Seasonal Adjustment to Facilitate Forecasting: Empirical Results*

William R. Bell† Ekaterina Sotiris‡

Abstract

In this paper we consider how well seasonal adjustment methods (X-11 and ARIMA model-based), and certain of their variations, satisfy one objective of seasonal adjustment: facilitating short-term forecasting of nonseasonal movements in time series. We do this via an empirical study using a number of seasonal time series of major U.S. economic aggregates. For these series we examine how forecast accuracy is affected by the following choices: (1) alternative choices of simple models for forecasting the seasonally adjusted series (or trend estimates); (2) use of seasonally adjusted series versus trend estimates; (3) use of time series of unrevised versus time series of revised seasonally adjusted data; (4) use of X-11 versus ARIMA model-based adjustment; and (5) use of seasonally adjusted data in forecasting versus directly forecasting the unadjusted series.

Key Words: X-11 seasonal adjustment, ARIMA model-based seasonal adjustment, seasonal adjustment revision, trend estimation

1. Introduction

Seasonal adjustment is generally done without explicit, specific objectives. This is not to say, however, that a specific objective for seasonal adjustment has yet to be identified. According to Burman (1980, p. 321), the most common purpose of seasonal adjustment "is to provide an estimate of the current trend so that judgmental short-term forecasts can be made." Much earlier, Julius Shiskin, one of the developers of the popular X-11 method of seasonal adjustment, expressed similar sentiments when he said:

> A principal purpose of studying economic indicators is to determine the stage of the business cycle at which the economy stands. Such knowledge helps in forecasting subsequent cyclical movements and provides a factual basis for taking steps to moderate the amplitude and scope of the business cycle. ... In using indicators, however, analysts are perennially troubled by the difficulty of separating cyclical from other types of fluctuations, particularly seasonal fluctuations (Shiskin 1957).

In Shiskin (1961), he thus stressed the importance to short-term economic forecasting of having a comprehensive system of seasonally adjusted data.

Bell (1995) formally considered how model-based seasonal adjustment could be done to facilitate the forecasting of nonseasonal movements of time series. He showed that, from a theoretical perspective, this objective could be best served by not revising the concurrent seasonally adjusted data. What was lacking, though, was an empirical investigation to determine if this approach would realize any advantages

*Disclaimer: This report is released to inform interested parties of ongoing research and to encourage discussion. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau or the University of Maryland.

†U.S. Census Bureau, 4600 Silver Hill Road, SRD, Room 5K142A, Washington, DC 20233.

‡U.S. Census Bureau, Statistical Research Division, and University of Maryland.

when applied in practice, in comparison to similarly forecasting the usual seasonally adjusted series that is regularly revised.

This paper provides such an empirical investigation, but with a substantially expanded scope. In this investigation, we address several important questions related to how certain fundamental choices about seasonal adjustment affect subsequent forecasting results. Using data for a number of U.S. aggregate economic time series, we produce seasonally adjusted data and trend estimates, use these results to forecast the time series in various ways, and empirically assess the accuracy of the alternative forecasts. This allow us to examine how forecast accuracy is affected by the following:

1. alternative choices of simple models for forecasting the seasonally adjusted series (or trend estimates);

2. use of seasonally adjusted series versus trend estimates;

3. use of time series of unrevised versus time series of revised seasonally adjusted series or trend estimates;

4. use of X-11 versus canonical ARIMA model-based adjustment to calculate the seasonally adjusted series or trend estimates; and

5. use of seasonally adjusted series (or trend estimates) in forecasting versus directly forecasting the unadjusted time series.

In point 4, "canonical ARIMA model-based adjustment" refers to the seasonal adjustment method proposed by Hillmer and Tiao (1982) and Burman (1980), as is implemented in the TRAMO/SEATS sofware of Gomez and Maravall (1997), and more recently in the X-13-ARIMA/SEATS program at the Census Bureau (Monsell 2007). This approach uses ARIMA (autoregressive-integrated-moving average) time series models (Box and Jenkins 1970) and signal extraction to determine seasonal adjustments and trend estimates, in contrast to the empirical moving averages used in X-11 seasonal adjustment. See Ladiray and Quenneville (2001) for discussion of X-11. We implement X-11 adjustment here via the X-12-ARIMA program (Findley et al. 1998, U.S. Census Bureau 2009), which allows use of RegARIMA time series models (regression models with errors that follow ARIMA time series models) to estimate calendar effects, detect and adjust for outliers, and forecast extend the series for seasonal adjustment of recent data. We implement the X-11 procedure in X-12-ARIMA with default options, allowing the program to choose the seasonal and trend moving averages that are applied.

The plan of the paper is as follows. Section 2 describes how we conducted the empirical investigation: the series used, how the seasonal adjustment and forecasting were done, and how the forecasting results were assessed. Section 3 then discusses the empirical results relevant to points 1–5 given above. Section 4 closes by reviewing the general conclusions.

2. Overview of the Empirical Investigation

2.1 Data used

Table 1 shows the 19 time series used in this empirical investigation. All are monthly time series of major U.S. economic aggregates. For each time series, Table 1 shows

