Errors Associated with Period Reporters in Retail Trade Estimates[1]

Kelly Dixon
U.S. Census Bureau, SSSD, Room 2754-3, Washington DC 20233-6500

Abstract

The Census Bureau's Advance Monthly Retail and Food Service survey estimates sales on a calendar month basis. However, some "period reporters" report sales covering either a 4-week or 5-week period that rarely coincides with the exact beginning and ending dates of the calendar month. These period reports are adjusted to a calendar month basis using trading day weights computed as part of the seasonal adjustment process. This research examines whether the current methodology adequately accounts for the increased holiday sales that occur in November and December. During these months, survey analysts attempt to re-contact large period reporters to convert their original period report to a calendar basis. The companies that have both a period and a calendar report are used in this research to estimate the effects of unadjusted period reports, the current adjustment method and a simple adjustment based on the number of days contained in the period. The purpose is to measure the error caused by adjusting the period reports and discuss ways to improve measurement of the error.

Keywords: Retail Trade, Period Reporters, Measurement Error

1. Introduction

Business surveys face special measurement problems. Many firms may not be able to provide the information requested in surveys because their record systems do not fit the survey's needs. Ignoring this problem could lead to nonresponse and measurement errors in surveys. Nonresponse occurs if the respondent skips the question or refuses to participate altogether because getting the information is too difficult. Measurement error occurs when the respondent provides an answer that does not match the request. The survey designer has the options of changing the questionnaire to better reflect the bookkeeping methods of its respondents, requesting additional information, or developing a way to identify the incorrect responses and make them fit the request through an adjustment procedure. Implementing any of these solutions adds steps to the survey process and may introduce additional errors. The respondent may still misunderstand the re-worded questionnaire or refuse the additional data requests. The adjustment process may be imperfect and become another source of error. The survey designer's job is a tough one: to find a solution that minimizes error and does not adversely affect survey processing or costs.

In MARTS, the Advanced Monthly Retail Trade Survey, the Census Bureau asks companies to report their sales for the prior month. Most companies give their reports on a calendar basis -- meaning their sales for November are for November 1-30. Companies that report for other time spans are called "period reporters." The Census Bureau's solution for period reports is to adjust the report to a calendar basis using adjustment factors produced as a byproduct of the seasonal adjustment process. Although the theory behind this method is solid, the error introduced by the period reporter adjustment needs to be evaluated with actual survey data. There has not been a study relating the theory to the practice in the actual surveys for the retail sector.

The remaining sections of the paper will define period reporters and how they affect MARTS, give an overview of the current adjustment method, describe a special data collection that facilitated the study of measurement error, and discuss ways to better understand the effects of period reporters in retail surveys.

1.1 Reporting to the Advanced Monthly Retail Trade Survey

The statistical model for measurement error in establishment surveys presented by Biemer (Biemer, 1995) states that measurement error comes from six sources: 1) the information system, 2) the respondent, 3) the mode of data collection, 4) the interviewer, 5) the survey questionnaire, and 6) the interview setting. The first item in Biemer's measurement error model, the information system, is where most of the problems associated with period reporters occur. Many companies choose to keep their accounting books in 4 week or 5 week periods instead of by the calendar. Some companies keep thirteen 4 week periods. Others choose to report in '445' or '454' periods where there are exactly 13 weeks in each quarter. Publicly traded companies are required to publish quarterly earnings and profits, so this gives comparability across periods.

Sales data for the calendar month may not exist in a retail company's records. The respondent gives the period report as the closest approximation to the survey request. Freedman (1988) refers to this type of error as specification error and describes that the respondent is using an existing concept that does not really fit the concept being measured. Biemer advises "Incompatibilities between the survey and the business record system are best remedied by revising the survey

[1] This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical, methodological, technical or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

item definition because respondents are unlikely to revise their record systems." (Biemer, 1995.) The MARTS questionnaire currently used has built in the flexibility to accept calendar or period reports. The period reports are entered into the processing database and are converted to a calendar basis in the estimation phase. If this adjustment did not take place, then there would be measurement error. Similarly, if the adjustment is imperfect then MARTS remains subject to measurement error through the adjustment procedure.

