the values of correlations between estimates one or more months apart, and other relevant conditions are kept the same. Applying formulae for the CVs of composite estimators found in Wolter (1979), $CV(U_t)$ for retail sales is about 24.8% larger than $CV(P_t)$ and about 40.1% larger than $CV(F_t)$ under these simplified conditions. But for estimating month-to-month trend, $CV(U_t/U_{t-1})$ is only about 1.7% larger than $CV(P_t/F_{t-1})$. The latter result follows because all respondents in the fixed-panel design report in consecutive months.

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). Since that time, however, the 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 of the larger sample sizes and mandatory reporting in the annual surveys and the Economic Census, these benchmarking operations improve the estimates of monthly level under a fixed-panel design as well. Therefore, when assessing the change in the sample design, we are more concerned with the effects on estimates of month-to-month trend. It should be noted 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, we believe the CVs for monthly level will decrease but leave a similar relative difference between the rotating and fixed-panel designs. The CVs for trend are not affected by the benchmarking.

Addressing a different point, results in Section 5 show that in retail sales current-month estimates appear to be biased downward relative to prior-month estimates. Is this reason to forego the fixed-panel design – where monthly estimates are based only on current month data? To simplify the answer, we ignore the effects of panel imbalance. Suppose (i) the current-month estimate is biased downward, that is, $E(U_{t,j,1}) = r_1 \times m_t$, where $.95 < r_1 < 1$; but (ii) the prior-month estimate is unbiased, that is, $E(U_{t,j,2}) = m_t$. Using Equation (6) and a similar result for F_t , it is easy to show that, although the fixed-panel estimate U_t is biased downward, its bias is *much smaller* than that of the current composite estimators. For the MRTS (MWTS), the downward bias of U_t is only about 25% to 29% (35% to 38%) that of P_t , and about 31% to 36% (50% to 55%) that of F_t .

The Census Bureau plans to eliminate its rotating panels and implement a design with a single fixed panel in its monthly surveys of retail and wholesale trade with the introduction of new samples in early 1997. The chief drawback is the expected increase in the variance of estimates of monthly level. However, reducing the size of the projected revisions by mitigating the effects of panel imbalance and differential response bias was an important factor in making the decision.

7. References

Caldwell, C.V., Monsell, B.C., Piesto, N.A., and Shimberg, M.Z. (1994). Using Seasonal Adjustment Models to Correct for Panel Imbalance in the Monthly Wholesale Trade Survey. Proceedings of the American Statistical Association, Section on Business and Economic Statistics, 83–88.

U.S. Bureau of the Census (1997). *Current Business Reports, Monthly Retail Trade: Sales and Inventories, January*. Publication no. BR/97-01, Washington, D.C.: Government Printing Office.

Waite, P.J. (1974). An Evaluation of Nonsampling Errors in the Monthly Retail Trade Sales Data. *Proceedings of the American Statistical Association, Section on Business and Economic Statistics*, 602–607.

Wolter, K.M. (1979). Composite Estimation in Finite Populations. *Journal of the American Statistical Association*, 74, 604–613.

Woodruff, R.S. (1963). The Use of Rotating Samples in the Census Bureau's Monthly Surveys. *Journal of the American Statistical Association*, 58, 454–467.

Received February 1996

Revised April 1997