Statistics Canada / Statistique Canada

Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

Canada

[redacted]

Working Paper No TSRAD-88-003E
Time Series Research and Analysis Division
Methodology Branch

DETERMINISTIC AND STOCHASTIC MODELS FOR THE ESTIMATION OF TRADING-DAY VARIATIONS

by

Estela Bee Dagum and Benoit Quenneville

Time Series Research & Analysis Division
13-K, R.H. Coats Bldg.
Statistics Canada
Ottawa (Ontario)
Canada
K1A 0T6

January, 1988.

Invited paper presented at the IV Annual Research Conference of the U.S. Bureau of the Census, March 1988.

ABSTRACT

DETERMINISTIC AND STOCHASTIC MODELS FOR THE ESTIMATION OF TRADING-DAY VARIATIONS

by

Estela Bee Dagum and Benoit Quenneville
Statistics Canada

A large class of flows and stocks series related to production, shipments, sales and inventories are affected by trading-day or calendar variations. Trading-day variations represent the "within-month variations" due to the number of times a particular day or days of the week occur in a calendar month. These variations are systematic and may strongly influence month-to-month comparisons. Whenever present they must be removed together with seasonality to obtain a clear signal of the short-term trend-cycle component, which is used for decision making by socio-economic players.

The X-11-ARIMA (Dagum, 1980) and the Census Method II-X-11 variant (Shiskin, Young and Musgrave, 1967) use the deterministic trading-day model developed by Young (1965). This model assumes that the daily weights and the weekly pattern remain constant throughout the chosen span of the series. For some socio-economic time series this assumption might be questionable.

This paper presents two stochastic models for trading-day variations that allow a moving behavior of the daily coefficients. The estimation method is discussed and the deterministic and stochastic models are applied on both simulated and real series.

RESUME

MODELES DETERMINISTE ET STOCHASTIQUES POUR L'ESTIMATION DE LA COMPOSANTE DE ROTATION DES JOURS

par

Estela Bee Dagum et Benoit Quenneville
Statistique Canada

Un grand nombre de séries de flux et de stock reliés à la production, ventes et inventaires sont affectées par la rotation des jours ou l'effet du calendrier. La composante de rotation des jours représente les variations intra-mensuelles causées par la distribution des jours dans le mois (ex: 4 lundis,..., 4 vendredis, 5 samedis et 5 dimanches dans un mois de 30 jours). Ces variations sont systématiques et peuvent fortement influeçées les comparaisons entre les mois. Ainsi, lorsque ces variations sont présentes dans une série, elles doivent être identifiées avec la saisonnalité pour obtenir un estimé plus préçis de la tendance-cycle.

Les méthodes X11-ARMMI (Dagum, 1980) et X11 (Shiskin, Young et Musgrave, 1967) utilisent un modèle déterministe pour l'estimation de la composante de rotation des jours developpé par Young (1965). Ce modèle assume que les poids des jours et leur patron hebdomadaire restent constants à travers une période pré-determinée de la série. Pour certaines séries socio-économiques cette hypothèse peut être discutable.

Cet article présente deux modèles stochastiques pour l'estimation de la composante de rotation des jours. La méthode d'estimation est discutée et les modèles déterministes et stochastiques sont appliqués sur des séries simulées et réelles.

I. Introduction.

A large class of flows and stocks series related to production, shipments, sales and inventories are affected by trading-day or calendar variations. Trading-day variations represent the "within-month variations" due to the number of times a particular day or days of the week occur in a calendar month. These variations are systematic and may strongly influence month-to-month comparisons. Whenever present, they must be removed together with seasonality to obtain a clear signal of the short-term trend (trend-cycle) of the series.

The X-11-ARIMA (Dagum, 1980) and the Census Method II-X-11 variant (Shiskin, Young and Musgrave, 1967) estimate trading-day variations using a simple regression model developed by Young (1965). This model assumes a deterministic behaviour in the sense that the daily weights and the weekly pattern remain constant throughout the chosen span of the series. For some socio-economic time series, however, this assumption may be too restrictive and a stochastic model for trading-day variations be more adequate.

The main purpose of this study is to introduce two stochastic models of trading-day variations for gradually moving daily coefficients. Section 2 gives a definition of trading-day variations. The two stochastic models are discussed in section 3 together with the deterministic model. Section 4 deals with the estimation procedure of the stochastic models. In Section 5, the deterministic and the two stochastic trading-day variation models are tested on simulated data. In section 6, the three models are applied to two real series affected by trading-day variations. Finally, section 7 gives the conclusions of this study.

2. Definition of Trading-Day Variations.

Let ξ_{it} $i=1,2,\ldots,7$ represent the effects of daily activity on Monday, Tuesday,..., and Sunday in month t. The overall effect attributed to the number of times each day of the week occurs in month t defines what is known as trading-day variations or effects. That is,

$$E_t = \sum_{i=1}^{7} \xi_{it} X_{it} \tag{2.1}$$

where X_{it} $i=1,\ldots,7$ denotes respectively, the number of Mondays, Tuesdays,..., and Sundays in month t.

Let $\bar{\xi}_t = 1/7 \sum_{i=1}^{7} \xi_{it}$ be the average of the daily effects and X_t be the number of days in month t. Then we can reparametrize (2.1) as follows:

$$E_t = \sum_{i=1}^{6} (\xi_{it} - \bar{\xi}_t)(X_{it} - X_{7t}) + \bar{\xi}_t (X_t - 365.25/12) + \bar{\xi}_t (365.25/12) \tag{2.2}$$

Equation (2.2) decomposes the overall effect E_t in three parts: (i) The trading-day effect, (ii) the length-of-month effect and (iii) the month effect.

The trading-day effect in month t is given by the first term of equation (2.2), that is:

$$D_t = \sum_{i=1}^{6} (\xi_{it} - \bar{\xi}_t)(X_{it} - X_{7t}) \tag{2.3}$$

If all the ξ_{it}'s are equal, there is no trading-day effect. Similarly, for the month of February, except in leap year, $X_{it} = X_7$ $i=1,\ldots,6$, and there is no trading-day effect. For notational convenience, let $\delta_{it} = \xi_{it} - \bar{\xi}_t$ $i=1,\ldots,6$ and $T_{it} = X_{it} - X_{7t}$. The δ_{it}'s represent the difference between the Monday, Tuesday,..., and Saturday effects ξ_{it} and the average of the daily effects $\bar{\xi}_t$, for month t. The difference between the Sunday effect and the average of the daily effects is $\xi_{i7} - \bar{\xi}_t = -\sum_{i=1}^{6} \delta_{it}$.

The second term $\bar{\xi}_t (X_t - 365.25/12)$ represents the length of month effect and is usually attributed to seasonality.

Finally, the third term $\bar{\xi}_t (365.25/12)$ represents the average effect in month t if all the months would be of equal length and is usually attributed to the trend-cycle component.

Under the assumption that the trend-cycle and seasonal variations have been adequately estimated and removed from the data, the trading-day effects definition (2.1) reduces to definition (2.3), i.e. $E_t = D_t$.

3. Deterministic and Stochastic Models for Trading-day Coefficients or Daily Weights.

Given a time series, say y_t, where already the trend-cycle and seasonal fluctuations have been removed, we assume:

$$y_t = D_t + e_t, \quad t=1,\ldots,T \tag{3.1}$$

where $e_t \sim NID(0,\sigma^2)$ and T is the number of observations.

In this section, we introduce three models for the estimation of D_t, namely, a deterministic model, a random walk model and a random walk model with a random drift.

The deterministic trading-day variations model developed by Young (1965) assumes that $\delta_{it} = \delta_t$ for all t. In this case, equation (2.3) reduces to:

$$D_t = \sum_{i=1}^{6} \delta_i T_{it} \tag{3.2}$$

where the δ_i's are considered as fixed parameters and estimated using ordinary least squares (OLS).

The random walk model proposed by Monsell (1983), can be written as follows:

$$D_t = \sum_{i=1}^{6} \delta^1_{it} T_{it} \tag{3.3.a}$$

with

$$\underline{\delta}^1_t = \underline{\delta}^1_{t-1} + \chi_t \tag{3.3.b}$$

where

$$\underline{\delta}^1_t = (\delta^1_{1t},\ldots,\delta^1_{6t})' \tag{3.4.a}$$

$$\chi^1_t = (\chi_{1t},\ldots,\chi_{6t})' \tag{3.4.b}$$

$$\chi^1_t \sim NID(0,\sigma^2_\chi I_6). \tag{3.4.c}$$

Here, I_6 is the identity matrix of order 6.