Table 1. Some Monthly U.S. Aggregate Macroeconomic Time Series

Census Bureau Series[1]	Dates	Calendar Effects[2]	ARIMA Model[3]
Total Housing Starts	1/59 - 1/86 - 12/08	td, lom	(0,1,1)
Total Building Permits	1/59 - 1/71 - 12/08	td	(0,1,1)
Total Exports (NAICS)	1/90 - 1/97 - 12/08	td	(2,1,0)
Total Imports (NAICS)	1/90 - 1/97 - 12/08	td	(0,1,3)
Oil Imports	1/86 - 1/96 - 12/08	td	(0,1,1)
Retail Sales of Department Stores (NAICS)	1/92 - 1/99 - 12/08	td, easter	(0,1,1)
Total Retail Sales (NAICS)	1/92 - 1/99 - 12/08	td, easter	(0,1,2)
Total Manufacturers Shipments (SIC)	1/58 - 1/70 - 12/00	td	(0,1,1)
Total Manufacturers Shipments (NAICS)	1/92 - 1/98 - 12/08	td	(0,1,1)
Total Manufacturers Inventories (SIC)	1/58 - 1/80 - 12/00	-	(3,1,0)
Total Manufacturers Inventories (NAICS)	1/92 - 1/99 - 12/08	-	(2,1,0)
Total Retail Inventories (NAICS)	1/92 - 1/99 - 12/09	tdstock	(0,1,1)
Bureau of Labor Statistics Series[1]			
Consumer Price Index (CPI-U)	1/50 - 1/86 - 12/08	-	(2,1,0)
Unemployment Rate	1/50 - 1/86 - 12/08	-	(2,1,0)
Total Nonfarm Employment (NAICS)	1/50 - 1/86 - 12/08	user-defined td	(2,1,0)
Federal Reserve Board Series[1]			
Industrial Production Index - historical (SIC)	1/50 - 1/70 - 12/85	td	(2,1,0)
Industrial Production Index - recent (NAICS)	1/86 - 1/96 - 12/09	td	(3,1,0)
Consumer Credit	1/50 - 1/87 - 12/09	tdstock	(2,1,0)
Money Supply (M1)	1/59 - 1/87 - 12/07	tdstock	(3,1,0)

1. Updated versions of the series, with revised historical data and further information (such as information on sampling and nonsampling errors), can be obtained from the agency web sites. For the Census Bureau series see http://www.census.gov/briefrm/staticindipage.html (particularly the "Program Overview" links there relevant to the above series.) For the BLS series see the overview tabs at http://www.bls.gov/cpi/, http://www.bls.gov/cps/, and http://www.bls.gov/ces/. For the Federal Reserve Board series follow links for the particular series from http://www.federalreserve.gov/econresdata/releases/statisticsdata.htm.

All series are logged except for housing starts, oil imports, unemployment rate, and nonfarm employment.

The North American Industry Classification System (NAICS) is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy. NAICS was developed under the auspices of the Office of Management and Budget (OMB), and adopted in 1997 to replace the Standard Industrial Classification (SIC) system. For many time series there is an overlap period for which tabulations are available under both NAICS and SIC, and for some series historical data were reconstructed under NAICS. The NAICS versus SIC distinction affects those time series indicated.

2. The calendar effect regression variables are described in the X-12-ARIMA program manual (U.S. Census Bureau 2009) except for the "user defined" trading-day regressors for the nonfarm employment series. These were used to model trading-day effects specific to that series, and were provided by Stuart Scott of the Bureau of Labor Statistics.

3. For all series the seasonal part of the model was (0,1,1). The column above thus gives the orders of only the nonseasonal parts of the models.

three dates. The first and last dates listed are the starting and ending dates of the data from each series used in the analysis here. The second date listed is the date of the first forecast origin used here. Data between this first forecast origin and the ending date were used in assessing the forecast performance of the various procedures compared here, though in a somewhat involved manner that is discussed in Section 2.4. The last two columns of Table 1 provide information on RegARIMA models developed for the series, as discussed in Section 2.2. Finally, the notes to Table 1 indicate the sources of the series used here, where additional information about the series can be found (e.g., on sampling and nonsampling error in the series), and whether or not logarithms were taken as part of the modeling.

Use of such major aggregate series has advantages and disadvantages for this investigation. Advantages include the fact that these series are fairly recognizable, making the results more interpretable than they would be with more disaggregate series. Also, such aggregate series tend to be more stable than very disaggregate series, facilitating their modeling, seasonal adjustment, and subsequent forecast analysis. In particular, these series tend to have very low levels of sampling error (and some have no or essentially no sampling error). The major disadvantage to use of these aggregate series is that such series are not generally directly seasonally adjusted. Instead, more detailed sub-aggregate series are seasonally adjusted, and these results are aggregated to produce the seasonally adjusted aggregate series. Thus, our direct adjustment of aggregate series here does not conform to standard practice. Our choice to analyze aggregate series thus involved a trade-off, and leaves open questions about how results from the analyses here might differ from similar results that could be obtained using more disaggregate series. This could provide a topic for further study.

2.2 RegARIMA modeling

The first step in the analysis of each series was to develop a RegARIMA model for the series that could be used in seasonal adjustment via both X-12 (for X-11 adjustment) and X-13 (for ARIMA model-based adjustment using the SEATS procedure). The RegARIMA models developed had the general form

$$\begin{aligned} y_t &= x_t'\beta + z_t \\ \phi(B)(1-B)(1-B^{12})z_t &= \theta(B)(1-\Theta B^{12})a_t \end{aligned} \tag{1}$$

where y_t is the observed time series or, more commonly, its logarithms, x_t is a vector of regression variables at time point t, β is the corresponding vector of regression parameters, and the regression residuals z_t follow the seasonal ARIMA model given in (1). In the latter, $\phi(B)$ and $\theta(B)$ are (nonseasonal) AR and MA polynomials in the backshift operator B ($Bz_t = z_{t-1}$), $(1-\Theta B^{12})$ is the seasonal MA polynomial, $(1-B)$ and $(1-B^{12})$ are the nonseasonal and seasonal differencing operators, and a_t is white noise ($i.i.d.$ $N(0,\sigma_a^2)$). Note from (1) that, for every series, the model used has one nonseasonal and one seasonal difference, no seasonal AR operator, and a seasonal MA operator of order 1. These restrictions on the models seemed appropriate for these series, and they also facilitated the ARIMA model-based seasonal adjustment. In the $(p,d,q)(P,D,Q)_{12}$ notation of Box and Jenkins (1970), the models used were thus all of the form $(p,1,q)(0,1,1)_{12}$, with the values of p and q varying across the series.

The regression variables in x_t are used here for modeling calendar effects (trading-day or holiday effects) when present, and also for modeling additive outliers and level shifts when such outliers are detected. Table 1 provides a summary of the models used, indicating what calendar effects were included in the model for each series, whether logs were taken, and the orders p and q of $\phi(B)$ and $\theta(B)$.