2. Current Methodology: Overview

The methodology underlying the period reporters adjustment comes from the seasonal adjustment and economic time series research conducted by the Census Bureau in the 1960s by Shishkin, Young and Margrave (Shiskin et al.) The seasonal adjustment method currently used at the Census Bureau is the X-12-ARIMA program (Findley et al, 1998). Seasonal adjustment attempts to parse apart the original unmodified series (*O*) into seasonal (*S*), cyclical (*C*), trading day (*TD*) and irregular (*I*) components. A simple multiplicative model is $O = S \times C \times TD \times I$ (Shiskin et al, 1967). The models currently used involve more complicated exponential functions of the same four components. Economists are primarily interested in the cyclical and seasonal series (*S, C*), so the goal of seasonal adjustment is to isolate these components.

Estimating the trading day component is an important part of the seasonal adjustment process. Young (1965) suggested modeling the trading day effects as the second step in the seasonal adjustment process. Over the decades, the models have changed from multiplicative to logarithmic models and the order has changed to adjust for the trading day variation first. Calendar composition, which is the number of Fridays, Saturdays, Sundays, etc. in a month, is known ahead of time. Therefore, its effects on monthly data can be predicted. Once you control for the trading day variation, it is easier to understand the true changes in the series. The period reporter adjustment comes from the trading day weights that are derived as part of modeling the trading day variation.

2.1 Current Methodology: Trading Day Weights

Trading day variation is believed to be systematic and stable over the years, and thus can be controlled for. Trading day variation is defined as the monthly variation in a series related to the within-month variation of calendar composition. For example, November 2003 has five Saturdays (usually the biggest day of sales for retail trade) while November 2004 has four Saturdays. Thus, the two Novembers have different calendar composition. To compute the trading day factor, months of the same calendar composition are grouped together, since it is the number of each type of day of the week in the month that creates trading day variation (Young, 1965). This is the important feature of trading day component that makes it distinctive from the seasonal component: seasonal adjustment groups adjacent months together (October, November, December), whereas trading day adjustments group months together based on similarity of composition. For example, Young groups all 31-day months beginning on a Sunday together, all 30 day months beginning on a Saturday together, etc. There are 22 types of months and the time series for retail go back to the 1960s.

Young (1965) first seasonally adjusted the data and then attempted to model the trading day variation with the irregular component according to this equation:

$$I_i = \frac{X_{1i}B_1 + X_{2i}B_2 + \ldots + X_{7i}B_7 + E_i}{N_i},$$ where I_i is the modified irregular for month i., and $E[I_i] = 1$. Conceptually I_i is the variation left in the series after the seasonal component is modeled. It consists of a trading day component, which will be modeled out, and the "true" irregular component for month i, E_i. X_{ji} is the number of times day-of-the-week j occurs in month i. (This number is 4 or 5.) B_j is the daily weight, where $\sum_{j=1}^{7} B_j = 7$. For example if there was no activity on Sunday, then B_1=0. Saturday typically has the highest weight in retail series. N_i is number of days in the month; either 31, 30 or 28.25 for February (because of the leap year situation). This is estimated as

$$\hat{I}_i = \frac{X_{1i}b_1 + X_{2i}b_2 + \ldots + X_{7i}b_7}{N_i},$$ where

b_j is the least-squares estimate of B_j.

If there is no trading day variation, then the estimated daily weights will all be 1. (Young, 1965, p.8). This is not the current assumption made for retail trade because years of time series data indicate that monthly sales level vary with calendar composition. More current research has shown that it is more effective to model the irregular and trading day factors first, but the method is largely the same (Soukup and Findley, 2000.)