Finally, the second stochastic model discussed in this paper assumes that the vector of daily coefficients follows a random walk model with a random drift. That is:

$$D_t = \sum_{i=1}^{6} \delta^2_{it} T_{it} \qquad (3.5.a)$$

with

$$\underline{\delta}^2_t = \underline{\delta}^2_{t-1} + \underline{\rho}_{t-1} + \underline{\chi}_t \qquad (3.5.b)$$

$$\underline{\rho}_t = \underline{\rho}_{t-1} + \underline{\psi}_t \qquad (3.5.c)$$

where

$$\underline{\delta}^2_t = (\delta^2_{1t}, \ldots, \delta^2_{6t}) \qquad (3.6.a)$$

$$\underline{\rho}_t = (\rho_{1t}, \ldots, \rho_{6t}) \qquad (3.6.b)$$

$$\underline{\chi}_t = (\chi_{1t}, \ldots, \chi_{6t}) \qquad (3.6.c)$$

$$\underline{\psi}_t = (\psi_{1t}, \ldots, \psi_{6t}) \qquad (3.6.d)$$

$$\underline{\chi}_t \sim NID(0, \sigma^2_\chi I_6) \qquad (3.6.e)$$

$$\underline{\psi}_t \sim NID(0, \sigma^2_\psi I_6) \qquad (3.6.f)$$

Here $\underline{\chi}_t$ and $\underline{\psi}_t$ are mutually independent. Equations (3.5.b) and (3.5.c) give a local approximation to a linear trend in the daily coefficients. The level and slope of the trend are assumed to be generated by stochastic processes.

The two stochastic models (3.3.b) and (3.5.b-3.5.c) are written in state-space forms and the estimates of $\underline{\delta}^i_t$ i=1,2, together with their mean squared error matrices are estimated with the Kalman filter. Smoothed estimated are obtained using the fixed interval smoother. Finally, maximum likelihood estimators are used to estimate the remaining hyper-parameters σ^2, σ^2_χ/σ^2 and σ^2_ψ/σ^2.

4. Estimation of the Stochastic Models.

The estimation of the stochastic models is made using the Kalman filter and the fixed interval smoother. A brief description of these two now follows.

4.1 The Kalman Filter and Fixed Interval Smoother.

The state space model consists of a measurement equation, namely,

$$y_t = \underline{z}'_t \underline{\alpha}_t + \epsilon_t, \quad t=1,\ldots,T \qquad (4.1)$$

and a transition equation, namely,

$$\underline{\alpha}_t = G\underline{\alpha}_{t-1} + \underline{\eta}_t, \quad t=1,\ldots,T \qquad (4.2)$$

where $\underline{\alpha}_t$ is an mxl state vector, $\underline{z}_t$ is a mxl fixed-vector, G is a fixed mxm matrix and the errors ϵ_t and $\underline{\eta}_t$ are independent. It is further assumed that $\epsilon_t \sim NID(0,\sigma^2)$ and $\underline{\eta}_t \sim NID(0,\sigma^2 Q)$ where Q is a fixed mxm matrix and σ^2 is a

scalar. Although Q may depend on unknown parameters it is regarded as being fixed and known for the purpose of the Kalman filter.

Let $\underline{a}_{t-1}$ be the minimum mean squared estimator (MMSE) of $\underline{\alpha}_{t-1}$ based on all the information up to and including t-1, and let $\sigma^2 P_{t-1}$ be the MSE matrix of $\underline{a}_{t-1}$, i.e., the covariance matrix of $\underline{a}_{t-1}-\underline{\alpha}_{t-1}$. The MMSE of $\underline{\alpha}_t$, given $\underline{a}_{t-1}$ and P_{t-1}, is then given by:

$$\underline{a}_{t|t-1} = G\underline{a}_{t-1} \quad (4.3)$$

with MSE matrix:

$$P_{t|t-1} = GP_{t-1}G' + Q. \quad (4.4)$$

Once y_t becomes available, this estimator of $\underline{\alpha}_t$ can be updated as follows:

$$\underline{a}_t = \underline{a}_{t|t-1} + P_{t|t-1}\underline{z}_t v_t / f_t \quad (4.5)$$

with MSE matrix:

$$P_t = P_{t|t-1} - P_{t|t-1}\underline{z}_t\underline{z}_t' P_{t|t-1}/f_t \quad (4.6)$$

where

$$v_t = y_t - \underline{z}_t'\underline{a}_{t|t-1} \quad (4.7)$$

$$f_t = \underline{z}_t' P_{t|t-1}\underline{z}_t + 1. \quad (4.8)$$

Starting values $\underline{a}_0$ and P_0 are needed to implement the Kalman filter given by (4.3) to (4.8).

The Kalman filter yields the MMSE of $\underline{\alpha}_t$, given the information available up to time t. However, once all the observations are available, a better estimator can be obtained. One of the techniques for computing such estimators is the fixed interval smoother. The fixed interval smoother is a set of recursions which start with the Kalman filter estimates $\underline{a}_T$ and P_T, and work backwards. If $\underline{a}_{t|T}$ and $\sigma^2 P_{t|T}$, denote the smoothed estimate and its covariance matrix, the smoothing equation is given by:

$$\underline{a}_{t|T} = \underline{a}_t + P^*_t(\underline{a}_{t+1|T} - G\underline{a}_t) \quad (4.9)$$

with

$$P_{t|T} = P_t + P^*_t(P_{t+1|T} - P_{t+1|t})(P^*_t)' \quad (4.10)$$

where

$$P^*_t = P_t G'(P_{t+1|t})^{-1}. \quad (4.11)$$

4.2 <u>State Space Representation of the Two Stochastic Models.</u>

A convenient state-space representation of the random walk stochastic model for the trading-day coefficients (3.3) and (3.4) along with equation (3.1) is obtained through the following equivalences with the transition equation (4.2) and the measurement equation (4.1):

$$\underline{\alpha}_t = \underline{\delta}^1{}_t,\ \underline{z}'_t = (T_{1t},\ldots,T_{6t}),\ \epsilon_t = e_t, \qquad (4.12)$$
$$G = I_6,\ \underline{\eta}_t = \underline{\chi}_t \text{ and } Q = \sigma^2{}_\chi/\sigma^2 I_6.$$

For the random walk model with a random drift, described by equations (3.5) and (3.6), the equivalences are given by:

$$\underline{\alpha}_t' = [\underline{\delta}^2{}_t',\ \underline{\rho}_t'],\ \underline{z}'_t = (T_{1t},\ldots,T_{6t},\underline{0}_6'),\ \epsilon_t = e_t,$$
$$G = \begin{bmatrix} I_6 & I_6 \\ 0 & I_6 \end{bmatrix},\ \underline{\eta}_t = \begin{bmatrix} \underline{\chi}_t \\ \underline{\psi}_t \end{bmatrix} \text{ and } Q = \begin{bmatrix} \sigma^2{}_\chi/\sigma^2 I_6 & 0 \\ 0 & \sigma^2{}_\psi/\sigma^2 I_6 \end{bmatrix} \qquad (4.13)$$

It is clear from (4.12) and (4.13) that Q depends on the unknown parameters $\sigma^2\chi/\sigma^2$ in the first model, and $\sigma^2\chi/\sigma^2$ and $\sigma^2\psi/\sigma^2$ in the second model. These parameters along with σ^2 are called hyper-parameters since they represent the parameters of the a-priori distribution of the state vector. Their estimation as well as the estimation of the initial conditions $\underline{a}_0$ and P_0 are discussed in the next section.

4.3 Estimation of the Initial Conditions and the Hyper-Parameters.

4.3.1 *Estimation of the Initial Conditions.*

One way of deriving the initial estimate $\underline{a}_0$ of $\underline{\alpha}_0$ with its covariance matrix $\sigma^2 P_0$ is by assuming that the state vectors $\underline{\alpha}_t$ are deterministic instead of stochastic over the first K observations. This leads to the regression model (3.2) for the random walk model and to $\underline{\delta}^2{}_t = \underline{\delta}^2{}_0 + t\underline{\rho}_0$ for the second stochastic model. The estimated covariance matrix of $\underline{a}_0$ from the regression provides an estimate of $\sigma^2 P_0$ from which P_0 is easily obtained. We will refer to this approach as Method 1.

