A Review of the Problem of Seasonal Adjustment Variances

William R. Bell

U.S. Census Bureau

William.R.Bell@census.gov

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Disclaimer:

The views expressed here are those of the author and not necessarily those of the U.S. Census Bureau.

Outline

1. Background and notation

2. Two choices of “seasonal adjustment variances”

3. Contributions to seasonal adjustment error
   a. from the time series components
   b. from forecast extension error
   c. from error in estimating model parameters

4. Variances for X-11 seasonal adjustments

5. Conclusions

Background and Notation

Additive "*seasonal* + *nonseasonal* + *sampling error*" decomposition:

$$y_t = S_t + N_t + e_t \qquad t = 1, \ldots, n$$

If y_t is not subject to sampling error, drop e_t.

Extend with $N_t = T_t + I_t$:

$$y_t = S_t + T_t + I_t + e_t.$$

Note that these additive decompositions are typically used after taking logarithms of the data.

Generic RegComponent model for examples:

$$y_t = Y_t + e_t$$

where

$$(1 - B)(1 - B^{12})[Y_t - x_t'\beta] = (1 - \theta_1 B)(1 - \theta_{12} B^{12}) b_t$$

$$Y_t = S_t + N_t \qquad \text{(canonical decomposition)}$$

ARMA model for e_t (when present)

Estimators of components (need modifications for regression terms $x_t'\beta$):

$$\hat{S}_t = \omega_S(B) y_t = \sum_j \omega_{S,j} y_{t-j}$$

$$\hat{N}_t = \omega_N(B) y_t = \sum_j \omega_{N,j} y_{t-j}$$

$$\hat{A}_t = y_t - \hat{S}_t = [1 - \omega_S(B)] y_t.$$

where $\hat{A}_t$, the "seasonally adjusted series," can be thought of as estimating the corresponding "true seasonally adjusted series"

$$A_t \equiv y_t - S_t = N_t + e_t.$$

If there is no sampling error in y_t, then $A_t = N_t$.

True seasonally adjusted series:

$$A_t \equiv y_t - S_t = N_t + e_t$$

Important Points:

1. If sampling error (e_t) is present in the series y_t, then $A_t \neq N_t$.

2. For standard software (X-11, SEATS, X-13), $\omega_N(B) = 1 - \omega_S(B)$, which implies that $\hat{N}_t = \hat{A}_t$.

3. Thus, when sampling error is present, it isn't clear whether these programs are estimating A_t or N_t.

Two Choices of Seasonal Adjustment Variances

- The error in using $\hat{N}_t$ to estimate N_t is

$$\begin{aligned}\hat{\varepsilon}_t^N &= N_t - \omega_N(B)[S_t + N_t + e_t] \\ &= [1 - \omega_N(B)]N_t - \omega_N(B)S_t - \omega_N(B)e_t.\end{aligned}$$

- The error in using $\hat{S}_t$ to estimate S_t is

$$\hat{\varepsilon}_t^S = -\omega_S(B)N_t + [1 - \omega_S(B)]S_t - \omega_S(B)e_t.$$

- The error in the seasonally adjusted series is

$$\hat{\varepsilon}_t^A = (y_t - S_t) - (y_t - \hat{S}_t) = -\hat{\varepsilon}_t^S.$$

