

**DÉPARTEMENT DE SCIENCE ÉCONOMIQUE
DEPARTMENT OF ECONOMICS**

CAHIERS DE RECHERCHE / WORKING PAPERS

0306E

**Evaluating Time Series Models
in Short and Long-Term Forecasting of
Canadian Air Passenger Data**

**by
Emir Emiray and Gabriel Rodriguez**

ISSN: 0225-3860



uOttawa

Faculté des sciences sociales
Faculty of Social Sciences

**CP 450 SUCC. A
OTTAWA (ONTARIO)
CANADA K1N 6N5**

**P.O. BOX 450 STN. A
OTTAWA, ONTARIO
CANADA K1N 6N5**

CAHIER DE RECHERCHE #0306E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER #0306E
Department of Economics
Faculty of Social Sciences
University of Ottawa

Evaluating Time Series Models in Short and Long-Term Forecasting of Canadian Air Passenger Data

by

Emir Emiray¹
Transport Canada

and

Gabriel Rodríguez²
Department of Economics
University of Ottawa

First version: February 28, 2003

This version: July 21, 2003

¹Emiray would like to thank the Air Forecasting team at Transport Canada, specifically Robert Duclos, chief - Aviation forecasts, for his encouragement and support throughout the project.

²Rodríguez thanks the Faculty of Social Sciences for financial support. We also thank Josée Villemaire for the revision and correction of the abstract in French. Address for correspondence: P. O. Box 450, Station A, Ottawa, Ontario, Canada, K1N 6N5; Telephone: (613) 562 5800 (1750), Fax: (613) 562 5999, E-mail address: gabrielr@uottawa.ca.

Abstract

This paper uses six time series models to forecast seasonally unadjusted monthly data of Canadian enplaned/deplaned air passengers, for the domestic, transborder and international sectors. We find that forecasting performance of the models varies widely across series and forecast horizons. Our forecasting results are compared with forecasts published by Transport Canada's Aviation Forecasting Division.

Keywords: Forecasts, Seasonal Unit Roots, ARIMA Models, Periodic Autoregressive Models, Structural Time Series Model.

JEL Classification: C2, C5.

Résumé

Ce travail utilise six modèles de séries chronologiques afin de prévoir les données mensuelles (sans ajustement saisonniers) des embarquements/atterrissages des passagers aériens canadiens pour les vols domestiques et internationaux. Nous trouvons que la performance dans les prévisions varie grandement selon les séries et selon l'horizon. Nos prévisions sont comparées avec celles publiées par la division des prévisions d'aviation de Transport Canada.

Mots-clés: Prévisions, racines unitaires saisonnières, modèle ARIMA, modèles périodiques auto-régressifs, modèles de séries chronologiques structurelles.

Code JEL: C2, C5.

1 Introduction

Until the 1980's, studies of aviation economics were largely descriptive rather than analytical. Most publications on air transport, highly institutional and rather restrictive, were concerned with the study of monopoly, regulation, and with infrastructure issues such as airport construction and financing. The first and perhaps most known analytical application was Box and Jenkins (1970) study of "airline passenger model" which applied time series analysis to model and forecast passenger numbers in the United States. In the last two decades, analytical applications to both American and Canadian air industry increased, yet they all focused on panel data, and estimation of price elasticities. Today, the long span of Canadian air data, and improved modelling methods of seasonal macro-economic time series, provide us with opportunities to model and test the forecasting performance of these time series methods. To the knowledge of the authors there are no applications of time series models to the Canadian airline passenger data.

In this paper we use six time series models for seasonality for the Canadian airline passenger data series (for the domestic, transborder and international sectors) and later examine the out-of-sample forecasting performance of each of these models. (should I include pros and cons of each model?) In our model evaluation we utilize out of sample forecast evaluation techniques. Here we find that simple models tend to perform better in the short term while more flexible, more adoptable, models perform better in the long term, periods of level shifts and periods of structural changes. In this respect, when we forecast passenger numbers for the short term (2006) and long term (2015), we take into account these findings. These forecasts are then compared to Transport Canada's short and long term forecasts which are produced using panel data techniques and published yearly and are available online³.