There are two other ways of estimating $\underline{\alpha}_0$ and P_0 for the random walk model. These follow:

Method 2: First we define a new series $w_t = y_{T-t+1}$. The w series is obtained by reversing the order of the y's series; that is, $w_1 = y_T$, $w_2 = y_{T-1}$ and so on until $w_T = y_1$. The random walk model for the trading day coefficients is fitted to the w series. The Kalman filter is applied on this transformed series to predict $\underline{\alpha}_0$ as $\underline{\alpha}^\#{}_{T+1|T}$ and P_0 as $P^\#{}_{T+1|T}$ where $\underline{\alpha}^\#{}_t$ and $P^\#{}_t$ are the estimates of the state vectors and covariance matrices of the w series. In applying the Kalman filter to this transformed series, the initial estimate of the state vector $\underline{\alpha}^\#{}_0$ is taken to be equal to $\underline{0}_6$, $P^\#{}_0$ equal to kI_6 where k is a large constant (21 in the simulation discussed in section 5) and the hyper-parameters of the w series are computed using the method described in section 4.3.2.

Method 3: Same as method 2, but with $\underline{\alpha}^\#{}_0$ and P_0 estimated as in method 1

using the w series.

4.3.2 *Estimation of the Hyper-Parameters.*

Maximum Likelihood Estimators (MLE) are considered for the estimation of the hyper-parameters. Using the prediction error decomposition (Harvey 1981), the likelihood function, L, can be written in the form:

$$-2\log L = T\log 2\pi + T\log\sigma^2 + \sum_{t=1}^{T} \log f_t + \sigma^{-2}\sum_{t=1}^{T} v^2{}_t/f_t \qquad (4.14)$$

where T is the number of observations and v_t and f_t are defined by (4.7) and (4.8).

Differentiation of (4.14) with respect to σ leads to $\hat{\sigma}^2$, the MLE of σ^2, given by:

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} v^2{}_t/f_t \qquad (4.15)$$

The scalar parameter, σ^2, may be concentrated out of the log-likelihood function leaving the concentrated log-likelihood function:

$$-2\log L_c = T\log 2\pi + T + T\log\hat{\sigma}^2 + \sum_{t=1}^{T} \log f_t \qquad (4.16)$$

Numerical optimization has to be carried out with respect to the remaining parameters ($\sigma^2{}_\chi/\sigma^2$ for the first model and $\sigma^2{}_\chi/\sigma^2$, $\sigma^2{}_\psi/\sigma^2$ for the second model) to minimize the right hand side of the equation (4.16). This can be done by using the Fibonacci line search method for the random walk model and the Davidon-Fletcher-Powell algorithm (Bazaraa and Shetly, 1979) for the random walk model with random drift. In both cases the parameters are bounded between 0 and 1. For the random walk model, this assumes that the noise in the signal $\underline{\alpha}_t$ is less than the noise in the measurements y_t.

4.4 Outliers detection and accomodation.

In practice, the hyper-parameters are estimated from the data and the Kalman filter is used conditional on the estimated values of the hyper-parameters. In the application of the first stochastic model to real series (section 6), we found that outliers in the data strongly influence the estimation of both the hyper-parameters and the state vectors. In both cases, a strategy for outliers detection and accommodation has to be used.

Whenever an observation is identified as an outlier, its innovation (v_t) is set equal to its expectation, namely zero, and the Kalman gain

$(P_{t|t-1}z_t/f_t)$ is also set to zero. That is, the observation is treated as if it was missing but counted in the total number of observations.

Once the hyper-parameters are estimated, an estimate of σ^2 is available. It can be shown that the innovation sequence v_t~NID$(0,\sigma^2 f_t)$. The outlier identification for the purpose of estimating the state vector α_t is straightforward. Any observation whose innovation is outside a confidence interval ($2.5\sigma f_t^{1/2}$ in the application described in section 6) built around zero is declared an outlier and its value is set equal to its closest bound.

4.5 <u>A Test Procedure for the Selection of the Stochastic Model.</u>

Based on the assumption that the daily coefficients change slowly through time, the initial state vector is obtained with Method 1 of section 4.3. A test procedure to select the more adequate stochastic model for a given series is applied. The test is based on the hypothesis that $\rho_0 = \underline{0}$ in the initial state vector which is estimated assuming deterministic coefficients over the first K observations. (K=36 in the Application section). Such a test is easily performed and it is not discussed any longer (ref. Drapper and Smith (1981) section 2.10).

5. Simulations.

5.1 <u>Numerical Methods for the Estimation of the Hyper-Parameters.</u>

Given the importance of the hyper-parameters value in the estimation procedure of the two stochastic models a pilot test was carried out for three well known algorithms on a set of four simulated series with different values of σ^2_χ/σ^2 and σ^2_ψ/σ^2. The series were generated using the initial state vector α_0 = (-25, -20, -15, 20, 25, 30) and σ^2=25. The numerical methods were programmed with SAS using PROC MATRIX.

Table 1A shows the hyper-parameter values obtained for the three models, deterministic, random walk (called Model One) and random walk with a random drift (called Model Two). The deterministic model is estimated with the regression approach, the Model One hyper-parameters are estimated with the Fibonacci line search and Newton-Raphson; the Model Two hyper-parameters are estimated using the Davidon-Fletcher-Powell and the Newton-Raphson algorithms. Table 1A shows that there are no large differences in the hyper-parameter values given by the three algorithms

except in case 3 between the Fibonacci and Newton-Raphson algorithms. This discrepancy is reflected in table 1B showing that the mean squared error between the simulated series and the estimated values favours the Fibonacci algorithm. Both tables also indicate how the deterministic model deteriorates if the series is generated from a random walk model with a random drift (cases 3 and 4). Finally, the CPU time (in seconds) needed for the various procedures is shown in table 1C. For Model One, the Fibonacci algorithm takes less time in cases 3 and 4. For Model Two, the Davidon-Fletcher-Powell algorithm takes less time in cases 1 and 2.

The Newton-Raphson algorithm involves the evaluation of the second derivatives of the log-likelihood function, the Davidon-Fletcher-Powell algorithm involves the first derivatives and no derivatives are needed for the Fibonacci line search. Mainly because of implementation considerations, we selected the Fibonacci algorithm in Model One and the Davidon-Fletcher-Powell algorithm for Model Two. These two algorithms are applied on a larger set of values for the hyper-parameters in the next section.

5.2 Validation of the models.

A larger study has been done to evaluate the performance of the two stochastic and the deterministic models for the trading-day coefficients on simulated data. The data were simulated with a fixed initial vector $\underline{\alpha}_0 = (-25, -20, -15, 20, 25, 30)$ and varying values of σ^2, σ^2_χ/σ^2 and σ^2_ψ/σ^2. Each simulated series corresponds to a monthly series from January 1977 to December 1986. On a given set of data, the adjustment using the three methods is compared with the adjustment using the true values of the initial state vector and the hyper-parameters. The comparison criteria are the estimates of the hyper-parameters and the mean squared error (MSE) of the adjustment defined as:

$$MSE = 1/120 \sum_{t=1}^{120} (y_t - \hat{y}_t)^2 \qquad (5.1)$$

The significance level of the test procedure described in section 4.5 was also computed. Results are presented in Table 2.

Cases 1 to 10 simulate the deterministic models i.e. σ^2_χ/σ^2 and σ^2_ψ/σ^2 are zero. The true values are obtained by using the Kalman filter and fixed interval smoother of Model One with σ^2_χ/σ^2 equal to zero. In the deterministic model, the estimates of σ^2 are quite close to the actual

values of σ^2, the MSE's are increasing as σ^2 increases but are all quite low. In Model One most of the estimates of the hyper-parameter σ^2_χ/σ^2 are equal to .000313 which is the smallest value that could be obtained with the Fibonacci algorithm using 16 evaluations of the log-likelihood function. In all but case 10, the significance level of the test described in section 4.5 is below 10%. In Model Two, all the estimates of the hyper-parameters σ^2_χ/σ^2 and σ^2_ψ/σ^2 are zero. This is simply because the algorithm did not move from its initial starting value, which is obtained by a grid search in the square delimited by the points (0,0), (0,.5), (.5,0) and (.5,.5).

Cases 11 to 26 simulate the random walk model, i.e. Model One or Model Two with σ^2_ψ/σ^2 equal to zero. In those cases the true values are obtained by using the Kalman filter and fixed interval smoother of Model One with the actual values of the hyper-parameters and the initial state vector. When the hyper-parameter σ^2_χ/σ^2 is very small (Cases 11, 12, 16) the deterministic model works well. In the other cases the deterministic model deteriorates as σ^2_χ/σ^2 increases. The adjustment obtained with Model One is quite good when the MSE is compared with the MSE of the true values. As in the deterministic cases, the estimates of the hyper-parameters σ^2_χ/σ^2 and σ^2_ψ/σ^2 in Model Two did not move from their initial starting values.