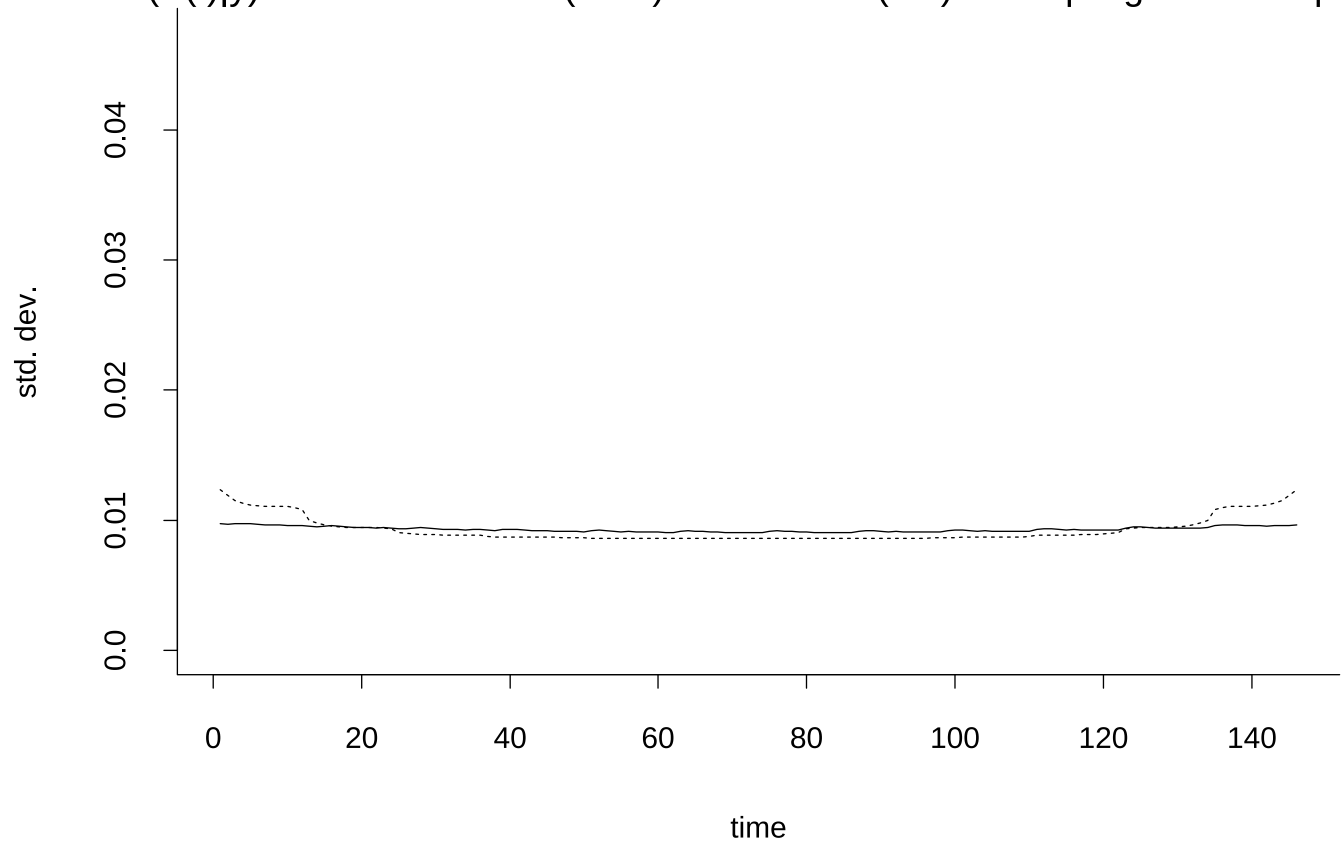
- The two choices of seasonal adjustment variances are

$$\mathrm{Var}(\hat{\varepsilon}_t^N) \qquad \text{or} \qquad \mathrm{Var}(\hat{\varepsilon}_t^A) = \mathrm{Var}(\hat{\varepsilon}_t^S).$$

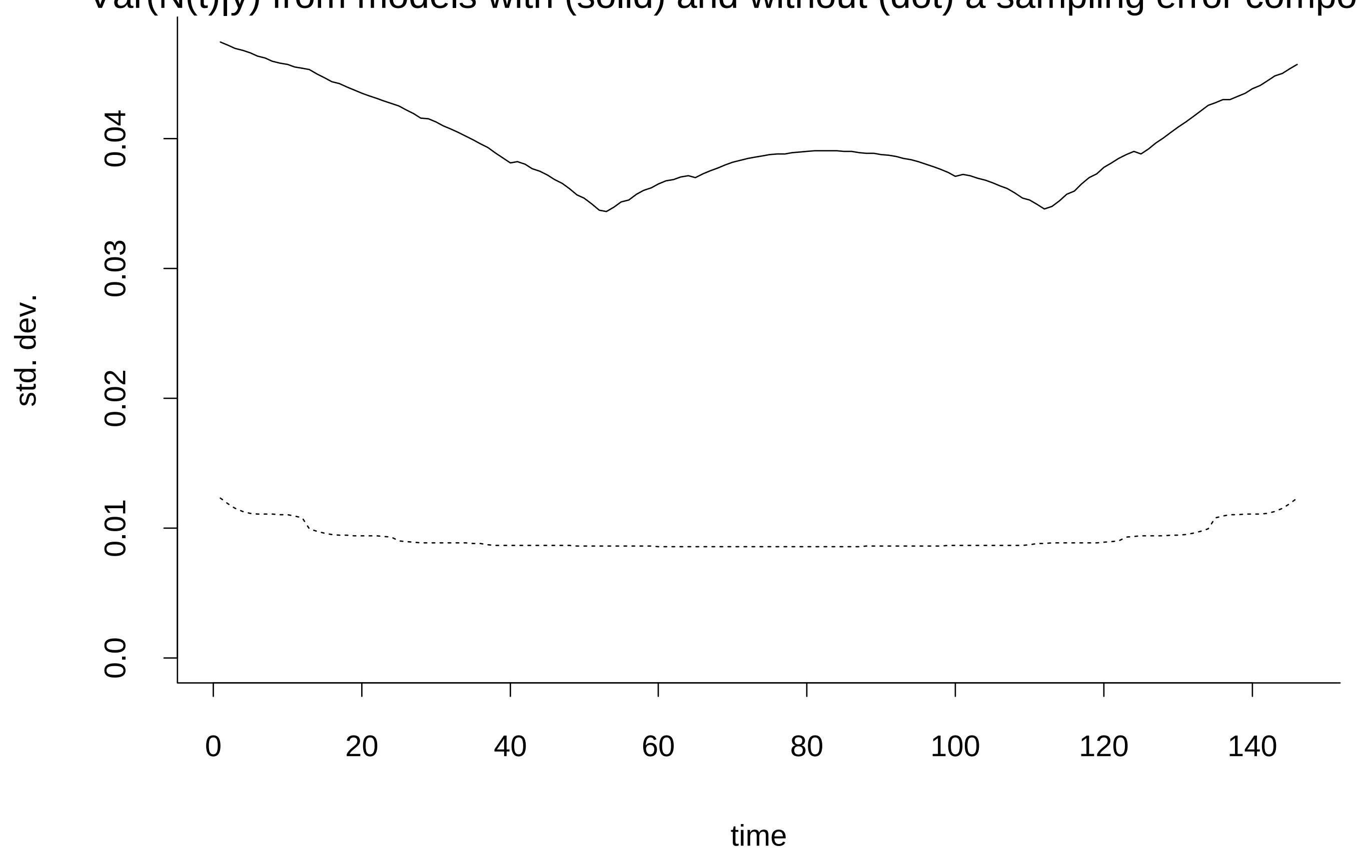
If sampling error e_t is present, these two choices are different.