The outline of the paper is as follows. Section 2 discusses the monthly Canadian enplaned/deplaned (hereafter E/D) passenger data and aviation forecasting in Canada. Section 3 discusses models of seasonality and results for each sector. Equations used in model evaluations and long term forecasting are provided in appendix B and C. Section 4 discusses the forecasts evaluation technique and the results for short and long term sector specific E/D forecasts with comparisons to the Transport Canada forecasts. Section 5 concludes.

³<http://www.tc.gc.ca/POL/EN/airforecasting/airforecasting.htm>.

2 Canadian Air Passenger Data

In this section, we describe the data sets used in this paper. The series we consider represent the seasonally unadjusted E/D air passengers in Canada for the domestic, transborder, and international sectors, at the monthly frequency covering the period between January 1984 and September 2002 (225 observations). The series are constructed by adding enplaned and deplaned passenger numbers, which are obtained from Statistics Canada and Transport Canada archives (see appendix A for data definitions and sources).

All three series, which are shown in levels in the first column of graphs in Figure 1 below, contain three common dominant features: distinct seasonal pattern, strong growth during late eighties and mid nineties and, a sharp drops after 1991 recession as well as after September 2001. All the series show pronounced seasonal pattern, as air travel is typically high during the summer time (May-August) and below average during the fall and winter. We also observe that each series have several dominant outliers. For example in the domestic sector, 1990:12, 1999:5 and 2001:9 are obvious striking outliers, while 1990 december and 1991 december are outliers for both the transborder and international sectors.

Looking at the second set of graphs below (second column in Figure 1), which corresponds to the seasonally differenced log-transformed series, one can clearly observe recessions and expansions⁴. For example we can observe from these figures that as the Canadian economy entered a recession in 1991 demand for air travel also slowed down, while between 1995 and 2000 as the Canadian was expanding, so was demand for air travel. What we cannot clearly observe from these graphs are the characteristics of these cyclicalities which may show asymmetric behaviour. Indeed examining the growth rates before recessions and after expansions show that air demand cycles are characterized by steep decreases during recessions, followed by slow(er) increases during expansions. This behaviour is due the fact that air travel demand is composed of business and leisure travel, where both are directly related to GDP and employment rates⁵. It is not surprising that in good times business travel is higher as there is more economic activity while in good times, leisure travel is higher as leisure travel is usually treated as

⁴There is no institution equivalent to the NBER in Canada that publishes the recession and expansion dates, nevertheless one can make educated guesses examining Canadian GDP growth rates.

⁵Macroeconomic variables such as GDP are most common determinants used to explain air travel demand in the literature, which shows evidence of asymmetric behaviour (see for example Dijk et al, 2002).

a luxury good (Brons et al, 2002).