Cases 27 to 42 simulate the random walk model with random drift where σ^2_χ/σ^2 equals zero in all cases. Although not shown here (for space reasons) trading-day coefficients generated from this model where very unrealistic. This is reflected in the MSE of the deterministic and Model One adjustments. It has to be noticed that the estimate .99967 of the hyper-parameter σ^2_χ/σ^2 in Model One is the largest value that could be obtained with the Fibonacci algorithm with 16 measurements. Even if the fit obtained under Model Two is acceptable in comparing the MSE with the MSE under the true values the estimates of the hyper-parameters are usually quite poor.

These results seem to indicate that for real applications the random walk model (Model One) and the deterministic model should be adequate for most of the cases observed.

5.3 <u>Statistical Properties of the Estimation Method for the Random Walk Model (Model One).</u>

This section concentrates on the statistical properties of the

estimation methods for the initial state vector $\underline{\alpha}_0$ and the hyper-parameters σ^2_χ/σ^2 and σ^2 of the random walk model. Three methods for the estimation of the initial state vector $\underline{\alpha}_0$ were presented in section 4.3.1. From those methods, five sets of estimates of the hyper-parameters are available. We shall call them Method 4 to Method 8 and they are:

i) Method 4: Use $\underline{a}_0$ and P_0 of Method 1 and estimate the hyper-parameters applying the Fibonacci line search algorithm on the y series.

ii) Method 5: The hyper-parameters of the y series are estimated by the estimated hyper-parameters of the w series used to derive $\underline{\alpha}_0$ in Method 2.

iii) Method 6: Using $\underline{a}_0$ of Method 2, estimate the hyper-parameters applying the Fibonacci algorithm on the y series.

iv) Method 7: Same as Method 5 but with Method 3 instead of Method 2.

v) Method 8: Same as Method 6 but with Method 3 instead of Method 2.

A simulation is done by generating 100 replicates with fixed $\underline{\alpha}_0 = (-0.8, -0.4, 0, 0.3, 0.6, 0.4)$, $\sigma^2_\chi/\sigma^2 = .0028$ and $\sigma^2 = .21$. Each replicate is a time series of 120 observations corresponding to a monthly series from January 1977 to December 1986. For each replicate the eight methods, described above, are applied by taking the first five years and adding one year at a time until the ten years are used. Since $\underline{\alpha}_0$ is a vector, the mean absolute deviation (MAD) of $\underline{a}_0$ defined by:

$$MAD = 1/6 \sum_{i=1}^{6} |a_{0i} - \alpha_{0i}| \qquad (5.2)$$

is computed for each method (1,2,3) in each replicate. The mean of the MAD and the standard error of the mean by number of years of data used were computed over all replicates. Results are given in Table 3A. The mean, the standard error and the 95% confidence intervals of σ^2_χ/σ^2 are given in Table 3B and 3C respectively. The mean and standard error of the mean of σ^2 are given in Table 3D.

Looking at Table 3A it might be surprising that the entries under Method 1 are different. However, this occurs because the regression is done over the first 36 non-outliers observations, and depending on the length of the series the outliers are not the same. For the purpose of estimating the initial state vector $\underline{\alpha}_0$, it can be seen that Methods 2 and 3 are much better than Method 1. Methods 2 and 3 are both comparable. The advantage of Method 2 over Method 3 is that the former does not need the evaluation of the initial state vector $\underline{\alpha}^{\#}_0$ of the w series since it is set

to zero.

From Table 3C it can be seen that the value .0028 of the hyper-parameter σ^2_χ/σ^2 is contained in the confidence intervals of Method 4 (9 Years), Method 5 (All Years), Method 6 (All Years), Method 7 (7 to 10 Years) and Method 8 (9 Years). Methods 4 and 8 (except 9 Years) underestimate the parameter σ^2_χ/σ^2. From Table 3B it has to be noted that the standard errors for the mean for Methods 5 and 6 are larger than the other methods. This explains why the confidence intervals under Methods 5 and 6 do contain the true value of .0028. The effect of the initial states vectors $\underline{\alpha}^{\#}_0$ and $\underline{a}_0$, and their MSE matrices $P^{\#}_0$ and P_0 can be seen by comparing Method 5 with Method 7 and Method 6 with Method 8. In both cases the mean of the estimates of the hyper-parameter under Methods 5 and 7 are much larger than under Methods 6 and 8.

From Table 3D it is obvious that the estimator of σ^2 is biased downward. Furthermore, the value .21 is not included in any of the 95% confidence intervals constructed around the means.

6. Application.

The deterministic and the two stochastic models for the estimation of the trading-day component were applied on a large sample of real series affected by trading-day variations. The series belonged to the sectors of Retail Trade, Wholesale Trade, Imports and Exports.

All the series were first seasonally adjusted using the X-11-ARIMA method without ARIMA extrapolations and assuming the multiplicative model. Therefore, the input y series is the irregular series I obtained from Table B-13 and transformed by the following equation:

$$y_t = (I_t/100 - 1)N_t \qquad (6.1)$$

where N_t is the number of days in month t.

The test procedures for the selection of the stochastic model lead, in most cases, to the acceptance of the null hypothesis. Also, in most cases, the estimation of the hyper- parameters σ^2_χ/σ^2 and σ^2_ψ/σ^2 under the second stochastic model gave the initial starting value (0,0).

In this section we shall discuss two representative cases. The first case is the series of Total Retail Trade Sales for Department Stores in Canada (D650062) from January 1977 to December 1986 and the second case

is the series of Total Retail Trade Sales for All Stores in Nova-Scotia (D650350) from January 1977 to December 1986. The data (ref: equation 6.1) for both series are given in Tables 4A and 4B.

Method 2 and Method 6 are selected for the estimation of the initial state vector and hyper-parameters of the random walk model. The hyper-parameters are computed with and without the outliers strategies (section 4.4) using 8, 9 and 10 years of data for both series. The results are given in Table 5.

It is apparent that the outliers replacement reduces the estimates of the hyper-parameter σ^2_χ/σ^2 since too large innovations will be reduced. It can also be seen for the Total Retail Trade - Canada that the outliers replacement stabilizes the estimates of σ^2_χ/σ^2 (This was also noticed in the other cases studied).

Graphs of the daily coefficients for both series are provided in figures 1A and 1B. Clearly the random walk model, with or without smoothing, allows for a moving behaviour of the daily coefficients. In the case of total Retail Trade-Canada, the estimate of the hyper-parameter σ^2_χ/σ^2 has a value of .00031 and the smoothed daily coefficients are almost a straight line. On the other hand, the hyper-parameter σ^2_χ/σ^2 of total Retail Trade-Nova Scotia has a value of .00971 and the smoothed daily coefficients show a well marked evolution through the years.

The trading-day coefficients and components were computed under the deterministic and the random walk models with and without smoothing to investigate their behaviour. Three tests of hypotheses were done; namely, H1) the trading-day components obtained under the random walk model with smoothing are equal to the trading-day component under the deterministic model, H2) the trading-day components obtained under the stochastic model without smoothing are equal to the trading-day components obtained under the deterministic model, and H3) the trading-day components obtained under the random walk model without smoothing are equal to the trading-day components obtained under the random walk model with smoothing. In the above three hypotheses, the second set of coefficients was considered to be known and fixed. Graphics for the three tests are provided in figure 2. The upper line (UL) and lower line (LL) give the 95% confidence intervals defined as follows:

$$UL = y_t^a - y_t^f + 1.96(var(y_t^a))^{1/2} \qquad (6.2)$$
$$LL = y_t^a - y_t^f - 1.96(var(y_t^a))^{1/2}$$

where for H1, y_t^a is the trading-day component under the random walk model with smoothing and y_t^f is the estimated trading-day component under the deterministic model. Similar equations are defined for H2 and H3. The graphics show that when the hyper-parameter σ^2_χ/σ^2 is large (case D650350), the hypotheses H1 and H2 are rejected. In fact, for H1 and H2 there are respectively 13 and 12 time points excluding zero. Notice that there are no trading-day components in February during non-leap year and this is why the two lines meet at zero. For the D650062 series, none of the hypotheses can be rejected.

Next a frequency demain analysis was conducted on the input series and the trading-day adjusted series defined as the series obtained by subtracting the estimated trading day component from the original series. Three estimates of the trading-day component were compared, they are: the deterministic, the random walk without smoothing and the random walk with smoothing. Graphs of the spectral densities are given in figures 3A and 3B.

The spectral density of the input series are characterized by a peak at the frequencies around .348. This is typical of such series as shown by Cleveland and Devlin (1980).

For the series Total Retail Trade - Canada, the spectral densities of the trading-day adjusted series are all similar and the power around the .348 frequency has been removed from the three adjusted series. For this series it does not seem that a more sophisticated method than the deterministic model is needed to remove the trading-day component. This is indicated by the small value of the estimate of σ^2_χ/σ^2 and by the spectral densities of the adjusted series.