In developing the RegARIMA models, we used all the available data for a series, from the starting date to the ending date shown in Table 1. A primary reason for this was to detect and estimate outlier effects, as well as estimate any calendar effects, for the full series. Each series was then adjusted for any calendar and outlier effects in its model, thus computing an estimate of the regression residual series z_t. This adjusted series was then used in the forecasting study. The adjustment for outliers over the whole series, including the forecast period, was done to try to avoid large outliers having undue influence on the forecast accuracy measures. The adjustment for calendar effects over the whole series simplified forecasting considerably without adversely affecting the forecast accuracy comparisons. While re-estimating the calendar effects as part of the forecasting would have been more realistic, this would typically have added only a small amount to the forecast error, especially at the more advanced lead times. More importantly, this addition to forecast error would have been the same for all the different forecasts produced for a given series, and so should not have appreciably affected the forecast accuracy comparisons.

2.3 Producing forecasts using seasonally adjusted data

Having thus obtained (an estimate of) the calendar and outlier adjusted series z_t, suppose we have these observations for some time points denoted $t = 1, \ldots, n$. We now seasonally adjust z_t using this span of data by either the X-11 or ARIMA model-based methods. Either way, we estimate an additive decomposition, $z_t = S_t + N_t$. We can then take the seasonally adjusted series (denote this $\hat{N}_t$), forecast it ahead beyond time point n, and take the results as forecasts of N_t. This forecasting can be done by relatively simple methods, such as using simple nonseasonal ARIMA models (as discussed in Section 2.5). Given also forecasts of S_t, we could then add these to the forecasts of N_t to obtain forecasts of z_t.

How, in practice, would forecasts of S_t be obtained? Fortunately, both the X-11 procedure of the X-12-ARIMA program, and the SEATS procedure of the X-13 program, will produce (one year of) forecasts of S_t. These forecasts are traditionally called "projected seasonal factors," and, if made available to data users, would allow them to forecast the observed series as just discussed. In addition, Bell (1995) points out that, for ARIMA model-based adjustment using the model (1), the forecasts of S_t beyond one-year-ahead would simply repeat the year of projected factors, something data users could certainly do. This is the approach we shall use here.

The forecasting approach just described can equally well be applied to time series of trend estimates instead of seasonally adjusted data. This uses the three component decomposition, $z_t = S_t + T_t + I_t$, where T_t is the trend and I_t is the irregular component. Let $\hat{T}_t$ denote the trend estimates. Note that if, as is customary, the irregular component I_t is assumed to be white noise, then its forecasts would be zero, and so forecasts of N_t could be taken as forecasts of T_t and vice versa. In fact, if forecasts of N_t and T_t were obtained directly from the model (1) and its decomposition into component models, with I_t modeled as white noise, these forecasts of N_t and T_t would be identical. (Forecasts made from simpler approaches using $\hat{N}_t$ or $\hat{T}_t$ as data would not be identical, however.)

In the remainder of the paper, when we mention forecasting with seasonally adjusted data, it should be understood that we could use trend estimates instead.

Section 3 presents, for the series in Table 1, empirical results on forecast accuracy for forecasts obtained both ways. In fact, we have a combination of eight alternative approaches to forecasting depending on whether we use (*i*) seasonally adjusted data or trend estimates, (*ii*) unrevised versus revised seasonally adjusted data, and (*iii*) X-11 or ARIMA model-based seasonal adjustment. The unrevised versus revised distinction is discussed in Section 2.4.

2.4 Assessing forecast accuracy

Suppose we take observations of our series z_t for some time points labelled $t = 1, \ldots, n$ and forecast z_t using the procedure described above to produce forecasts, $\hat{z}_n(\ell)$, for some set of forecast leads $\ell = 1, \ldots, L$. We shall use $L = 24$ here so there will be two years of forecasts. If we held out some observed data beyond time n, and so have observations $z_{n+\ell}$ for $\ell = 1, \ldots, L$, we can compare the forecasts to the held-out data, producing forecast errors

$$e_n(\ell) = z_{n+\ell} - \hat{z}_n(\ell) \qquad \text{for} \qquad \ell = 1, \ldots, L. \tag{2}$$

For each of the series listed in Table 1, we start this process with n corresponding to the first forecast origin date. We then update the forecast origin to the next time point, $n + 1$, redo the seasonal adjustment or trend estimation (using data $z_1, \ldots, z_{n+1}$), use these results to produce forecasts $\hat{z}_{n+1}(\ell)$ from origin time $n + 1$, and compute forecast errors $e_{n+1}(\ell) = z_{n+1+\ell} - \hat{z}_{n+1}(\ell)$ for $\ell = 1, \ldots, L$. We continue updating the forecast origin and calculating the forecast errors as in (2) until we get to $\bar{n} - L$, where time point $\bar{n}$ corresponds to the ending date of the available data.

To clarify this process and explain it further, consider how it applies to the time series of total housing starts, the first series listed in Table 1. For this series, time point 1 corresponds to 1/59, and the first forecast origin (time point n) corresponds to 1/86. Thus, to start, we seasonally adjust the data from 1/59 – 1/86. We do not retain the seasonally adjusted data for this entire span, however. Rather (for a reason of comparability to the approach discussed in the next paragraph), to construct the "revised seasonally adjusted series," we drop the first eight years of the results, and retain the seasonally adjusted data only for 1/67 – 1/86. We then use this span of adjusted data to forecast two years ahead, up through 1/88, and calculate forecast errors over this two-year span. We then update the forecast origin to 2/86 (corresponding to time $n + 1$), seasonally adjust the data from 1/59 – 2/86, save the results for 1/67 – 2/86, use this data to forecast through 2/88, and compute forecast errors over this span. We then update the forecast origin to 3/86, etc. The last forecast origin used is 12/06, from which we forecast through 12/08, the end of the series, and calculate the corresponding forecast errors. The analogous procedure applies to forecasting using time series of revised trend estimates.

To do forecasting and calculate forecast errors using unrevised seasonally adjusted data (or trend estimates) is more involved. We start by seasonally adjusting only the data from 1/59 – 1/67 (eight years plus one month). From this, we retain only the last adjusted value, this being for 1/67. We then seasonally adjust the data from 1/59 – 2/67, retaining the adjusted value only for 2/67. We continue this way through the first forecast origin, 1/86, thus building up a time series of unrevised seasonally adjusted data from 1/67 – 1/86. We then forecast the unrevised seasonally adjusted series from origin 1/86, and compute forecast errors from 2/86 – 1/88. We then seasonally adjust the data from 1/59 – 2/86, retaining the adjusted value only for 2/86, giving us now the time series of unrevised seasonally adjusted data from 1/67 – 2/86. We then forecast from origin 2/86 and compute forecast errors for 3/86 – 2/88. We then seasonally adjust the data from 1/59 – 3/86, and so on.