2.2 Current Methodology: KPR Factors

The mathematical statisticians working on the retail surveys run X-12 seasonal adjustment for each kind of business. "Kind of business," denoted as KB in tables of this paper, is the basic industry tabulation level for the retail surveys. The trading day weights are output and used to construct a constant period reporter factor (KPR factor.) There are two KPR factors produced: one for 4-week reporters and one for 5-week reporters. The KPR is the sum of all the trading day weights in a month, plus any special holiday effect for holidays occurring at the boundary points of the month, divided by either 28 or 35. A holiday effect is added onto the days before and after the holiday. Major holidays in the retail area are: New Years Day, Easter, Memorial Day, July 4th, Labor Day, Thanksgiving and Christmas.

November 2003

S	M	T	W	Th	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

November 2003 KPR factors are computed below as an example. As stated above, the sum of the trading day weights for one week equals 7. Equivalently, the sum of 4 weeks of trading day factors equals 28. For November 2003, there are 4 weeks, plus one extra Saturday, one extra Sunday, and a major holiday (see calendar.) Suppose that a KB's trading day weights are 1.147 for Saturday, 0.750 for Sunday, and an additional 0.10 is computed for the days around the holiday, then the 4-week and 5-week KPR factors are:

$$KPR_4 = \frac{28+1.147+0.750+0.1}{28} = 1.071,$$

and

$$KPR_5 = \frac{28+1.147+0.750+0.1}{35} = 0.857 \cdot$$

If company ABC reported $1,500,000 in sales for a 4-week period, then its monthly sales would be tabulated in the survey as $1,500,000 x 1.071=1,606,500. If company XYZ also reported $1,500,000 in sales but for a 5-week period, then its sales would be tabulated as $1,500,00 x 0.875=$1,285,500.

2.3 Current Methodology: Limitations

The drawback of the current method is that it makes the assumption that a day of the week is assumed equal at any point in the year. In the example above, the 1.147 would be used in the KPR adjustment of its KB for every Saturday (except holidays) in the year. Around the holiday times especially, this is a strong assumption. Also, adjusted period data are used in X-12 to produce the trading day weights for the next year. This is a "chicken-and-egg" situation where trading day estimates are produced using data with forecasted trading day weights. This is done because period reporters in some industries dominate the sales totals, and if they were omitted the X-12 results would suffer (Soukup and Findley, 2000). The trading day estimates are not good for highly irregular series, or series with too few observations input into the program. Another thing to consider as we move into the future is that business being conducted 24 hours a day, 7 days a week in a global environment. Retail stores used to be closed on Sundays, people used to work 9 to 5, and many retail locations closed earlier than 7 pm. This is not the case anymore and perhaps peoples' shopping patterns will change in ways that have not been considered

3. Motivation

Table 1 shows the characteristics of period reporters in the MARTS sample and shows why this is an important group to study. Although period reporters represent a small proportion of the total sample overall (22.3%), they account for a larger percentage of the sample in the grocery, apparel and general merchandise (GM) categories. In addition, period reporters represent the vast majority of sales in that category (87.4%) and overall (76.6%). General merchandise includes establishments like department stores, discount stores, and dollar stores.

Table 1. Period Reporters Percentages of Sample Size and December 2003 Unweighted Sales

Kind of Business	Period Reporters	Total Sample	% Pd Rep of Sample	Pd Reps Sales As % of Total Sales
Grocery	223	574	38.9%	93.9%
Apparel	181	425	42.6%	85.2%
General Merchandise (GM)	66	111	59.5%	85.4%
Grocery, Apparel, and GM	470	1110	42.3%	87.4%
TOTAL SAMPLE	1114	5001	22.3%	76.6%

The next table shows differences in response rates between the total sample and period reporters. The response rates shown below are the cross sectional response rates for the given month. A company needs to respond in two consecutive months and pass rough data edits to be considered a respondent in MARTS. There is no imputation in this survey due to the short estimation period.