U.S. Retail Sales of Drinking Places

Var(A(t)|y) from models with (solid) and without (dot) a sampling error component



Var(N(t)|y) from models with (solid) and without (dot) a sampling error component



Contributions to Seasonal Adjustment Error from the Time Series Components

Errors in estimates of components (again):

$$\begin{aligned}\hat{\varepsilon}_t^N &= [1-\omega_N(B)]N_t - \omega_N(B)S_t - \omega_N(B)e_t \\ \hat{\varepsilon}_t^S &= -\omega_S(B)N_t + [1-\omega_S(B)]S_t - \omega_S(B)e_t \\ \hat{\varepsilon}_t^A &= -\hat{\varepsilon}_t^S.\end{aligned}$$

These expressions hold whether $\omega_N(B)$ and $\omega_S(B)$ are symmetric or asymmetric filters, model-based or X-11 filters.

From orthogonality of the components, the variances of these errors are

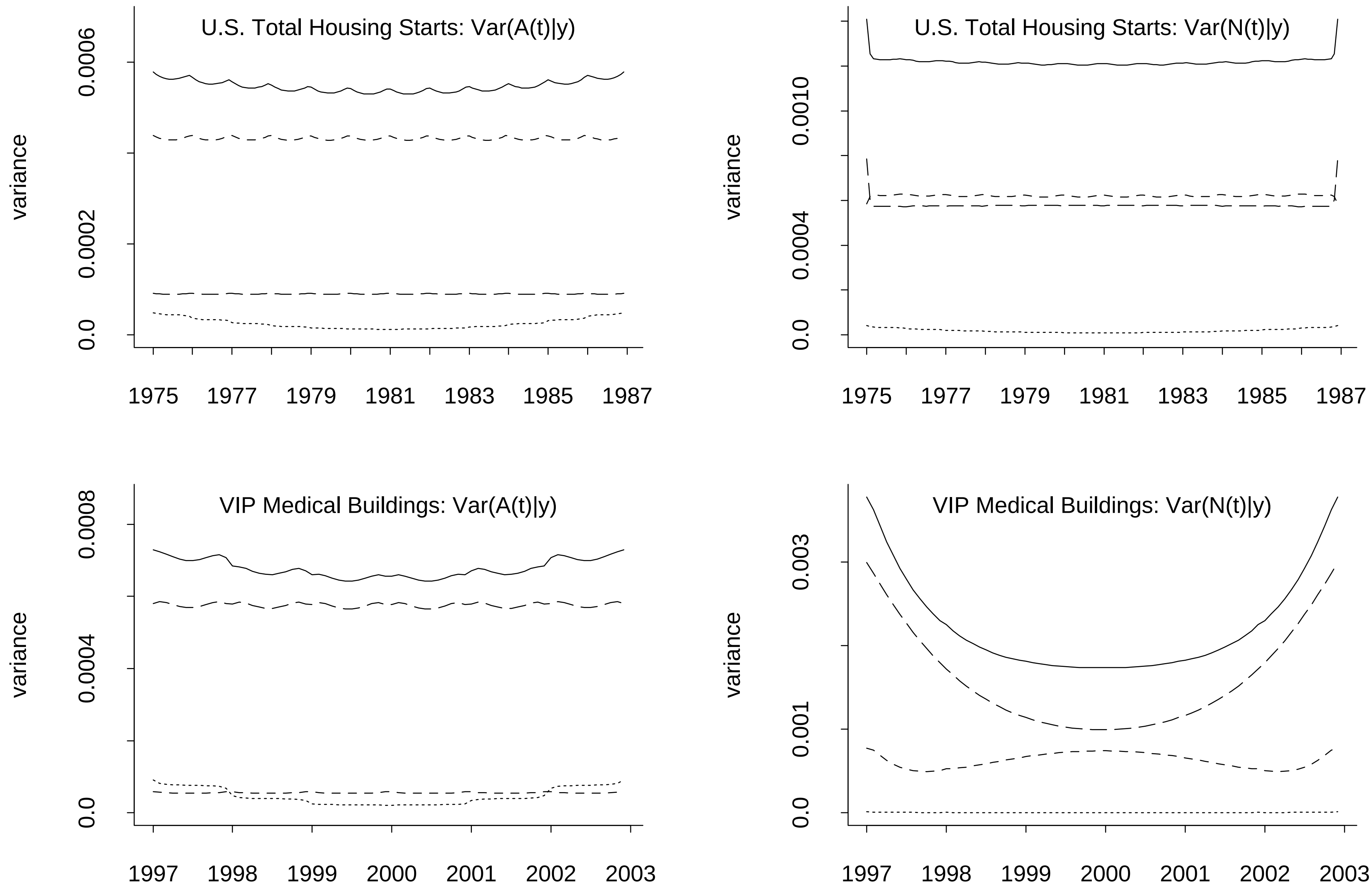
$$\text{Var}(\hat{\varepsilon}_t^N) = \text{Var}\{[1-\omega_N(B)]N_t\} + \text{Var}[\omega_N(B)S_t] + \text{Var}[\omega_N(B)e_t]$$