To measure forecast accuracy, we shall compute the empirical forecast mean squared error (MSE) and corresponding root mean squared error (RMSE) defined as

$$\text{MSE} \; = \frac{1}{K}\sum_{k=n}^{\bar{n}-L}[e_k(\ell)]^2 \qquad \text{RMSE} \; = \sqrt{\text{MSE}}$$

where $K = \bar{n} - L - n + 1$ is the number of terms in the sum. We do this calculation in the scale in which the modeling and seasonal adjustment were done which, for most of the series in Table 1, is the log scale. For these (logged) series, 100 times the RMSE can be interpreted as an approximate percentage error for the series on the original (unlogged) scale, as long as the RMSE is not too large. To see why, let Z_t be the series on the original scale (so $z_t = \log(Z_t)$), and let $\hat{Z}_n(\ell) = \exp(\hat{z}_n(\ell))$ be its forecast for some time point t. Then a Taylor series linearization of the (relative) forecast error on the original scale gives $(Z_{n+\ell} - \hat{Z}_n(\ell))/Z_{n+\ell} = 1 - \exp[\hat{z}_n(\ell) - z_{n+\ell}] \approx z_{n+\ell} - \hat{z}_n(\ell)$. Since RMSE is a measure of the magnitude of $z_{n+\ell} - \hat{z}_n(\ell)$, it is approximately a measure of the magnitude of the relative forecast error $(Z_{n+\ell} - \hat{Z}_n(\ell))/Z_{n+\ell}$. This fact aids interpretation of the empirical RMSE results for the logged series.

2.5 Simple models for forecasting the seasonally adjusted series

Users of seasonally adjusted data would most likely use simple methods of forecasting, such as judgmental extrapolation. Though we cannot study use of personal judgment here, we can roughly mimic the kind of thing users might do by using simple models to forecast the seasonally adjusted series or trend estimates. Let w_t be either the seasonally adjusted series or the time series of trend estimates. We shall use the following alternative models in forecasting:

- No slope:

$$(0,1,1): \qquad (1-B)w_t = (1-\theta B)a_t$$

- Constant slope:

$$\begin{aligned} &(0,1,0)c: \qquad (1-B)w_t = c + a_t \\ &(0,1,1)c: \qquad (1-B)w_t = c + (1-\theta B)a_t \end{aligned}$$

- Adaptive slope:

$$(0,2,1): \qquad (1-B)^2 w_t = (1-\theta B)a_t$$

All these models produce linear forecast functions, but they differ in how the slope of the forecast function is determined. We call the (0,1,1) model the "no slope" model because it produces constant forecasts for all future time points. In fact, the level of the forecasts is an exponentially weighted moving average of the observations (Box and Jenkins 1970, p. 145), thus, it corresponds to the forecasting method of "exponential smoothing." The (0,1,0)c and (0,1,1)c models produce linear forecast functions with slope given by the estimate of the trend constant c. Forecasts from these two models differ because their estimates of c will differ, and because the

MA(1) term in the (0,1,1)c model can affect the lead-1 forecast (for $\theta \neq 0$). We call the (0,2,1) model the "adaptive slope" model because it allows the slope of its forecast function to change (even with θ known) as the forecast origin changes. In fact, for this model, the slope of the forecast function will be an exponentially weighted moving average of the observations of $(1-B)w_t$, as can be seen from the fact that forecasts of w_t are obtained by accumulating forecasts of $(1-B)w_t$, and $(1-B)w_t$ follows the (0,1,1) model. For $\theta \geq 0$, the (0,2,1) model is the "local linear trend" model of Harvey (1989, pp. 37 and 45).

There are several connections among these models. The (0,1,1)c reduces to the (0,1,1) when $c = 0$ and to the (0,1,0)c when $\theta = 0$. The (0,2,1) reduces to the (0,1,0)c when $\theta = 1$. For this reason, differences in forecasts between these "connected" models can, depending on the estimated parameter values, be small. We estimate the parameters of the models by maximum likelihood using the X-12-ARIMA program. For the (0,2,1) model, we shall also provide results from forecasts when θ is fixed at 0.0 and when θ is fixed at 0.5. Since $\theta = 1$ produces the (0,1,0)c model, we thus cover the full range of nonnegative values of θ, which is the parameter that determines the decay rate in the exponentially weighted moving average of $(1-B)w_t$ that provides the slope of the linear forecast function.

With the four estimated models given above, plus the two additional versions of the (0,2,1) model (from setting θ to 0.0 or 0.5), we have six models to use in forecasting the seasonally adjusted series or trend estimates. Combining these with the 8 versions of seasonally adjusted series or trend estimates noted in Section 2.2, we thus have 48 alternative forecasts for each time series. The next section compares RMSEs between several of these alternatives for each of the time series in Table 1, to address the five points noted in the Introduction.

3. Empirical Results

Each of the five subsections here discusses the empirical results of forecast RMSE comparisons aimed at addressing one of the five points noted in the Introduction. For each of the 19 series of Table 1, we obtained forecast RMSEs from the sets of alternative seasonally adjusted series or trend estimates mentioned in Section 2.3, and for each of the six forecast models discussed in Section 2.5. There are many cases where the differences between forecast RMSEs from different models are small, and certainly of no practical importance. This is also true for many of the forecast RMSE differences between the use of different seasonally adjusted series or trend estimates in forecasting. In cases where there does appear to be a practically meaningful RMSE difference, we caution that we have no check on the statistical significance of these differences. While this is a limitation, since our major conclusions derive mostly from the large number of cases where the observed RMSE differences are of no practical significance (rendering the question of statistical significance moot), this is not a serious limitation overall.

3.1 How does forecast accuracy depend on alternative choices of simple models for forecasting the seasonally adjusted series?