Table 2. November 2003 and December 2003 Response Rates (RR) for the Sample and Period Reporters (Pd Rep)

			NOVEMBER			DECEMBER		
Kind of Business	Sample Size n	Pd Rep n	Total RR	Pd Rep RR	Non Pd Rep RR	Total RR	Pd Rep RR	Non Pd Rep RR
Grocery	574	223	65.3%	69.1 %	62.9%	68.5%	77.6 %	62.7%
Apparel	425	181	74.4%	86.7 %	65.3%	76.0%	86.2 %	68.4%
GM	111	66	82.9%	90.9 %	71.2%	79.3%	89.4 %	64.4%
Grocery, Apparel, GM	1110	470	70.5%	78.9 %	64.3%	72.6%	82.6 %	65.3%

As shown in Table 2, the period reporters have better response rates in both November and December. This may be explained by the fact that the group of period reporters contains large companies. Large companies typically are better respondents because they are conditioned to more survey requests and have larger staff to accommodate the requests. Period reporters are an important group because they account for a majority of sales volume and respondents in these three kinds of business.

3.1 Special Data Collection

Given the information in Tables 1 and 2, it is easy to see why the survey analysts would target period reporters in the grocery, apparel, and general merchandise kinds of business if they were trying to reduce the error due to the KPR adjustments. These KBs have the highest concentration, greatest sales percentage and best response rates among period reporters in the sample.

In November and December of 2003, the survey managers for MARTS requested additional data from large period reporter companies in an attempt to get their sales on a calendar month basis. For example, if the company gave a period report from November 2nd-29th, then they would ask for additional information to get the sales for November 1st and 30th to convert their sales to a calendar month basis for November. If a December report covered November 30th to January 3rd, they would be asked for sales for November 30th and January 1st-3rd. The sales from these days would be subtracted out of the period report to convert their report to a calendar month basis for December. The analysts request the additional data because they believe that the calendar data will improve the estimates during the highly publicized holiday shopping season. However, there was high nonresponse to this special data collection since it was an extra burden and some companies refused to give reports for the additional days. The analysts targeted companies that are "good period reporters" meaning that the company was a known period reporter that had a history of timely reporting and they had not reacted negatively to additional requests in the past. The extra data collection was done over the phone, and many companies were excluded because they prefer to respond via fax or mail, or simply refused to take phone calls.

There are complications with using this data because it is not a random sample from all possible respondents, and the good reporters differ from the other reporters in several key areas. Also, the data were not collected for the purpose of evaluating the KPR adjustments – the purpose was to get the best estimates possible and to take advantage of the daily sales figures that may be kept during the holiday season. The additional daily sales data were simply stored in an interviewer's comment field in the MARTS database. The interviewers recorded the daily data and there were no edits for reasonableness, nor was the data always collected in the same units (ie, thousands or millions). Biemer notes that "in many situations, the required randomizations and experimental manipulations would be too expensive, too impractical, or simply impossible to perform; "nonexperimental" or observational studies are conducted instead …" (Beimer, 1995,p. 269). Because it is not a random sample, we will need to consider how those targeted for the special collection differ from the rest of the sample.

Table 3 provides a description of the targeted group. The targeted companies were 31.5% of the subgroup and accounted for 36.3% of November sales and 39.0% of December sales. Although we do not have a random sample and cannot generalize results beyond the targeted group, it is good to see that the targeted group's means on important characteristics do not look radically different from the period reporter group as a whole.