For the series Total Retail Trade - Nova Scotia, the powers at the frequencies around .348 of the trading adjusted series are quite different. It can be seen that the power at the .35 frequency is reduced from 37.66 for the original series to 6.97 for the deterministic model, to 3.65 for the random walk without smoothing and to 2.84 for the random walk with smoothing.

7. Conclusions.

In this study a random walk model and a random walk model with random drift for trading-day coefficients are introduced and compared with the deterministic model used in X-11-ARIMA. We provide a solution to the problem of estimating the hyper-parameters and initial conditions of the stochastic models. A simulation with a large number of series indicated that in real applications the random walk model and the deterministic model should be adequate for most cases observed. The simulation also showed that the MLE of σ^2 tends to be biased downward but the estimator of σ^2_χ/σ^2 is unbiased in the random walk model.

Two real case studies are thoroughly discussed to illustrate where the deterministic and the random walk models can be adequate.

Bibliography

Bazaraa,S.B., and Shetty,C.M., (1979), **Nonlinear Programming, theory and algorithms**, New-York, Wiley.

Cleveland,W.S., and Devlin,S.J., (1980), "Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods", **JASA**, 75, 487-496.

Dagum,E.B., (1980), "The X-11-ARIMA Seasonal Adjustmemt Method", **Statistics Canada**, Catalogue No. 12-564E

Draper,N.R., and Smith,H., (1981), **Applied Regression Analysis**, New-York, Wiley.

Harvey,A.C., (1981), **Time Series Models**, Oxford, Philip Allan.

Monsell,B.C., (1983), "Using the Kalman Smoother to Adjust for Moving Trading Day", SRD Research Paper No: CENSUS/SRD/RR-83/04, Bureau of the Census.

Shiskin,J., Young,A.H. and Musgrave,J.C., (1967), "The X-11 Variant of Census Method II Seasonal Adjustment", Technical Paper No 15, **Bureau of the Census, U.S. Department of Commerce**.

Young,A.H., (1965), "Estimating Trading-Day Variation in Monthly Economic Time Series", Technical Paper No 12, **Bureau of the Census, U.S.** Department of Commerce.

TABLE 1
Simulation for Numerical Methods

1A Estimates of the Hyper-Parameters

Case	Input			DM[1]	Model One				Model Two					
					Fibo[2]		NR[3]		DFP[4]			NR[3]		
	σ^2	σ^2_χ/σ^2	σ^2_ψ/σ^2	$\hat{\sigma}^2$	$\hat{\sigma}^2$	$\hat{\sigma}^2_\chi/\sigma^2$	$\hat{\sigma}^2$	$\hat{\sigma}^2_\chi/\sigma^2$	$\hat{\sigma}^2$	$\hat{\sigma}^2_\chi/\sigma^2$	$\hat{\sigma}^2_\psi/\sigma^2$	$\hat{\sigma}^2$	$\hat{\sigma}^2_\chi/\sigma^2$	$\hat{\sigma}^2_\psi/\sigma^2$
1	25	0	0	22.44	20.10	.00031	19.98	.00023	19.97	0	0	17.98	0	0
2	25	.01	0	24.48	22.14	.00031	21.90	.00023	24.32	0	0	21.10	0	0
3	25	0	.001	166.76	13.18	.38100	137.24	.00076	12.38	.25	0	11.85	.25	0
4	25	.01	.001	159.56	13.88	.37350	14.78	.35526	12.80	.25	0	12.29	.25	0

(1) DM: Deterministic Model
(2) FIBO: Fibonacci Algorithm
(3) NR: Newton-Raphson Algorithm
(4) DFP: Davidson-Fletcher-Powell Algorithm

1B. Mean Squared Error

Case	DM	Model One		Model Two	
		FIBO	NR	DFP	NR
1	1.51	1.61	1.59	2.33	2.24
2	2.99	2.81	2.86	2.83	2.73
3	10.72	3.43	9.73	3.40	3.47
4	10.64	3.42	3.39	3.43	3.44

1C. SAS CPU Time (Seconds)

		Model One		Model Two	
Case	DM	FIBO	NR	DFP	NR
1	.24	2.55	2.59	7.62	21.98
2	.23	2.52	2.55	19.81	22.09
3	.24	2.53	2.64	19.82	15.79
4	.24	2.53	16.01	19.74	15.81

TABLE 2

VALIDATION OF THE MODELS

Case	INPUT			S.L. %	TRUE VALUES		DETERMINISTIC		MODEL ONE			MODEL TWO			
	$\sigma^2_\chi\|\sigma^2$	$\sigma^2_\psi\|\sigma^2$	σ^2		θ^2	MSE	θ^2	MSE	$\hat{\sigma}^2_\chi\|\sigma^2$	θ^2	MSE	$\hat{\sigma}^2_\chi\|\sigma^2$	$\hat{\sigma}^2_\psi\|\sigma^2$	$\hat{\sigma}^2$	MSE
1	0	0	100	83.17	100.06	0	95.98	2.97	.000313	92.25	3.12	0	0	89.49	3.79
2	0	0	100	97.16	87.18	0	84.10	2.69	.000313	79.79	2.81	0	0	75.85	3.48
3	0	0	50	58.18	41.68	0	41.21	1.59	.00093	39.22	1.77	0	0	37.91	2.51
4	0	0	50	95.50	51.10	0	51.29	1.54	.000313	49.86	1.22	0	0	48.94	1.81
5	0	0	25	43.97	21.11	0	21.44	.86	.000313	20.59	.88	0	0	20.83	1.37
6	0	0	25	28.55	28.53	0	29.42	.76	.000313	27.65	.84	0	0	29.30	1.09
7	0	0	2.5	39.39	2.46	0	2.42	.39	.000313	2.34	.45	0	0	2.37	.54
8	0	0	2.5	61.69	2.47	0	2.45	.37	.000313	2.32	.40	0	0	2.28	.50
9	0	0	.025	13.82	.0222	0	.0228	.023	.000313	.0216	.026	0	0	.0228	.0344
10	0	0	.025	.5	.0268	0	.027	.031	.0034	.0222	.055	0	0	.0268	.058
11	.00025	0	100	42.68	110.51	1.63	114.85	2.46	.000313	108.70	2.63	0	0	107.88	2.82
12	.0005	0	50	93.64	48.01	1.14	51.33	1.71	.00093	45.80	1.73	0	0	43.65	2.32
13	.001	0	25	22.64	29.03	.72	27.89	2.08	.00093	25.36	1.87	0	0	27.04	2.06
14	.01	0	2.5	3.38	2.33	.56	4.17	1.37	.029	1.72	.64	0	0	2.34	.69
15	1	0	.025	.061	.025	.14	4.49	2.06	.833	.033	.14	.5	0	.052	.14
16	.0025	0	100	57.53	96.29	2.90	106.72	2.96	.000313	101.15	2.64	0	0	98.60	3.69
17	.005	0	50	66.23	47.07	2.11	61.48	3.34	.00907	43.26	2.32	0	0	51.99	2.50
18	.01	0	25	89.70	26.83	1.73	41.49	3.76	.0234	20.69	2.22	0	0	26.74	2.28
19	.1	0	2.5	14.73	2.22	.94	14.75	3.65	.046	2.73	1.05	.25	0	1.67	.96
20	.025	0	100	8.82	104.99	4.63	324.94	13.36	.0165	109.129	4.81	0	0	164.06	7.12

TABLE 2 - Cont'd.