Here we summarize the results of forecast RMSE comparisons across the six forecast models within a given set of forecasts defined by using ARIMA model-based (SEATS) seasonal adjustments and trend estimates. We thus selected the best forecast model (lowest RMSE) for each of these four sets of estimates: (1) unrevised trend, (2) revised trend, (3) unrevised seasonally adjusted, and (4) revised seasonally adjusted. Table 2 lists the selected best model(s) by series and estimate type.

Table 2. Best Simple Forecasting Models, by Series and Type of Estimate

	Trend		Seasonally Adjusted	
Series	**Unrevised**	**Revised**	**Unrevised**	**Revised**
Total Housing Starts	(0,1,1)			
Total Building Permits	(0,1,1)			
Oil Imports	(0,1,1)			
Retail Sales of Dept. Stores	(0,1,1)			
Unemployment Rate	(0,1,1)			
Industrial Production Index recent (NAICS)	(0,1,1)c short leads (0,1,1) long leads			
Industrial Production Index historical (SIC)	(0,1,1)c			
Total Exports (NAICS)	(0,1,1)c			
Total Imports (NAICS)	(0,1,1)c			
Total Retail Sales	(0,1,1)c			
Manufacturers Shipments (SIC)	(0,1,1)c			
Manufacturers Shipments (NAICS)	(0,1,1)c			
Total Retail Inventories	(0,1,1)c			
Consumer Price Index	(0,1,1)c	(0,1,1)c	(0,2,1)	(0,2,1)
Manufacturers Inventories (SIC)	(0,2,1)* short leads (0,1,1) long leads			
Manufacturers Inventories (NAICS)	(0,2,1)* short leads (0,1,1) long leads			
Nonfarm Employment	(0,2,1)* short leads (0,1,1)c long leads			
Money Supply (M1)	(0,2,1)*			
Consumer Credit	(0,2,1)	(0,2,1)*	(0,2,1)	(0,2,1)*

* For these (0,2,1) models, θ is fixed at .5. For the others, θ is estimated.

Note: Blanks in the last three columns mean that the best model is the same as when using unrevised trend estimates.

No single model was best for all series, but the series can roughly be grouped into three categories by the type of model that worked best: the no slope model, (0,1,1); the constant slope models, (0,1,1)*c* and (0,1,0)*c;* and the adaptive slope model, (0,2,1). As shown in Table 2, for 17 of the 19 series, the best model(s) for a given series were best when using all four sets of estimates.

For 15 of the 19 series, we were able to find a model which performed best across all lead times. For example, for the Total Imports series, forecast RMSEs (using the unrevised trend estimates) were consistently lowest for the constant slope models – see Figure 1. (Note there that the RMSE curves for the (0,1,0)*c* and (0,1,1)*c* models essentially coincide, so only the former is visible, and either of these two models can effectively be regarded as "best".) For the other four series, however, performance of the models varied by forecast lead time. For example, for the Non-Farm Employment series, in forecasting using unrevised trend estimates, the (0,2,1) model did best at early leads, but the (0,1,1)c model was best at later leads – see Figure 2. (Note there that the RMSE curve for the (0,2,1) model with θ estimated almost coincides with that for the (0,2,1) model with $\theta = 0.5$. Similarly, the RMSE curves for the (0,1,0)c and (0,1,1)c models almost coincide, so that the former almost completely obscures the latter.)

While we selected the best model or pair of models for each of the series, for many series (such as Total Imports) there were other models which had practically identical forecast RMSEs to those of the best model. In fact, for four out of five series where the (0,1,1) model did best, both the (0,1,1)c and the (0,1,0)c models performed similarly. This is not terribly surprising given the connection between these models (when $c = 0$) noted in Section 2.5.

3.2 How is forecast accuracy affected by the use of seasonally adjusted series versus trend estimates?

In this and the succeeding sections, we focus mostly on forecast RMSEs obtained with the best models, in order to make comparisons of forecast accuracy across the use of different seasonally adjusted series or trend estimates in forecasting. In this section we examine how forecast accuracy (with the best models) is affected by the use in forecasting of seasonally adjusted series versus use of trend estimates. We made such comparisons using ARIMA model-based (SEATS) results, both unrevised and revised.

The general conclusion is that, for most series (16 out of 19), there was no appreciable difference in forecast RMSEs for forecasts made using seasonally adjusted series versus trend estimates, both for unrevised and revised estimates. For example, for the Manufacturers' Inventories (NAICS) series, forecast RMSEs from using unrevised seasonally adjusted series versus unrevised trend estimates were essentially the same across all forecast leads (Figure 3).

One example where forecast performance from using seasonally adjusted series appeared to differ from that when using trend estimates is the Retail Inventories series (for which the best model was the (0,1,1)c.) For this series, Figure 4 shows that forecasts using trend estimates appear possibly more accurate for leads of 1 and 2 months, but for leads over 5 months forecasts using seasonally adjusted estimates appear more accurate. (As noted earlier, such conclusions are subject to concerns about the statistical significance of such differences, which are particularly relevant here for leads ≤ 12 where the differences are small. Also, note that the RMSE curves for use of unrevised seasonally adjusted data essentially coincide with those for use of revised seasonally adjusted data, though the results are not identical. The same is true of the RMSE curves for use of unrevised trend estimates and use of revised trend estimates.) The difference between the RMSEs from using trend estimates versus seasonally adjusted data at lead 1 month is around .5%, and by lead 4 the difference has vanished. At the maximum forecast lead of 24 months, the RMSEs from using trend estimates and from using seasonally adjusted data are about 7% and 4%, respectively.

3.3 How is forecast accuracy affected by the use of time series of unrevised versus revised seasonally adjusted series?

Using the best forecasting models, we compared the RMSEs from using unrevised trend estimates to those from using revised trend estimates, as well as those from using unrevised seasonally adjusted data to those from using revised seasonally adjusted data. This was done with SEATS estimates.

The general conclusion is that, for most series (15 out of 19), there was no appreciable difference in RMSEs for forecasts made using unrevised versus using revised estimates, both trend and seasonally adjusted (with the best model). For this reason, we do not provide a plot to illustrate these results. (Such a plot would look like Figure 3.)