$$\begin{aligned}\text{Var}(\hat{\varepsilon}_t^S) &= \text{Var}[\omega_S(B)N_t] + \text{Var}\{[1-\omega_S(B)]S_t\} + \text{Var}[\omega_S(B)e_t] \\ &= \text{Var}(\hat{\varepsilon}_t^A).\end{aligned}$$

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Fig. 2 Seasonal adjustment variances and component contributions for Example 2

solid = Var(A(t)|y) or Var(N(t)|y), long dash ~ e(t), dot ~ S(t), dash ~ N(t)



Contributions to Seasonal Adjustment Error from Forecast Extension Error

Let

$$\hat{y}_t = \begin{cases} y_t & t = 1, \ldots, n \\ E(y_t | y_1, \ldots, y_n) & \text{otherwise.} \end{cases}$$

Denote the estimators of N_t, S_t, and A_t based on the finite data as

$$\tilde{N}_t \equiv \omega_N(B)\hat{y}_t \qquad \tilde{S}_t \equiv \omega_S(B)\hat{y}_t \qquad \tilde{A}_t \equiv y_t - \tilde{S}_t.$$

where here we assume that $\omega_N(B)$ and $\omega_S(B)$ are *symmetric*. The error in the estimator $\tilde{N}_t$ is

$$\begin{aligned} \tilde{\varepsilon}_t^N &= N_t - \tilde{N}_t \\ &= (N_t - \hat{N}_t) + (\hat{N}_t - \tilde{N}_t) \\ &= \hat{\varepsilon}_t^N + (\hat{N}_t - \tilde{N}_t). \end{aligned}$$

where $\hat{\varepsilon}_t^N$ is the error in the symmetric estimator $\hat{N}_t$.

The second term in $\tilde{\varepsilon}_t^N$ is the revision from $\tilde{N}_t$ to $\hat{N}_t$. It can be shown to depend linearly on the forecast and backcast errors (Pierce 1980).

The variance of the error, $\tilde{\varepsilon}_t^N = \hat{\varepsilon}_t^N + (\hat{N}_t - \tilde{N}_t)$, is

$$\text{Var}(\tilde{\varepsilon}_t^N) = \text{Var}(\hat{\varepsilon}_t^N) + \text{Var}(\hat{N}_t - \tilde{N}_t) + 2\,\text{Cov}(\hat{\varepsilon}_t^N, \hat{N}_t - \tilde{N}_t).$$

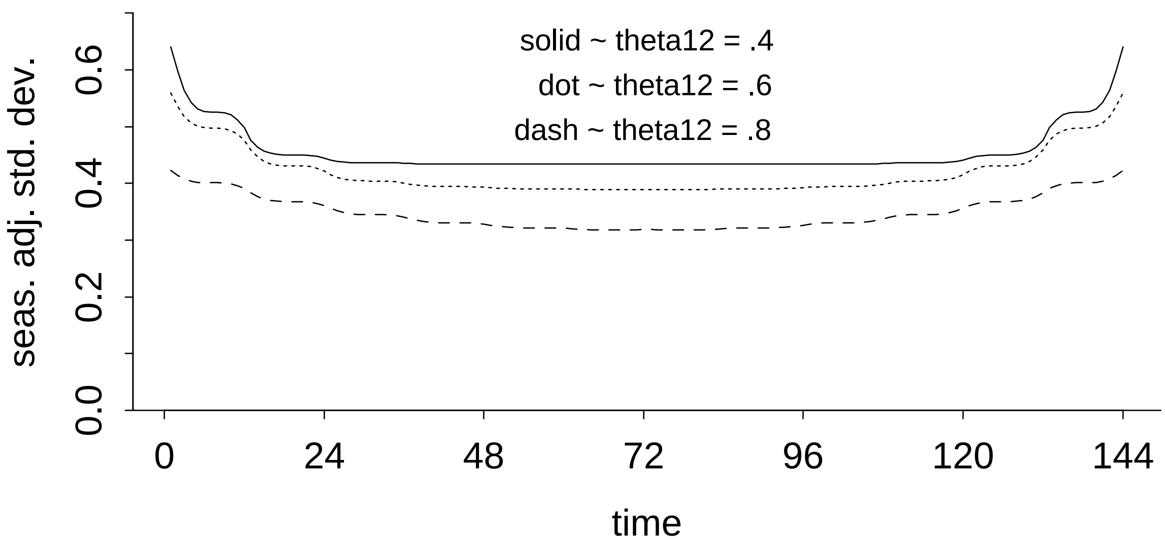
If $\hat{N}_t$ is the optimal estimator of N_t (as assumed in model-based adjustment), then the covariance term is zero.