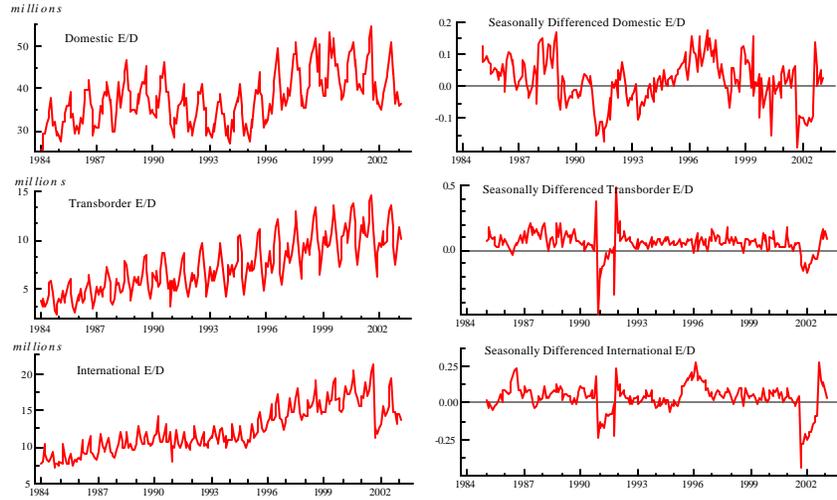


Figure 1. Canadian E/D passenger numbers in levels and Seasonal differences

Figure 2 below shows year annual growth rates of total Canadian E/D passenger data for each month over the sample period⁶. We observe the pronounced periodicity of the process as the mean of annual growth rates of total E/D passengers varies significantly over the months. The mean varies from 2.09% growth in November to 2.66% growth in July. The standard deviation varies from 5.78% in August to 8.25% in February. It is this observation that results in fitting a periodic autoregressive model (PAR) to each series.

⁶ Where Total E/D=Domestic E/D +Transborder E/D + International E/D.

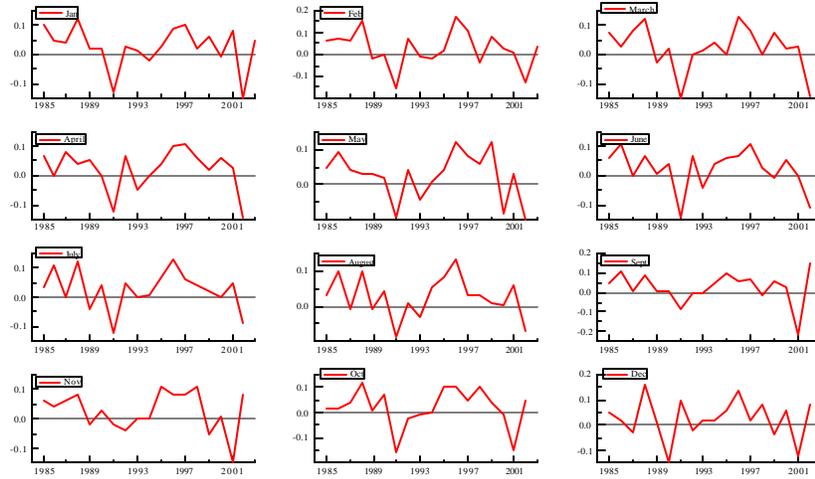


Figure 2. Annual growth rates of Total Air Demand

Lastly, figure 3 below shows the periodic autocorrelations functions for each month for the total E/D passenger data. We observe that even though the autocorrelations look alike for each month, there are obvious differences. This observation leads us to test the seasonal unit root hypothesis for each series and apply a specific differentiation rather than simply assuming unit roots.

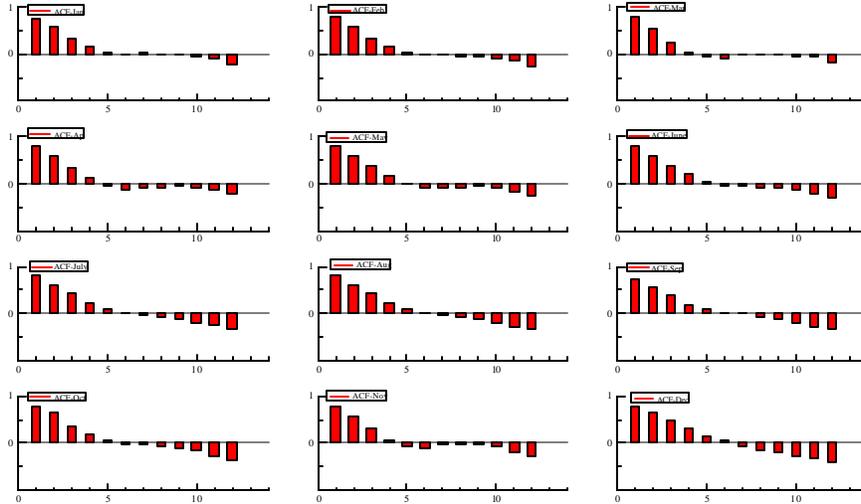


Figure 3. Periodic Autocorrelation functions for each month for the Total E/D passenger data