One example where forecast performance with unrevised versus revised estimates appeared to differ is the Manufacturers' Inventories (NAICS) series. For this series, we have two competitors for best model: (0,2,1) with $\theta = 0.5$, which performs best at low leads, and (0,1,1), which performs best at higher leads. For the (0,1,1) model, unrevised and revised estimates lead to forecasts with essentially the same accuracy. However, with the (0,2,1) model with $\theta = 0.5$, revised estimates show slightly lower forecast RMSEs than do the the unrevised estimates. This is true for both trend and seasonally adjusted estimates. Figure 5 shows the results for trend estimates; results are similar for seasonally adjusted estimates. The difference in the RMSEs grows with forecast lead, with the difference being around 0.7% at lead 24 months. Statistical significance of this difference is questionable, however.

3.4 How is forecast accuracy affected by the use of X-11 versus canonical ARIMA model-based adjustment?

Using the best forecasting models, we compared forecast RMSEs when using X-11 versus SEATS to produce each of the four types of estimates studied. We found that for 17 of our 19 series, there was no appreciable difference in RMSEs for forecasts made using X-11 versus SEATS estimates, for all four types of estimates.

One series for which forecast performance with X-11 versus SEATS estimates showed some difference is the Consumer Price Index series. For this series, forecasts made using SEATS trend estimates (both revised and unrevised) had slightly lower RMSEs than did those made using X-11 trend estimates. The difference in RMSEs grows with forecast lead time to around 0.33% at 24 months (Figure 6). Statistical significance of this difference is again questionable.

3.5 How is forecast accuracy affected by the use of seasonally adjusted series in forecasting versus directly forecasting the unadjusted time series?

Finally, we compared RMSEs for direct forecasts of the observed series to RMSEs for forecasts made using seasonally adjusted or trend estimates. Here we analyzed 18 series (omitting Total Nonfarm Employment due to a computer program problem). For each of the 18 series, forecasts were made using the four sets of SEATS estimates – unrevised trend, revised trend, unrevised seasonally adjusted, and revised seasonally adjusted – and their accuracy was compared to that of direct forecasts of the observed series. The direct forecasts were obtained with the seasonal ARIMA models listed in Table 1 (recall that estimated regression effects were previously removed from the data using the RegARIMA models).

For 13 of the 18 series, forecast RMSEs from directly forecasting the observed series were possibly higher than when using seasonally adjusted or trend estimates in forecasting. Figure 7 presents an example of one such series, Building Permits. Note that the RMSE curves in Figure 7 for all four sets of SEATS estimates essentially coincide, and are lower than the RMSE curve for the direct forecasts.

For 2 of the 18 series, forecast RMSEs from directly forecasting the observed series were possibly lower than when using seasonally adjusted or trend estimates in forecasting. Figure 8 presents an example of such a series, Manufacturers' Shipments (SIC). Notice, though, that RMSEs for the direct forecasts appear lower only for leads beyond 12 months. Notice also that the four RMSE curves for the forecasts obtained using the four SEATS estimates almost coincide.

For 3 of the 18 series, comparing forecast RMSEs suggested no clear advantage to directly forecasting the observed series versus using seasonally adjusted or trend estimates in forecasting. While the RMSE curves did not effectively coincide, their differences were not large and suggested no advantage to either approach.

4. Conclusions

This study aimed at addressing five important questions related to use of seasonally adjusted data or trend estimates in forecasting. Like any empirical study, this one is subject to certain limitations. These would include the limited number and type of time series examined, and the lack of statistical significance checks on RMSE differences. The latter limitation becomes moot, however, when differences in RMSEs, even if potentially statistically significant, are clearly not practically significant.

We started by examining, for each of our 19 time series, how forecast accuracy, as measured by RMSEs, varied across six simple models used to forecast the seasonally adjusted data or trend estimates. We found that some of the six models often tended to perform similarly to one another for a given series, while others (particularly when they determined the slope of the forecast function differently) could produce substantially different forecast RMSEs. (Figures 1 and 2 provide illustrative examples.) We also found that the best model choices varied across series, leading to three groups of series according to which model worked best: the no slope model, (0,1,1); the constant slope models, (0,1,1)c and (0,1,0)c; or the adaptive slope model, (0,2,1). Clearly, how one forecasts seasonally adjusted series (or trend estimates) can have important effects on the accuracy of the forecsts then obtained for the observed series.

On the other hand, when we examined how forecast accuracy varied across alternative choices of which seasonally adjusted series or trend estimates to use in forecasting (seasonally adjusted series versus trend estimates, unrevised versus revised, X-11 versus ARIMA model-based), for most series we found no indication that these choices made any appreciable difference to forecast accuracy. In the relatively few cases where we observed apparent differences in forecast RMSEs, these differences were often practically small, and their statistical significance, which we could not assess, was also questionable.

Finally, we compared the accuracy of forecasts made directly of the observed series using seasonal ARIMA models with the accuracy of forecasts obtained by forecasting seasonally adjusted series or trend estimates, and combining these with forecasted seasonal factors. (Note that we effectively used RegARIMA models in forecasting the observed series, since these models were used to first estimate and remove calendar and outlier effects before doing the ARIMA forecasting. In producing the seasonally adjusted data and trend estimates, we also used the RegARIMA models to remove these same estimated effects.) Here we found some apparent differences in the forecast RMSEs, and these more often, but not always, tended to favor forecasting using the seasonally adjusted data or trend estimates. This is at least suggestive that seasonal adjustment can, indeed, facilitate forecasting.

Acknowledgment: We thank Brian Monsell for providing some modifications to the X-12-ARIMA program that greatly facilitated some of the calculations presented.

REFERENCES

Bell, William R. (1995), "Seasonal Adjustment to Facilitate Forecasting–Arguments for Not Revising Seasonally Adjusted Data," *American Statistical Association, Proceedings of the Business and Economics Statistics Section*, 268-273.

Box, G.E.P. and Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.

Burman, J. Peter (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society, Series A*, **143**, 321-337.

Findley, David. F., Monsell, Brian C., Bell, William R., Otto, Mark C., and Chen, Bor-Chung, (1998) "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program (with discussion)," *Journal of Business and Economic Statistics*, **16**, 127-177, technical report version available online at http://www.census.gov/ts/papers/jbes98.pdf.