VALIDATION OF THE MODELS

Case	INPUT			S.L. %	TRUE VALUES		DETER-MINISTIC		MODEL ONE			MODEL TWO			
	$\sigma^2_\chi \mid \sigma^2$	$\sigma^2_\psi \mid \sigma^2$	σ^2		θ^2	MSE	θ^2	MSE	$\hat{\sigma}^2_\chi \mid \sigma^2$	θ^2	MSE	$\hat{\sigma}^2_\chi \mid \sigma^2$	$\hat{\sigma}^2_\psi \mid \sigma^2$	$\hat{\sigma}^2$	MSE
21	.05	0	50	29.81	57.61	4.26	167.94	10.16	.0372	61.09	4.47	0	0	98.09	7.00
22	.1	0	25	26.47	19.58	3.17	205.85	13.02	.064	22.52	3.20	.25	0	13.64	3.22
23	1	0	2.5	70.37	2.87	1.36	91.13	8.98	.41	4.82	1.39	.5	0	4.31	1.36
24	.25	0	100	0	89.97	7.58	1935.25	42.04	.1938	109.29	7.53	.25	0	100.87	7.75
25	.5	0	50	..001	47.38	4.92	22371	44.79	.7385	35.61	4.96	.5	0	57.33	4.99
26	1	0	25	12.85	24.59	4.21	6876.79	81.29	.9996	25.08	4.24	.5	0	43.76	4.54
27	0	.00025	100	.13	110.73	4.79	3803.64	59.14	.352	68.06	7.40	.25	0	57.35	7.13
28	0	.0005	50	0	50.23	2.62	6370	77.89	.36913	39.74	4.48	.25	0	33.65	4.27
29	0	.001	25	0	18.80	3.31	3222	55.17	.62	11.93	4.07	.5	0	12.17	4.02
30	0	.01	2.5	0	2.74	1.24	2140	44.83	.99967	5.18	1.57	0	.005	3.35	1.35
31	0	1	.025	0	.024	.134	3304	56.03	.99967	11.66	.620	1	.207	.1007	.1547
32	0	.0025	100	0	72.44	5.06	10016	96.39	.99967	52.57	7.13	.052	.0016	58.00	5.54
33	0	.005	50	0	51.34	4.50	32646	175.65	.99967	120.31	5.98	1	.06	11.59	5.92
34	0	.01	25	0	18.59	3.01	6873	80.05	.99967	25.49	3.84	.28	.007	12.13	3.55
35	0	.1	2.5	0	2.36	1.13	38311	190.78	.99967	90.43	1.63	.62	.047	2.29	1.33
36	0	.025	100	0	104.47	6.99	160899	390	.99967	365.76	8.52	.48	.02	66.79	7.73
37	0	.05	50	0	64.33	5.41	216242	453	.99967	594.55	7.87	.41	.027	58.03	6.06
38	0	.1	25	0	24.95	4.05	220361	457	.99967	615	6.42	.79	.10	15.69	4.36
39	0	1	2.5	0	2.76	1.33	322922	554	.99967	1223	5.01	1	.24	10.50	1.70
40	0	.25	100	0	99.17	8.69	4300301	2020	.99967	5445	15.99	1	.245	90.92	9.20
41	0	.5	50	0	57.22	5.36	3775864	1894	.99967	10310	16.61	1	.278	85.02	6.01
42	0	1	25	0	23.88	4.49	801191	872	.99967	5007.86	11.79	1	.24	84.97	4.99

Table 3A

Mean and Standard Error of the Mean for MAD by Number of Years and Methods

No. Years	Method 1 Mean	Method 1 STD	Method 2 Mean	Method 2 Std	Method 3 Mean	Method 3 Std
5	.193	.760	.163	.576	.160	.540
6	.194	.793	.162	.571	.162	.561
7	.190	.782	.155	.517	.157	.538
8	.189	.768	.151	.492	.151	.526
9	.192	.762	.149	.489	.149	.513
10	.193	.748	.147	.497	.145	.508

Table 3B

Mean[1] and Standard Error of the Mean[2] for $\hat{\sigma}^2_{\chi}/\sigma^2$ (.0028) by Number of Years and Methods

no. Years	Method 4 Means	Method 4 STD	Method 5 Means	Method 5 STD	Method 6 Means	Method 6 STD	Method 7 Means	Method 7 STD	Method 8 Means	Method 8 STD
5	0.180	0.232	0.834	3.081	0.520	2.378	0.193	0.272	0.184	0.284
6	0.200	0.220	0.696	2.368	0.411	1.595	0.213	0.244	0.188	0.254
7	0.217	0.208	0.470	1.013	0.293	0.607	0.265	0.248	0.210	0.224
8	0.224	0.201	0.409	0.772	0.252	0.428	0.254	0.244	0.205	0.231
9	0.253	0.228	0.398	0.637	0.257	0.384	0.291	0.302	0.232	0.287
10	0.242	0.192	0.371	0.704	0.244	0.409	0.262	0.230	0.207	0.251

(1) Entries have to be multiplied by 10^{-2}

(2) Entries have to be multiplied by 10^{-3}

Table 3C

95% Confidence Intervals[1] of $\hat{\sigma}^2_x/\sigma^2$ (.0028)
by Number of Years and Methods

No.	Method 4		Method 5		Method 6		Method 7		Method 8	
Years	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
5	0.135	0.226	0.230	1.438	0.054	0.986	0.140	0.246	0.128	0.240
6	0.157	0.243	0.232	1.160	0.099	0.724	0.165	0.261	0.138	0.238
7	0.176	0.257	0.272	0.669	0.174	0.412	0.217	0.314	0.166	0.254
8	0.184	0.263	0.258	0.561	0.168	0.336	0.206	0.302	0.160	0.250
9	0.208	0.298	0.273	0.523	0.182	0.333	0.231	0.350	0.176	0.289
10	0.204	0.279	0.233	0.509	0.163	0.324	0.216	0.307	0.158	0.256

(1) Entries habe to be multiplied by 10^{-2}

Table 3D

Mean and Standard Error[1] of the Mean for $\hat{\sigma}^2$ (.21)
by Number of Years and Methods

no.	Method 4		Method 5		Method 6		Method 7		Method 8	
Years	Means	STD	Means	STD	Means	STD	Means	STD	Means	STD
5	0.174	0.480	0.162	0.489	0.166	0.481	0.174	0.471	0.170	0.464
6	0.176	0.424	0.169	0.426	0.168	0.426	0.175	0.433	0.172	0.420
7	0.179	0.384	0.170	0.380	0.173	0.377	0.178	0.381	0.176	0.375
8	0.179	0.366	0.173	0.361	0.177	0.363	0.179	0.360	0.179	0.357
9	0.179	0.334	0.173	0.330	0.116	0.334	0.178	0.328	0.177	0.333
10	0.179	0.299	0.173	0.299	0.177	0.302	0.178	0.302	0.178	0.299

(1) Entries have to be multiplied by 10^{-2}

Table 4A

Total Retail Trade Sales - Department Stores -Canada
Transformed Irregular Series[1]

Year	Jan	Fev	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
77	-225	270	-243	871	-721	-765	234	288	490	-562	309	-173
78	-338	-65	810	-315	81	-31	-827	-202	1473	-652	78	217
79	-396	-731	1833	-1233	168	119	-655	1230	-598	-12	102	-585
80	962	470	-780	-440	1181	-1680	938	-8	170	248	-600	-565
81	1955	-437	-1076	352	-358	611	-149	-273	-317	1014	-362	298
82	-315	105	-569	603	-273	-642	623	-270	421	-370	-157	481
83	-240	-24	1476	-1735	-1455	1630	104	8	215	-353	-649	886
84	-719	897	63	-941	-27	1143	-1006	-222	293	-213	830	-514
85	118	-504	537	257	528	-1401	-090	1451	-995	-116	756	-676
86	738	-603	-87	492	1390	-1917	453	100	-191	436	-405	214

(1) Entries have to be multiplied by 10^{-3}

Table 4B

Total Retail Trade Sales - All Stores - Nova Scotia
Transformed Irregular Series[1]