For X-11 adjustment, $\hat{N}_t$ is *not* the optimal estimator of N_t, and the covariance term is *not* zero.
(See Bell and Kramer, 1999, *Survey Methodology*.)

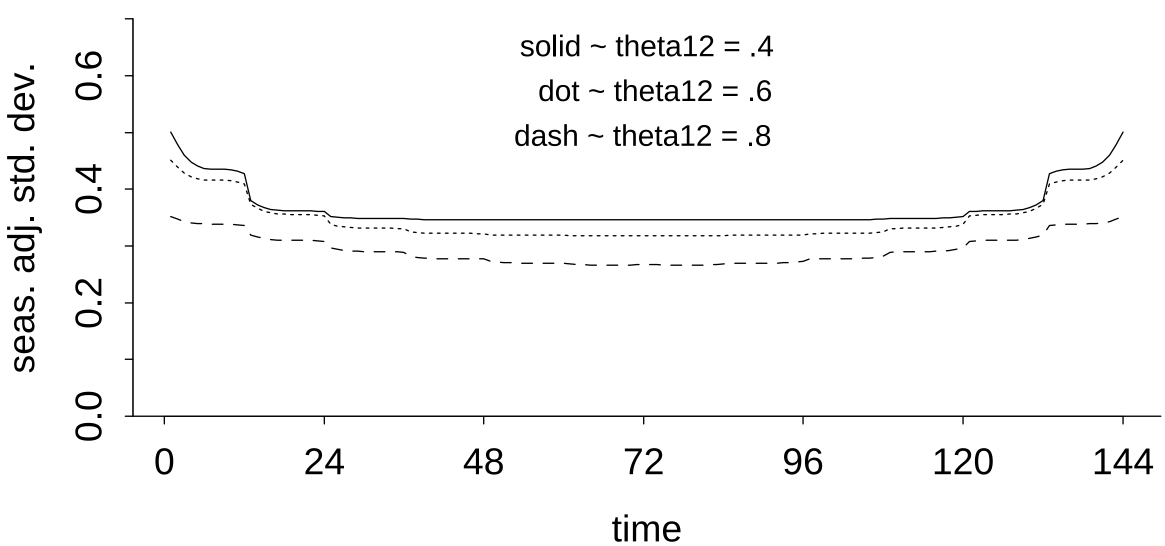
Airline model - canonical seasonal adjustment std. deviations

(innovation standard deviation = 1)

theta1 = 0.0

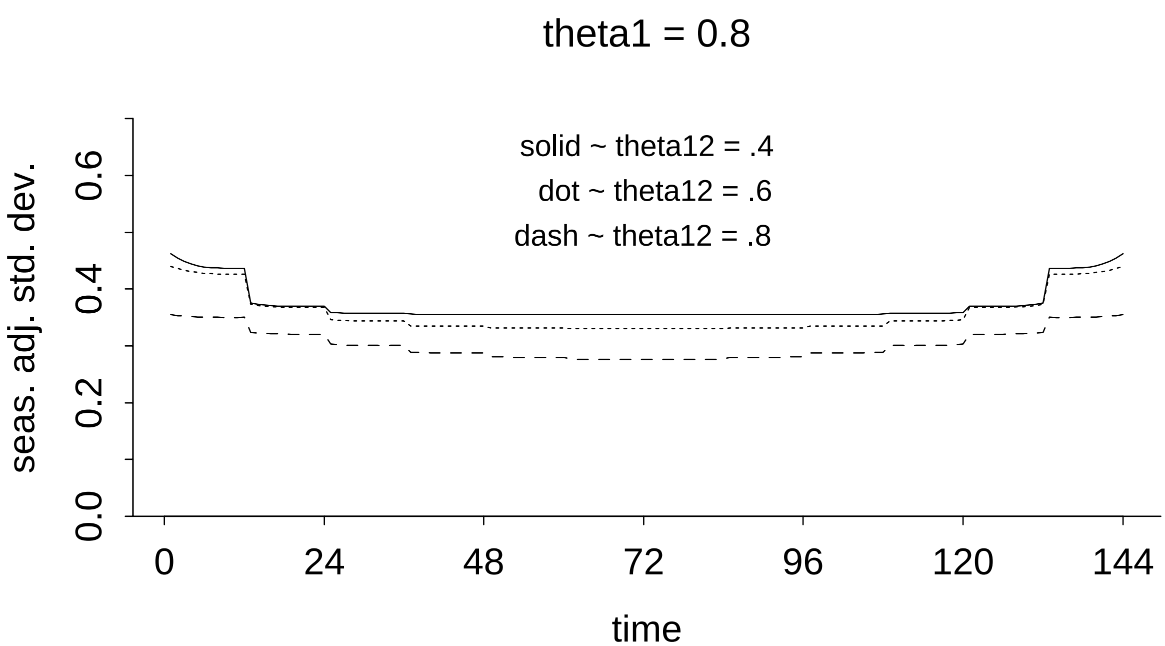


theta1 = 0.4



Airline model - canonical seasonal adjustment std. deviations

(innovation standard deviation = 1)