Gomez, Victor and Maravall, Agustín (1997), "Programs TRAMO and SEATS: Instructions for the User (Beta Version: June 1997)," Working Paper 97001, Ministerio de Economía y Hacienda, Dirección General de Análisis y Programación Presupuestaria, Madrid, Spain, available online at http://www.bde.es/webbde/es/secciones/servicio/software/tramo/manualdos.pdf.

Hillmer, Steven C., and Tiao, George C. (1982), "An ARIMA-Model-Based Approach to Seasonal Adjustment," *Journal of the American Statistical Association*, **77**, 63-70.

Ladiray, Dominique and Quenneville, Benoit (2001), *Seasonal Adjustment with the X-11 Method*, New York: Springer-Verlag.

Monsell, Brian C. (2007), "The X-13A-S Seasonal Adjustment Program," *Proceedings of the 2007 Federal Committee On Statistical Methodology Research Conference*, [Online], available at http://www.fcsm.gov/07papers/Monsell.II-B.pdf.

Shiskin, Julius (1957), "Electronic Computers and Business Indicators, *Journal of Business*, **30**, 219-267, available at http://www.nber.org/books/juli57-1.

Shiskin, Julius (1961), "Statistics for Short-Term Economic Forecasting," chapter 18 of *Business Cycle Indicators, Vol. 1*, ed. Geoffrey H. Moore, National Bureau of Economic Research, originally published by UMI, volume available at http://www.nber.org/books/moor61-1.

U.S. Census Bureau (2009), "X-12-ARIMA Reference Manual, Version 0.3," Time Series Staff, Statistical Research Division, Washington, DC, available at http://www.census.gov/srd/www/x12a/.

Figure 1. *Total Imports*: RMSEs for forecasts obtained from various models (using unrevised trend estimates)

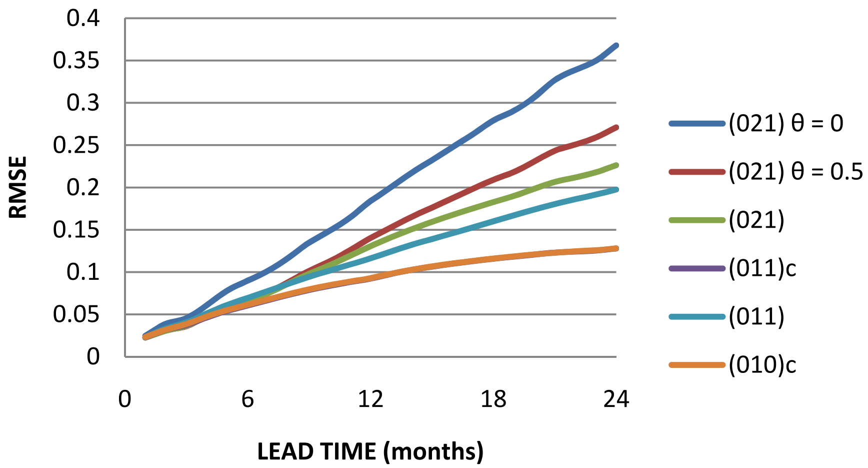

Figure 2. *Non-Farm Employment*: RMSEs for forecasts obtained from various models (using unrevised trend estimates)

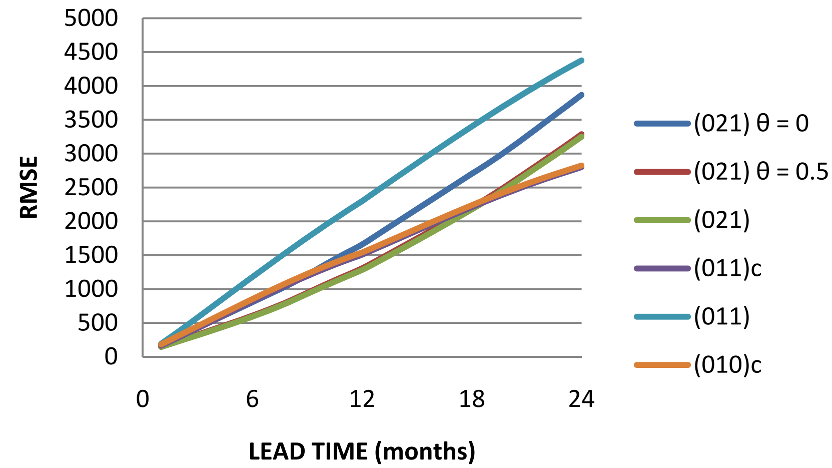

Figure 3. *Manufacturers Inventories (NAICS)*: RMSEs for forecasts obtained using Trend vs. Seasonally Adjusted series (unrevised estimates with best model)

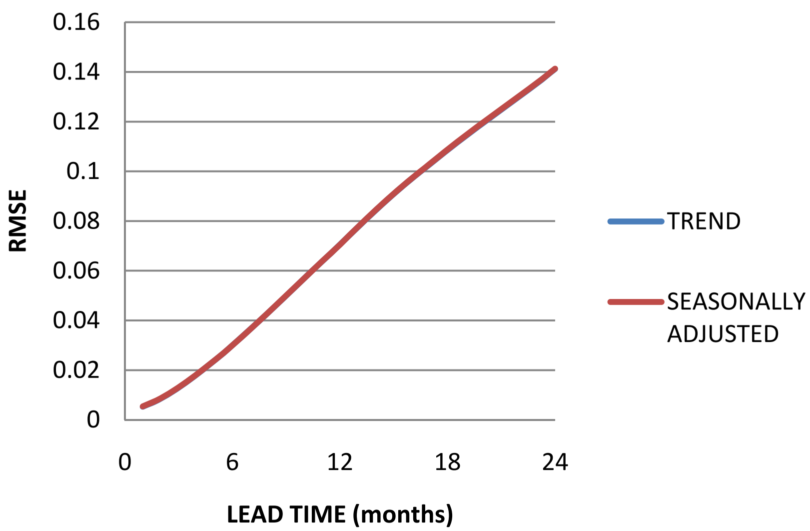

Figure 4. *Retail Inventories*: RMSEs for forecasts obtained using Trend vs. Seasonally Adjusted series (unrevised and revised with best model)