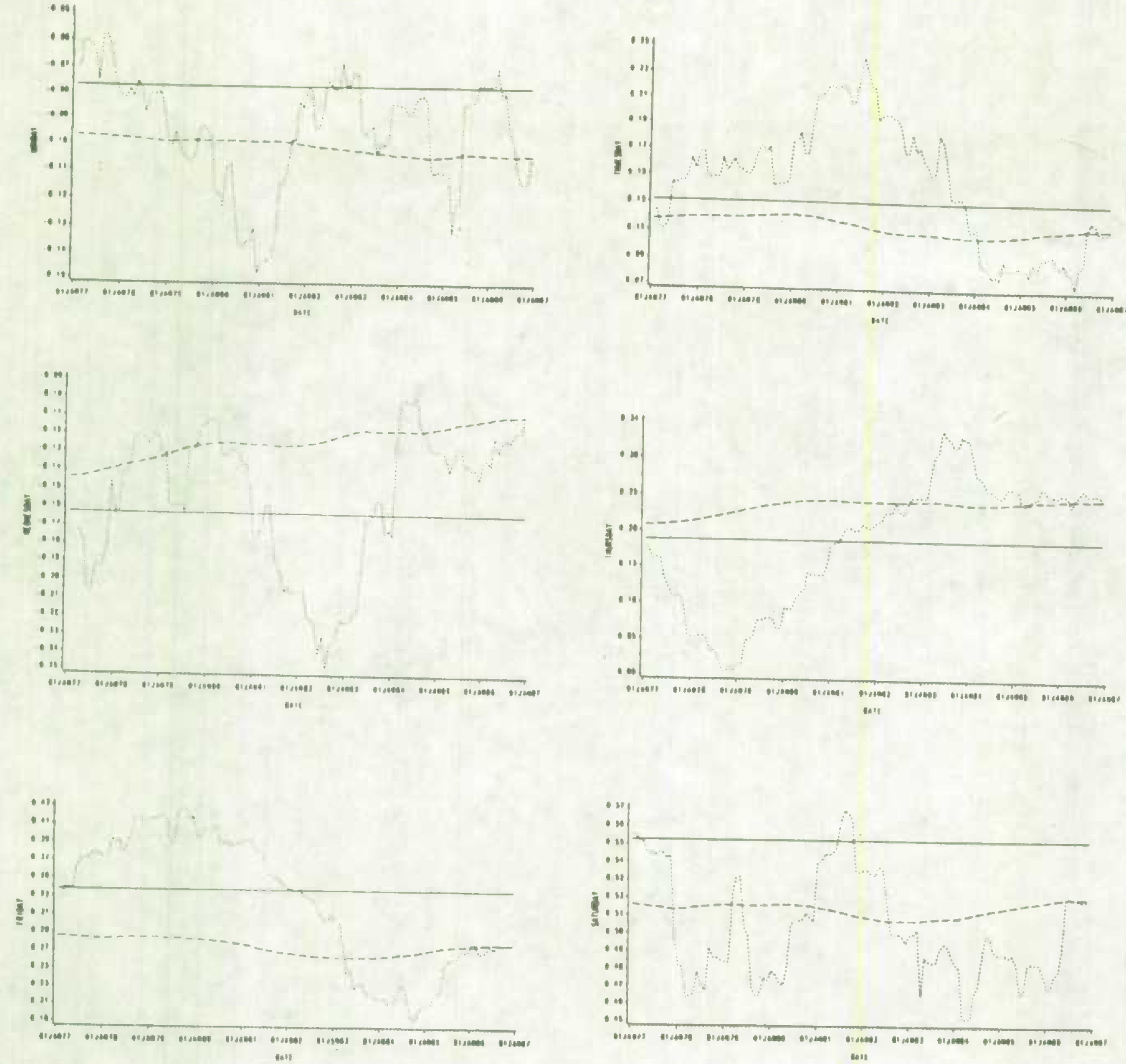
Year	Jan	Fev	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
77	161	-221	-300	574	-641	-363	480	-583	447	-223	342	693
78	-883	-34	243	-559	-11	558	-182	-99	585	-503	-155	-36
79	272	-428	935	-1113	171	1062	-921	1344	-1287	-230	1136	-954
80	44	712	-756	-136	1212	-903	-62	488	-275	696	-333	-1167
81	720	499	-657	695	409	-651	233	11	-817	514	-239	490
82	84	-443	-230	1251	-1097	0	682	-732	293	-168	-177	874
83	-563	-796	1613	-640	-1307	1451	73	-513	639	-739	-631	958
84	-865	638	407	-552	330	795	-1350	383	-206	-337	910	-928
85	693	-288	285	186	128	-475	-139	1008	-865	30	530	-1289
86	1514	161	-1202	705	608	-438	-184	-788	393	777	-640	-194

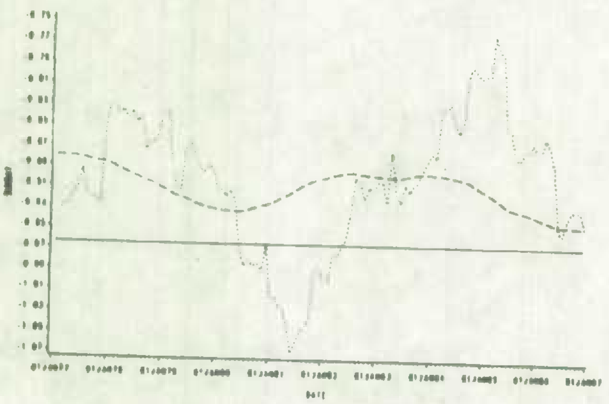
(1) Entries have to be multiplied by 10^{-3}

Table 5

Hyper Parameters Estimates

	Total Retail Trade Sales Department Stores - Canada				Total Retail Trade Sales All Stores - Nova Scotia			
	Without Replacement of Outliers		With Replacement of Outliers		Without Replacement of Outliers		With Replacement of Outliers	
no. Years	σ^2_χ/σ^2	σ^2	σ^2_χ/σ^2	σ^2	σ^2_χ/σ^2	σ^2	σ^2_χ/σ^2	σ^2
8	0.01033	0.21730	0.00031	0.21760	0.03225	0.09530	0.02098	0.09670
9	0.00407	0.24340	0.00031	0.20680	0.02912	0.09880	0.01158	0.10890
10	0.00031	0.26330	0.00031	0.20480	0.02098	0.11360	0.00971	0.12230