2.1 Aviation Forecasting in Canada

To the knowledge of the authors there are no time series analysis of Canadian passenger E/D data. Moreover, this study utilizes the longest monthly E/D data for Canada. All of the studies on Canadian air transportation utilize city pair panel data, where the focus lies on determining the elasticities of the air travel determinants⁷. Passenger forecasts are subsequently calculated by using these elasticities on forecasted explanatory variables. In these studies, the price of air travel is the most important determinant of air travel demand, see Gillen et al (2002) for literature review on air travel demand elasticities. There are various advantages of forecasting pax numbers using panel data. First, price elasticities can only be determined using panel data analysis. Policy makers use price elasticities to determine governmental policies such as air travel taxation, while airlines use them for marketing purposes (air travel pricing, etc.). A second advantage of panel studies is that pooling of information for all city pairs allows researchers to use fixed effects to capture city specific information and to calculate more

⁷Most common used determinants of air travel are price of air travel, GDP of origin and destination, air and/or car time of travel between the two cities, city-pair dummies, available flight frequency, personal disposable income (PDI), etc.

meaningful elasticities for each airport of study⁸. Nevertheless, due to the lack of availability of quality data on determinants of air travel, panel data studies of air travel may have major measurement errors⁹. Furthermore, forecasts from panel data studies are calculated using forecasted values of independent variables such as future price of air travel, PDI, etc., which by themselves introduce further uncertainty to the forecasts.

Time series forecasting of passenger numbers could eliminate these biases by exploring the past behavior of the data and assuming that same behaviour persists in the future. This is most useful when researchers are interested in accurate forecasts rather than specific elasticities of air demand determinants. For example, precise forecasts are needed in planning for capacity in airport terminals, airport runways, forecasts of government revenues from air travel taxes, and forecasts of total emissions from aviation on climate and atmospheric ozone¹⁰. Obviously, in planning and financing of airport infrastructure it is desirable that one has forecasts of passenger numbers for each airport which are disaggregated by the passenger destination and type¹¹. However, one can always distribute the total numbers between airports after the total forecasts are calculated¹². The distribution would be even more precise if one disaggregates total number of passengers into domestic, transborder and international passenger numbers, as the sector passenger share of each airport in Canada has been fairly constant over the years (see Table A.3 in appendix A).

In this paper we consider the aggregate passenger demand at all Canadian airports for each of the three sectors. The disaggregation also makes sense since it is expected that air passenger demand for each sector evolves differently over time¹³. The aggregated passenger numbers are calculated as:

⁸ Price elasticities tend to differ between city-pairs depending on population, GDP, level of competition, etc.

⁹ For example price of air travel is not widely available, as airlines keep their pricing information secure.

¹⁰ While emissions from aviation depends on the number of passengers, it also depends on the type of aircrafts being used and their efficiency in fuel use.

¹¹ Types of air travelers include: business, economy, discount. Further disaggregation of type of air travel could include family related travel, tourism, education related travel, etc.

¹² Indeed, airport based forecasts are of limited use, since as Nigel (2002) points out “existing route traffic will be heavily influenced by the current pattern of service and is of limited use when forecasting many years into the future.”

¹³ For example while the domestic market in the 1990’s had entered in a state of maturation, demand for international sector had high growth rates.

$$Aggregate\ pax = \sum_{a=1}^A \sum_{t=1}^T \sum_{o=1}^O pax_{a,t,o}$$

where A denotes the number of airports, T accounts for the types of passengers, and O are the origins of passengers. Forecasts of total passenger numbers in Canada are produced by several institutions. For example Transport Canada’s (TC) Air Forecasting department¹⁴ publishes yearly passenger forecasts for each of the Canadian city-pairs by origin and destination. These numbers are further disaggregated by type of passenger for a horizon of fifteen years. TC also includes different ‘scenarios’ to accommodate uncertainty in these forecasts. The last column in Table A.6 in appendix A provides the 2003-June TC forecasts for each sector aggregated for all airports. These forecasts are used by TC, the airports authorities, and other governmental departments for planning, and policy making. We will use them as a reference to compare with our forecasts.