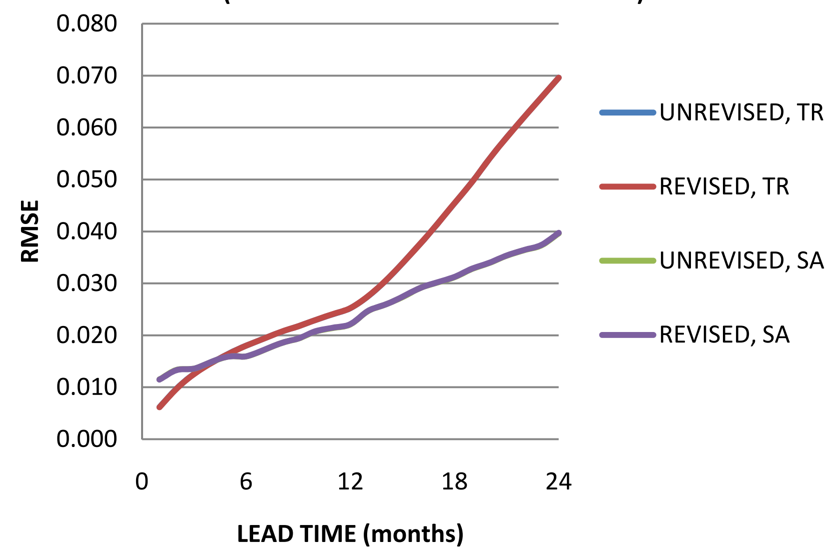

Figure 5. *Manufacturers Inventories* (*NAICS*): RMSEs for forecasts obtained using Unrevised vs. Revised trend estimates (with best model)

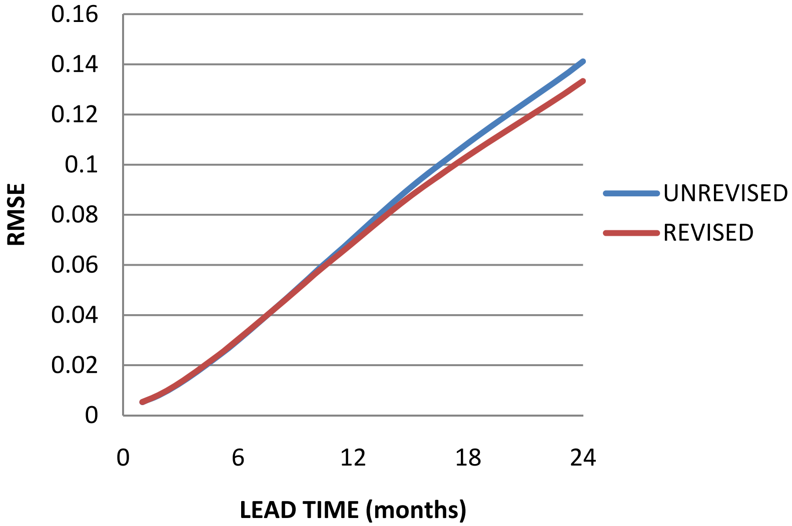

Figure 6. *Consumer Price Index*: RMSEs for forecasts obtained using X11 vs. SEATS revised trend estimates (with best model)

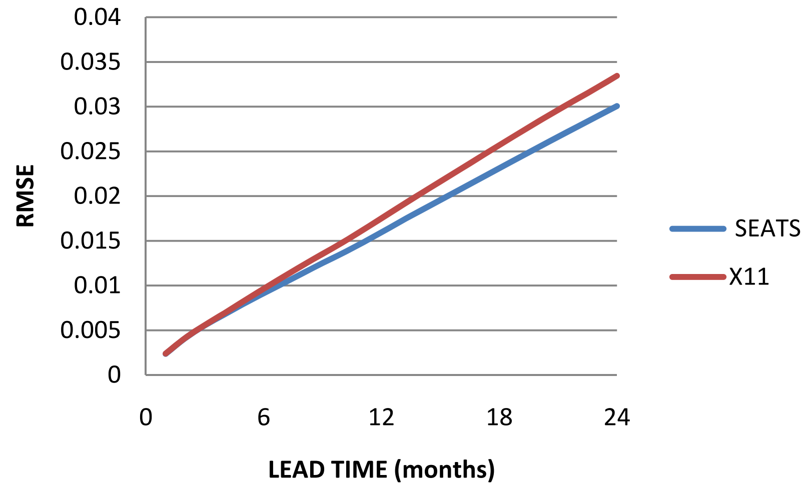

Figure 7. *Building Permits*: RMSEs for direct forecasts of the Original Series vs. forecasts made using Seasonally Adjusted or Trend estimates (with best model)

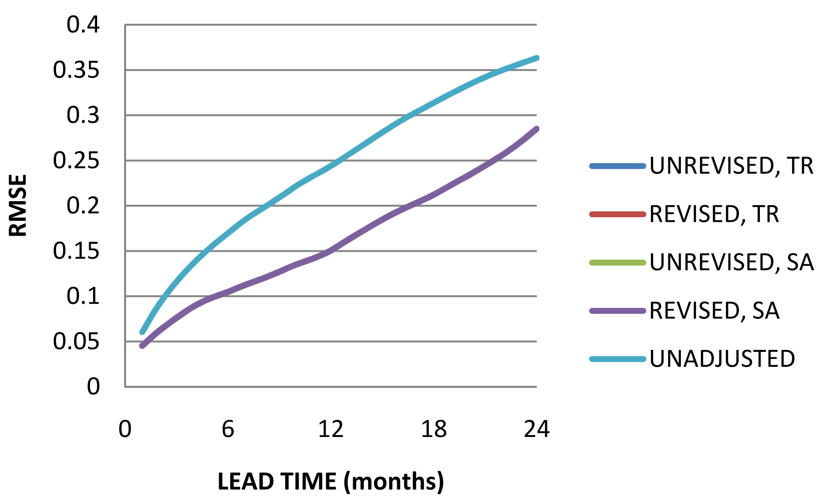

Figure 8. *Manufacturers' Shipments (SIC)*: RMSEs for direct forecasts of the Original Series vs. forecasts made using Seasonally Adjusted or Trend estimates (with best model)