Contributions to Seasonal Adjustment Error from Error in Estimating Model Parameters

For model-based adjustment, or for X-11 adjustment with model-based estimation of regression effects (e.g., trading-day), let

$$\hat{N}_t = \text{estimate of } N_t \text{ when model parameters are known}$$
$$\bar{N}_t = \text{estimate of } N_t \text{ when model parameters are estimated.}$$

The error in the estimator $\bar{N}_t$ is

$$\bar{\varepsilon}_t^N = (N_t - \hat{N}_t) + (\hat{N}_t - \bar{N}_t)$$

and its variance is

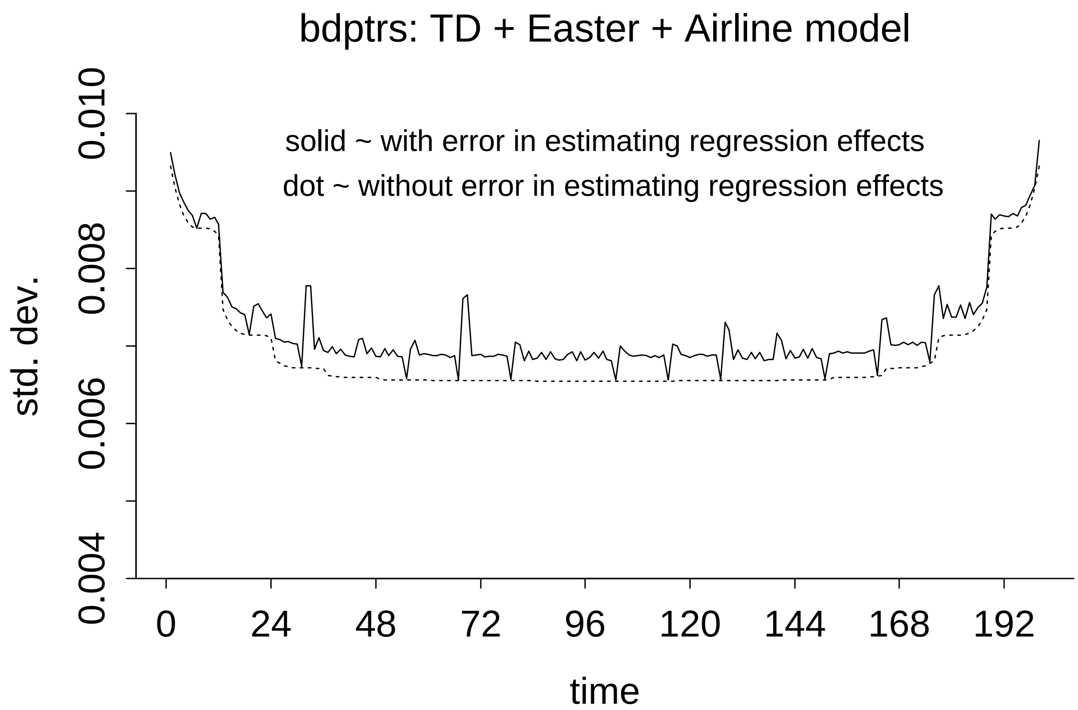
$$\text{Var}(\bar{\varepsilon}_t^N) = \text{Var}(\hat{\varepsilon}_t^N) + \text{Var}(\hat{N}_t - \bar{N}_t) + 2 \times \text{Cov}(\hat{\varepsilon}_t^N, \hat{N}_t - \bar{N}_t).$$

If $\hat{N}_t$ is the optimal estimator of N_t then the covariance term is zero.

For X-11 adjustment, $\hat{N}_t$ is *not* the optimal estimator of N_t, and the covariance term is *not* zero.