Table 3. Comparing Targeted Cases to the Period Reporters(Pd Rep) in Grocery, Apparel and GM

Measure	Grocery,Apparel, and GM Pd Rep	Targeted Pd Rep	Targeted As % Of Total
Sample size (n)	470	148	31.5%
Mean 1997 Census Sales	$4,525,373	$4,296,472	29.9%
Mean Nov. Unweighted Sales	$135,824,874	$156,767,069	36.3%
Mean Dec. Unweighted Sales	$183,111,771	$226,757,641	39.0%

The response to the special data collection has two levels. The first level is responding to the regular monthly survey request. The second level is providing daily sales data. If the company gave a period report and were targeted by the analysts, the analyst would then ask them for their daily sales data to turn their period report into a calendar report. For example, if company XYZ gave a 4-week report covering November 1st-29th then the analyst would ask for their daily sales on November 30th. Providing the daily sales is an added burden to the respondent and providing such detailed data may not be possible or it could go against company policy. Additionally, MARTS estimates are published only ten working days from the end of the month, so the data collection period is very short. Of the ones that refused, most were hard refusals, but there were some instances of soft refusals, where the company initially agreed to give daily data, but never could come up with the figures before data closeout. Some cases had to be counted as non-respondents in the analysis stage because the daily data they gave did not pass rough edits. It was too hard to tell if their daily figures were in different units then their original report or if there was an error in one of these reports. For all these reasons, the overall response rate to the special collection for the three kinds of business focused on was only 56.8% for November. The poor response among the grocery stores (29.3%) was largely responsible for this figure.

Table 4. Response to the Special Data Collection in November by Kind of Business (KB)

KB	Responded (1)	Targeted (2)	Targeted and Responded (3)	RR to Special Collection (3)/(2)
Grocery	26	41	12	29.3%
Apparel	68	75	47	62.7%
General Merchandise	32	32	25	78.1%
Total	126	148	84	56.8%

3.2 Analysis Plan

Despite the problems associated with the data, the appeal of using it for analysis is that it contains both the original period report and the calendar report. The analysis was limited to the three kinds of business where the concentration of period reporters is highest: apparel, grocery stores and general merchandise, which includes establishments like department stores, discount stores, and dollar stores.

Biemer (1989) explains measurement error with the following useful equations. We can assume that each company has a true monthly sales value, μ_j. What we tabulate in the survey is y_j.

$$y_j = \mu_j + \varepsilon_j \text{, where}$$

$$\varepsilon_j \sim \left(0, \sigma_j^2\right), \mu_j \sim \left(\mu, \sigma_\mu^2\right) \text{ and all covariances}$$

between different j units are 0. The error introduced by the adjustment is captured in ε_j. The adjustment also will lead to a variance inflation.

$$Var(\bar{y}) = \frac{\sigma_\mu^2}{n} + \frac{\sigma_\varepsilon^2}{n}$$

where σ_μ^2 is the variance of the true values and σ_ε^2 =E(σ_j^2)is the mean of the individual variances, σ_j^2, and n is the sample size. This is a reasonable model and similar to the model Ponikowski and Meily (1996) use in their study of response error in establishment surveys. They reason that reports are obtained from records and either mailed or phoned into the survey organization, which means that there is little interaction with other respondents or with survey interviewers.

In the analysis, an actual calendar report was treated as the true value, μ_j. The adjusted period report was treated as the predicted value, y_j. The bias was calculated as the sum of the differences between the true value and predicted value over all period reporters in a particular kind of business for a given adjustment method. There are three adjustment methods: the current method (KPR adjustment), an adjustment based on estimated daily sales in the period report (ad hoc adjustment), and no adjustment at all. The estimated daily sales is an ad-hoc method but has some merit. With this method, we will simply estimate a company's calendar report by multiplying the period report by the number of days in the month divided by the number of days in the period report. This adjustment method essentially sets all trading day factors to 1. The adjustments are as follows:

1. *KPR, adjusted=period report × KPR factor*
2. *Ad-hoc, adjusted=period report × (# days in month/#days in period)*
3. *Unadjusted=period report*

Consider the following example. Company XYZ originally reported $620,000 in sales for a 4-week period, November 1st - 29th. In the special data collection, they reported sales on November 30th as $40,000, which means their calendar report for November 1-30th is $620,000+$40,000= $660,000. If the KPR factor was 1.071 then their KPR adjusted sales is $664,020 and their ad-hoc adjusted sales

$$= \$620{,}000 \times \frac{30\,days}{28\,days} = \$664{,}285 \cdot$$

Each adjustment method was evaluated based on its bias and mean square ratio, MSR (defined below). The bias measures systematic differences of the true response from the adjusted response.