FIGURES 1A

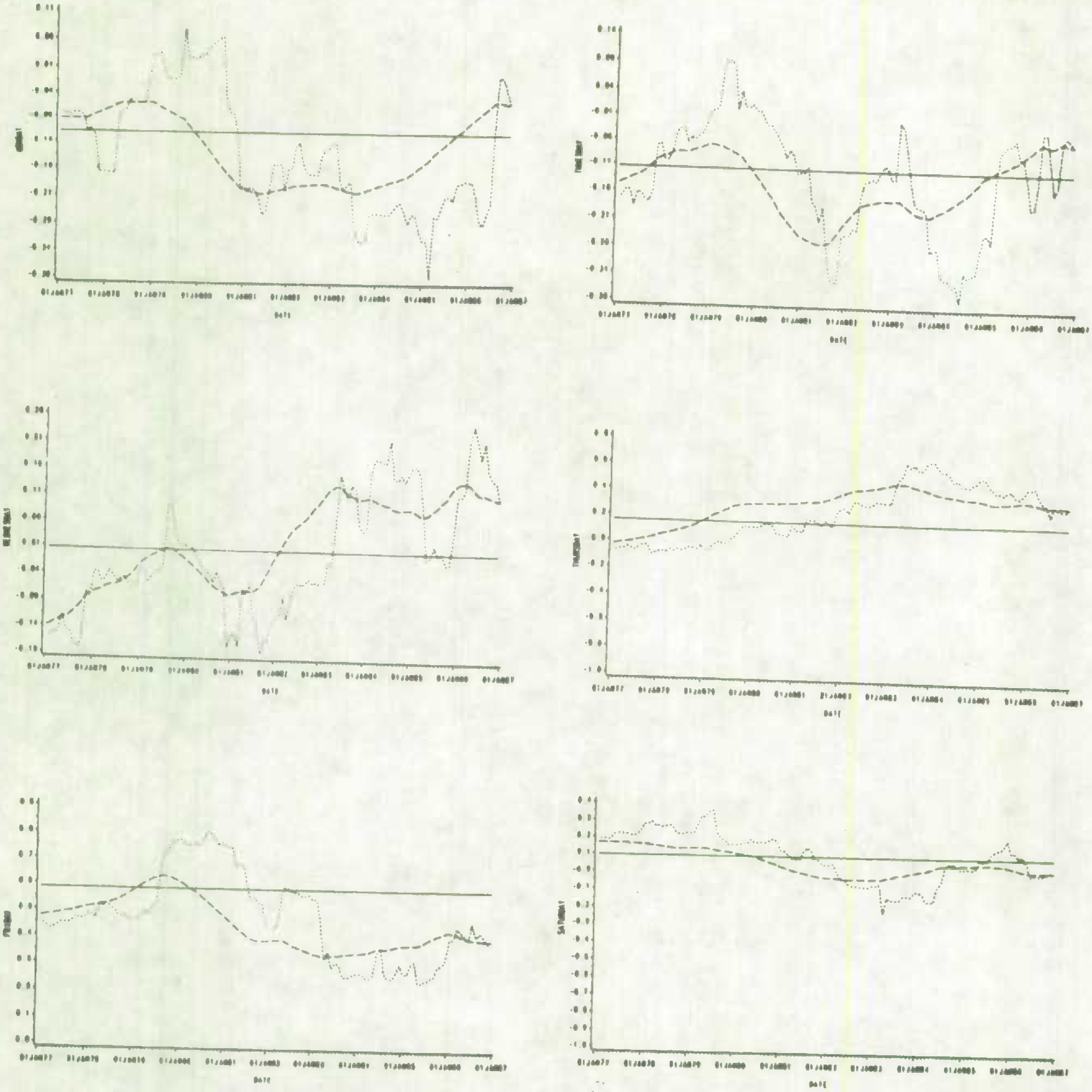
D650062 – DEPARTMENT STORES – CANADA

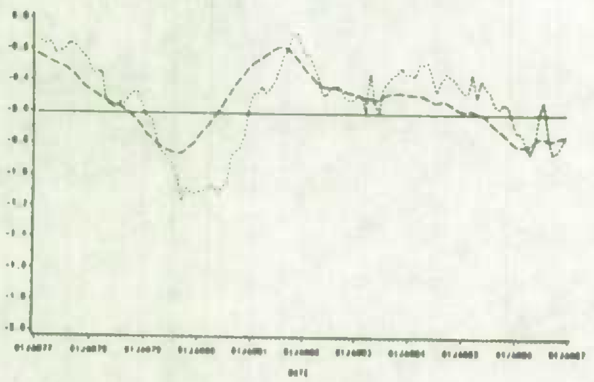
TRADING–DAY COEFFICIENTS

———: deterministic

.........: random walk without smoothing

– – – –: random walk with smoothing





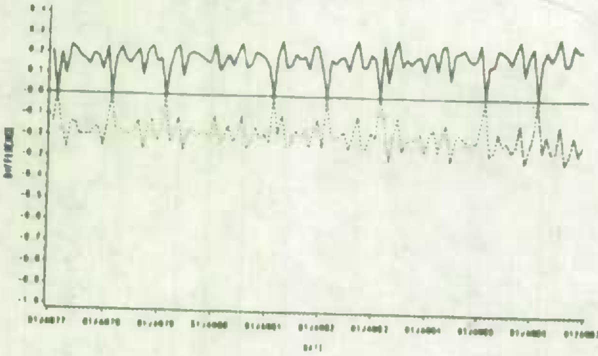
FIGURES 1B
D650350 – TOTAL RETAIL TRADE SALES – NOVA SCOTIA
TRADING–DAY COEFFICIENTS

———: deterministic
………: random-walk without smoothing
– – – –: random walk with smoothing

FIGURES 2
TEST OF HYPOTHESES

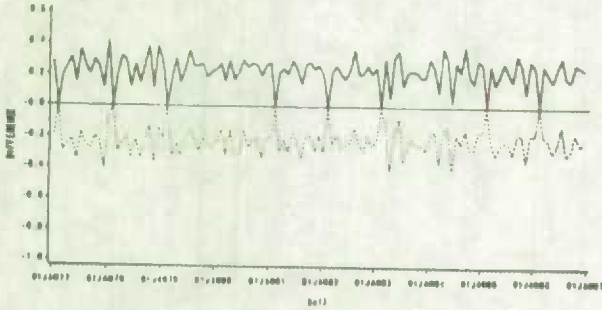
D 650062 – DEPARTMENT STORES SALES – CANADA

TEST OF HYPOTHESES: H1



UL = Random Walk With Smoothing - Deterministic + 1.96STD(RWS)
LL = Random Walk With Smoothing - Deterministic - 1.96STD(RWS)

TEST OF HYPOTHESES: H2



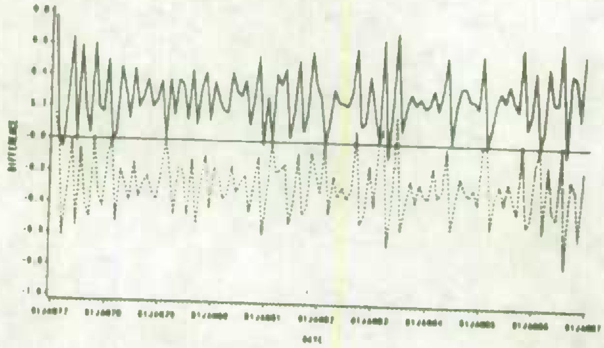
UL = Random Walk Without Smoothing - Deterministic + 1.96STD(RW)
LL = Random Walk Without Smoothing - Deterministic - 1.96STD(RW)

TEST OF HYPOTHESES: H3

UL = RWS - RW + 1.96STD(RW)
LL = RWS - RW - 1.96STD(RW)

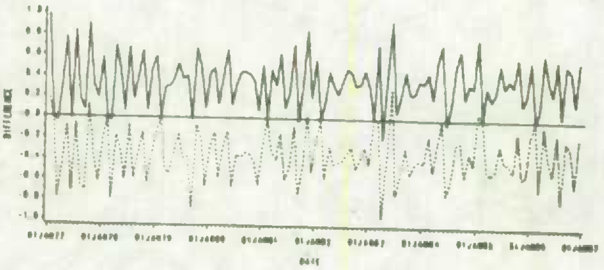
D 650350 – RETAIL TRADE SALES – NOVA SCOTIA

TEST OF HYPOTHESES: H1



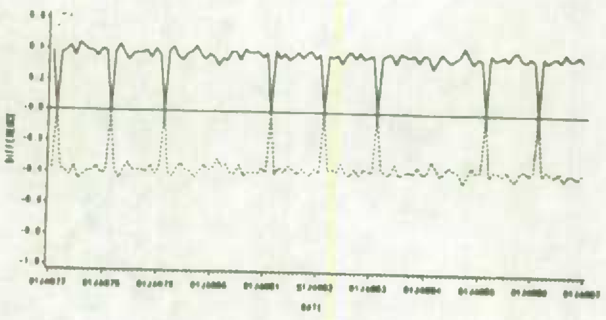
UL = Random Walk With Smoothing - Deterministic + 1.96STD(RWS)
LL = Random Walk With Smoothing - Deterministic - 1.96STD(RWS)

TEST OF HYPOTHESES: H2



UL = Random Walk Without Smoothing - Deterministic + 1.96STD(RW)
LL = Random Walk Without Smoothing - Deterministic - 1.96STD(RW)

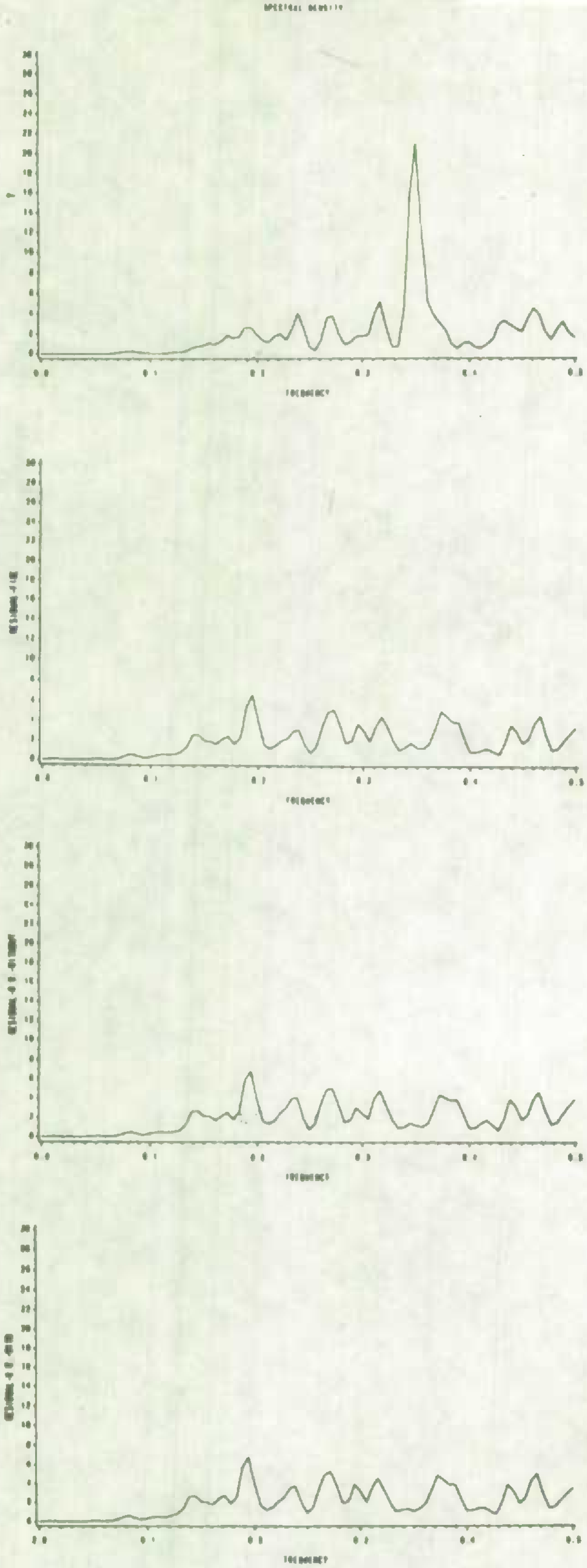
TEST OF HYPOTHESES: H3



UL = RWS - RW + 1.96STD(RW)
LL = RWS - RW - 1.96STD(RW)

FIGURES 3
SPECTRAL DENSITIES

D 650062 DEPARTMENT STORES - CANADA



D 650350 - RETAIL TRADE SALES - NOVA-SCOTIA