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Fig. 3 Seasonal adjustment standard deviations with and without accounting for error in estimating regression effects



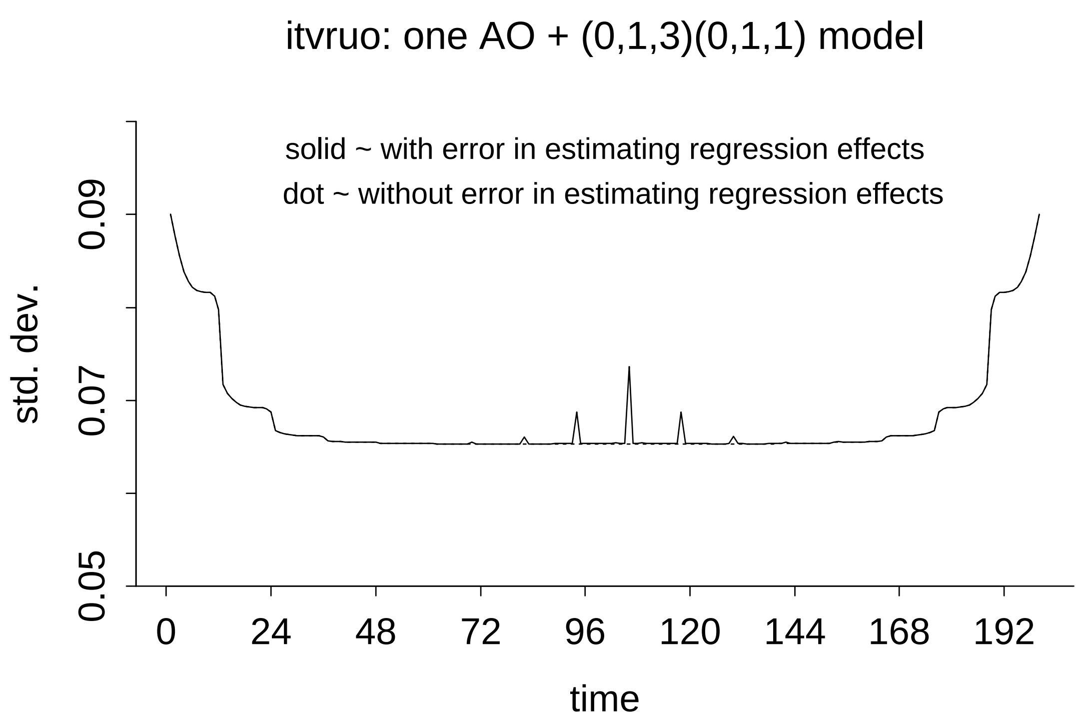
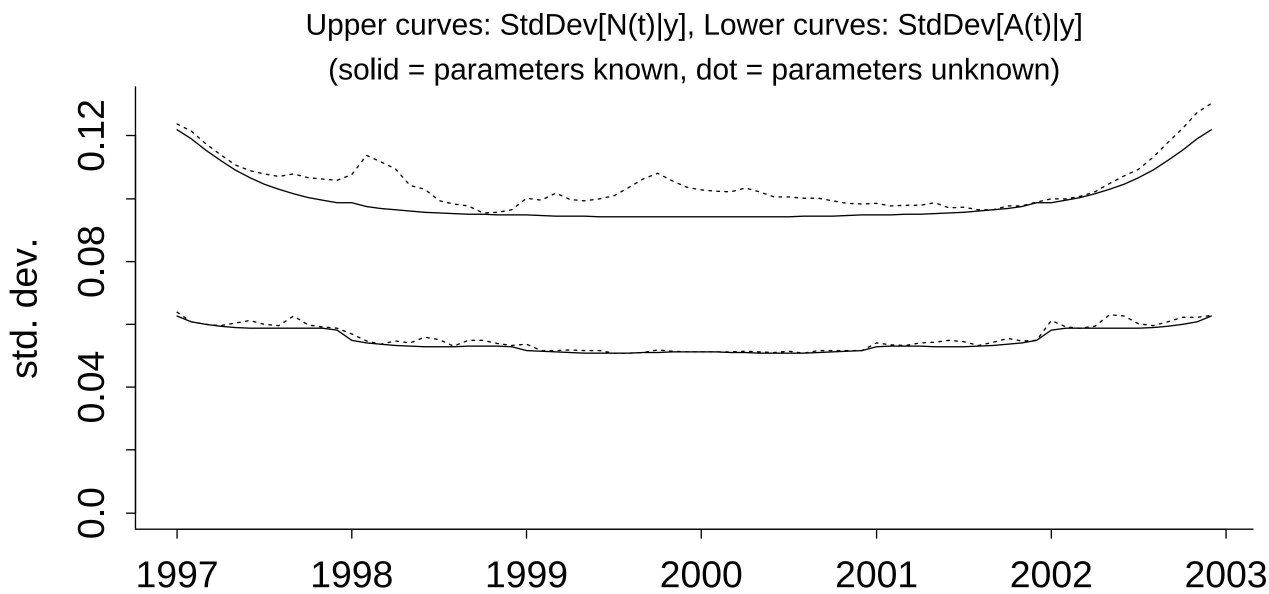


Fig. 4 VIP Other Educ.: Seasonal adjustment std. devs.



Variances for X-11 Seasonal Adjustments

Proposed approaches:

- WM: Wolter and Monsour (1981)
- DP: Pfeffermann (1994)
- BK: Bell and Kramer (1999)
- SPS: Scott, Pfeffermann, and Sverchkov (2012)
- CTB: Chu, Tiao, and Bell (2012); Bell, Chu, and Tiao (2012)

The above approaches differ in which contributions to error are recognized, and to what extent forecast extension error is accounted for.