$Bias(Method) = \sum_{k=1}^{n}(calendar - adjusted)$, where KB is the kind of business, the method is the type of adjustment used, and n_{KB} is the number of period reporters targeted and responded in the special data collection in a kind of business. The MSR is the ratio of the MSE (mean square error) for two estimates. The mean square ratio is a useful statistic for comparing the MSE of an adjusted estimate to a standard estimate (Brick and Keeter, 1996). For this study, each of the adjustment method's (KPR, ad-hoc and unadjusted) MSE estimates was compared to the calendar estimate's MSE. As stated above, we assume that the calendar method has bias equal to 0.

$$MSE(Method) = Bias^2(Method) + Variance(Method)$$
$$= \left[\sum_{k=1}^{n_{KB}}(calendar - adjusted)\right]^2 + \left[\frac{\sum_{k=1}^{n}(y_j - \bar{y})^2}{n-1}\right]$$
$$MSE(Calendar) = Bias^2 + Variance(Calendar) = 0 + Variance(Calendar)$$
$$MSR(Method) = \frac{MSE(Method)}{MSE(Calendar)} = \frac{Bias^2(Method) + Variance(Method)}{0 + Variance(Calendar)}$$
$$= \frac{Bias^2(Method)}{Variance(Calendar)} + \frac{Variance(Method)}{Variance(Calendar)}$$
$$= \frac{Bias^2(Method)}{(Std\,Dev(Calendar))^2} + VIF,$$

where VIF is the variance inflation factor due to adjustment.

If the trading day effects model holds, then the KPR adjustment method should yield the lowest biases. Variance inflation will also occur if the adjustment method fails to bring the period report close to the true calendar report. Not adjusting at all should have the largest bias and largest MSE since it fails to adjust the period reports to a comparable level to the calendar reports. It is hypothesized that the ad-hoc method should fall between the unadjusted and KPR-adjusted methods.

4. Results

Table 5 does not show a clear pattern of biases for each method. The ad hoc adjustment method had the lowest bias measures for November, but the ad-hoc and KPR methods performed about the same for December. Breaking down the bias measures by the kind of business showed that the ad-hoc method overestimated sales for grocery and general merchandise stores, but underestimated for apparel stores. The overestimates and the underestimates nearly canceled each other out resulting in a very small net bias for the ad-hoc method in November. The absolute biases for the KPR and ad-hoc methods were fairly close (2.0% and 1.7%). Higher sales in December and the failure of the adjustment methods to match the calendar reports caused the magnitude of the bias to grow. The affect of not adjusting at all would be to overestimate sales by 9.3% in December. Both the ad-hoc method and the KPR adjustment improve upon the unadjusted bias. The ad-hoc and KPR underestimated sales for each of the three kinds of business in December. The hypothesis was that not adjusting and the ad hoc method would yield higher biases than the KPR adjustment. The unadjusted method clearly is the most biased, but it is hard to distinguish between the other two methods in December. The absolute bias in the table below is simply the sum of the absolute values of the calendar minus adjusted reports.

Table 5. Bias Measures for the Different Methods

Method	Measure	November Unweighted	December Unweighted
KPR Adjustment	Bias	1.3%	2.9%
	Absolute Bias	2.0%	6.0%
Ad Hoc	Bias	0.2%	2.9%
	Absolute Bias	1.7%	6.4%
No adjustment	Bias	6.3%	-9.3%
	Absolute Bias	6.5%	9.6%

Table 6 shows information about the variability of the methods. Although the bias was much smaller for the ad-hoc method in November compared to the KPR adjusted estimates, the standard error for the ad-hoc estimate was higher than that for the KPR method. This resulted in the two methods having relatively close MSE and MSR estimates in November with the ad-hoc method performing slightly better. In December, the biases for the ad-hoc and KPR adjustment methods are very close, but the standard deviation is higher for the ad-hoc method, which results in higher MSE and MSR estimates for the ad-hoc method compared to the KPR method. Period reports and their adjustments can be a cause of more variability in the estimates. The MSR estimates show that the extra effort to collect calendar data from the targeted period reporters greatly reduced the overall error in December.

Table 6. Bias and Variability Estimates for the Different Methods

METHOD	NOVEMBER			
	BIAS	**STD DEV**	**MSE**	**MSR**
Calendar	$0.00E00	$5.10E08	$2.60E17	**1.00**
KPR adjusted	$1.56E08	$5.13E08	$2.88E17	**1.11**
Ad hoc	$2.17E06	$5.17E08	$2.67E17	**1.03**
Unadjusted	$7.45E08	$4.78E08	$7.84E17	**3.02**

METHOD	DECEMBER			
	BIAS	STD DEV	MSE	MSR
Calendar	$0.00E00	$8.40E08	$7.06E17	**1.00**
KPR adjusted	$6.91E08	$8.01E08	$1.12E18	**4.31**
Ad hoc	$6.86E08	$8.15E08	$1.13E18	**4.37**
Unadjusted	-$1.19E09	$8.21E08	$2.09E18	**8.05**

The relative difference for each case under the different methods can be defined as $\frac{calendar - adjusted}{calendar}$. The distribution of relative differences in Table 7 shows that the mean and median differences are almost 0 for the KPR adjustment method. However, there are some cases that have large relative differences that are affecting the bias calculation. Without calling the respondents back, it is hard to know if their period reports or daily reports were causing the problem. The other methods were also affected by these cases with large relative differences. It is good to see that the mean and median differences are close to 0, but the extreme values and small sample size cast doubt over the final bias calculations.

Table 7. Unweighted Relative Differences (Rel. Diff) of the KPR Adjustment Method to the Calendar Reports

Label	N	Minimum	Median	Mean	Maximum
November Rel. Diff	84	-0.74	0.00	0.03	1.00
December Rel. Diff	84	-0.24	0.02	0.04	0.44

5. Discussion

It was interesting that the ad-hoc method performed slightly better than the KPR method for November and nearly the same in December. Why would such an obvious adjustment work as well as a more complex one? One reason may be that this method treats a day in the month exactly the same as any other day in the month. The adjustment method treated November 30th the same regardless if it was a Saturday or a Tuesday. This may be a reasonable assumption for these kinds of businesses during the holiday period. The KPR method, as described in the background section, treats all Sundays the same way throughout the year, but the KPR adjustment would be different if November 30th was a Saturday or a Tuesday. Perhaps daily sales are more similar within a month then throughout the year. The variability of the ad-hoc method is a concern. The low bias in November was a result of overestimating for one kind of business and underestimating for the others. This led to greater variability overall for this method.

Both adjustment methods greatly improved the estimates relative to not adjusting at all, which shows why identifying period reporters in the sample is very important. Not adjusting for period reporters leads to both larger bias and variance estimates. Adjusting the reports is an attempt to bring all the companies up to the same level. If there are some companies tabulated as 5-week reporters, while others are tabbed as 4-week reporters, then there will be wide fluctuation in the data.

The KPR adjustment method may be a source of error in the monthly sales levels among this group of targeted respondents, but it is clearly superior to not adjusting at all. However, because of the limitations of this study it is difficult to estimate the bias attributable to the KPR adjustment among period reporters as a whole. One desirable property of this method is that it appears to be consistent. This method tended to slightly underestimate sales for all kinds of business in both November and December. Perhaps the holiday effect for Christmas is underestimated and the time series staff should look into increasing the holiday effect for December. The KPR adjustment method also had the lowest variance estimates among the adjustment methods.