For X-11 seasonal adjustment $\omega_N(B) = 1 - \omega_S(B)$, and the errors in the estimators of N_t and $A_t = y_t - S_t$ are:

$$\begin{aligned} \hat{\varepsilon}_t^N &= -\omega_N(B)S_t + \omega_S(B)T_t + \omega_S(B)I_t - \omega_N(B)e_t \\ \hat{\varepsilon}_t^A = -\hat{\varepsilon}_t^S &= -\omega_N(B)S_t + \omega_S(B)T_t + \omega_S(B)I_t + \omega_S(B)e_t. \end{aligned}$$

Approaches to variances for X-11 seasonal adjustments

Approach	seas. adj. error	**accounting for forecast extension error**
WM	$-\omega_N(B)e_t$	partial
DP, SPS	$\omega_S(B)I_t - \omega_N(B)e_t$	partial
BK	$-\omega_N(B)e_t$	full
CTB*	$\hat{\varepsilon}_t^N$	full

* CTB, in their two papers, consider only the case with no sampling error component, though the approach can be extended to accomodate one.

For examples comparing the DP and BK approaches see Scott, Pfeffermann, and Sverchkov (2012).

Conclusions

1. When sampling error is not present ($y_t = S_t + N_t$)
   - $A_t \equiv y_t - S_t = N_t$
   - $\hat{\varepsilon}_t^N = N_t - \hat{N}_t = A_t - \hat{A}_t = \hat{\varepsilon}_t^A$
   - Var($\hat{\varepsilon}_t^N$) =Var($\hat{\varepsilon}_t^A$)
2. When sampling error is present ($y_t = S_t + N_t + e_t$)
   - $A_t \equiv y_t - S_t = N_t + e_t$
   - $\hat{\varepsilon}_t^N \neq \hat{\varepsilon}_t^A$
   - Var($\hat{\varepsilon}_t^N$) $\neq$Var($\hat{\varepsilon}_t^A$), and these can be quite different.

3. For the case when e_t is present, we need to decide whether "seasonal adjustment variance" means Var($\hat{\varepsilon}_t^A$) or Var($\hat{\varepsilon}_t^N$). Proposed approaches to X-11 variances are most consistent with defining $\hat{\varepsilon}_t^N$ to be the seasonal adjustment error.

4. Contributions from the components (S_t, N_t, and e_t) to seasonal adjustment variances can be calculated and compared. Often when e_t is present, its contribution is quite important.

5. Proposed approaches to variances for X-11 seasonal adjustment omit the contributions of some components, effectively assuming they are negligible. X-11 variances can be developed to account for the contributions from all the components, but this requires a model and substantial calculations.

6. To produce seasonal adjustment variances, model-based approaches should recognize sampling error (when present).

7. Seasonal adjustment variances raise some "presentation issues"

- different variances for many time points
- erratic nature of contributions from parameter estimation error
- most proposed approaches to X-11 variances can produce dips in the variances near the ends of the series, which are unreasonable

References

[1] Bell, William R., Chu, Yea-Jane, and Tiao, George C. (2012), "Comparing Mean Squared Errors of X-12-ARIMA and Canonical ARIMA Model-Based Seasonal Adjustments," chapter 7 in *Economic Time Series: Modeling and Seasonality*, eds. William R. Bell, Scott H. Holan, and Tucker S. McElroy, Boca Raton, FL: CRC Press-Chapman and Hall, pp. 161–184.

[2] Bell, William R. and Kramer, Matthew (1999) "Towards Variances for X-11 Seasonal Adjustments," *Survey Methodology*, **25**, 13–29, available online at http://www.statcan.gc.ca/pub/12-001-x/1999001/article/4709-eng.pdf.

[3] Chu, Yea-Jane, Tiao, George C., and Bell, William R. (2012), "A Mean Squared Error Criterion for Comparing X-12-ARIMA and Model-Based Seasonal Adjustment Filters," *Taiwan Economic Forecast and Policy*, **43**, 1–32.

[4] Pfeffermann, Danny (1994) "A General Method for Estimating the Variances of X-11 Seasonally Adjusted Estimators," *Journal of Time Series Analysis*, 15, 85–116.

[5] Pierce, David A. (1980), "Data Revisions With Moving Average Seasonal Adjustment Procedures," *Journal of Econometrics*, **14**, 95–114.

[6] Scott, Stuart, Pfeffermann, Danny, and Sverchkov, Michail (2012), "Estimating Variance in X-11 Seasonal Adjustment," chapter 8 in *Economic Time Series: Modeling and Seasonality*, eds. William R. Bell, Scott H. Holan, and Tucker S. McElroy, Boca Raton, FL: CRC Press-Chapman and Hall, pp. 185–210.

[7] Wolter, Kirk M. and Monsour, Nash J. (1981), "On the Problem of Variance Estimation for a Deseasonalized Series," in *Current Topics in Survey Sampling*, ed. D. Krewski, R. Platek, and J.N.K. Rao, New York: Academic Press, 367–403.