What if the analysts did not target these period reporters and attempt to get their true calendar month sales in November and December? The mean square ratios show that the effect on the estimates was small in November for this group of respondents. However, in December the extra effort clearly led to error reduction. If they had let the period reports stand and go through the current KPR adjustment method, then the mean square error would have been more than 4 times greater. This shows that the special data collection in December may well be worth the extra cost and burden because it improves the estimates. Hopefully the results for December are an aberration due to the high sales volume. It is likely that the adjustment error is more similar to November's results throughout the year. Adjusting for period reporters introduces some error into the estimates, but is a great improvement over not adjusting at all. If the error is shown to be small and consistent over the year, then the adjustment is a good solution.

To better estimate the errors caused by adjusting for period reporters, the next study should incorporate four design changes 1) a random sample of targeted companies, 2) daily data stored in the MARTS database so that edits can be performed quickly, 3) a larger sample size, and 4) a time period other than the holidays. Many of the data problems encountered in this study can be corrected by addressing the four items listed above. If it were a probability sample we could make better generalizations to the other period reporters. In this small-scale study, the daily data were stored in a text field which made it difficult to analyze and led to approximately 15 dropped units in the analysis. A larger sample size would be useful so that analysis could be done on the grocery, apparel and general merchandise separately, and not together as a whole. Using a time period other than the holidays would be useful to see if the response and bias differed throughout the year.

Acknowledgements

The author would like to thank Aref Dajani, Bill Davie and Jim Liu of the Census Bureau and Roger Tourangeau of the Joint Program in Survey Methodology for reviewing the paper; and Ruth Detlefsen, Lora Gillott, Scott Scheleur and Mike Shimberg of the Census Bureau for explaining technical details. This paper was written as a research project for the Total Survey Error class of the Joint Program in Survey Methodology.

References

Biemer, P. P. and Fresco, R. S "Evaluating and Controlling measurement Error in Business Surveys" in Cox et al. (eds) *Business Survey Methods,* 1995 New York: John Wiley & Sons, pp.257-281.

Brick, J.M and Keeter, S.A. "Using Data on Interruptions in Telephone Service As Coverage Adjustments". *Survey Methodology*, 22 (2), 1996, pp. 185-197.

Findley, D.F., Monsell, B.C., Bell, W. R., Otto, M.C., and Chen, B.C.(1998) "New Capabilites and Methods of the X-12-ARIMA Seasonal Adjustment Program". *Journal of Business and Economic Statistics*, Vol 16, Number 2, 127-177.

Freedman, S. R. "Control and Measurement of Specification Error in Establishment Surveys". Proceedings of the Survey Research Methods Section, American Statistical Association, 1988.

Groves, R. M., "*Survey Errors and Survey Costs*", New York: John Wiley & Sons, 1989.

Ponikowski, C. H. and Meily, S. A. "Controlling Response Error in an Establishment Survey". Proceedings of the Survey Research Methods Section, American Statistical Association, 1989.

Shiskin, J., Young, A. H., and Musgrave, J. C. U.S. Bureau of the Census. *The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Paper No. 15. (1967 revision)* U. S. Government Printing Office, Washington, D.C., 1967.

Soukup, R. J. and Findley, David F. "On the Spectrum Diagnostics Used by X-12-ARIMA To Indicate the Presence of Trading Day Effects After Modeling or Adjustment". Proceedings of the Joint Statistical Meetings, 2000.

Wilcox, D. W. "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work". *The American Economic Review Vol. 82 no.4 (Sep 1992),* 922-941.

Young, A. U.S. Bureau of the Census. *Estimating Trading Day Variation in Monthly Economic Time Series. Technical Paper No.12*. U.S. Government Printing Office Washington, D.C, 1965.