3 Models of Seasonality

In this section, the time series models of seasonality are reviewed. All models described below are applied to monthly Canadian E/D passenger data. We first test for unit roots in each series using the Augmented Dickey-Fuller (Dickey and Fuller, 1979; Said and Dickey, 1984) test with a maximum of 14 lags chosen by BIC criteria. Table 1 in appendix A show that we cannot reject the null hypothesis of unit root and therefore all models except filtered ARMA model are assumed to have a unit root at zero frequency and, hence, the series are detrended by first differencing. For the Filtered ARMA model we actually determine whether the series have a unit root at frequency zero and therefore we test and find the appropriate frequencies that have unit root and filter it accordingly (Table A.2 in appendix A). Models are presumably listed in order of their simplicity.

3.1 An AR(p) Linear Model with Stable Deterministic Seasonality

The first model of seasonality assumes that the seasonal pattern in the time series y_t is constant over time and can be adequately captured by monthly

¹⁴See <http://www.tc.gc.ca/POL/EN/airforecasting/airforecasting.htm>.

dummy variables. Dummy variables are defined by $D_{t,s} = 1$ if $s = t$ and 0 otherwise. This model, which is labelled as *AR*, is specified as follows

$$\phi(L)\Delta_1 y_t = \delta_1 D_{t-1} + \delta_2 D_{t-2} + \dots + \delta_{12} D_{t-12} + \epsilon_t, \quad (1)$$

where $\phi(L)$ is a polynomial of order p in the lag operator L with all roots outside of the unit circle, Δ_1 is the differencing operator defined as $\Delta_k = y_t - y_{t-k}$ for all integers $k \neq 0$, and $\epsilon_t \sim i.i.d. N(0, \sigma^2)$.

Before estimating (1) it is necessary to choose the lag order such that the residuals are approximately white noise. The order of the AR model can be selected by conventional methods, such as the Akaike Information Criterion (AIC) defined as $AIC = T \ln \hat{\sigma}^2 + 2k$; or the Schwarz Information Criterion (BIC) specified as $BIC = T \ln \hat{\sigma}^2 + k \ln T$, where $\hat{\sigma}^2 = \sum_{t=1}^T \hat{\epsilon}_t^2$ with $\hat{\epsilon}_t$ the residuals from the estimated $AR(p)$ model, $k = p + 1$ the number of parameters in the model and T is the sample size. The value $p \in \{0, 1, \dots, p\}$ that minimizes the AIC or BIC is selected as the appropriate lag order, subject to passing a Lagrange Multiplier (LM) test for the first-to-seventh order residual autocorrelation at the 5% significance level. The estimation period for each series is 1985:12 to 2001:02 (182 observations). The last 24 observations are held for out-of-sample forecasting performance. Results for equation (1) are presented in appendix B (equations B.1, B.2 and B.3). Diagnostic checks suggest a well fit for the three sectors. R^2 is high, errors are all normally distributed (except for the transborder sector) and there is no problem of serial correlation. Further, the RESET test does not imply any misspecification and the ARCH test suggest no problem for autoregressive conditional heteroscedasticity.

3.2 An AR(p) Linear Model with Seasonal Unit Roots (SUR)

We label this model SUR. This model reads as follows:

$$\phi(L)\Delta_{12} y_t = \mu + \epsilon_t, \quad (2)$$

where $\phi(L)$ and ϵ_t are as defined before. Equation (2) differs from equation (1) by relaxing the assumption that seasonal patterns are constant over time. In fact, in equation (2) seasonality is assumed to be stochastic and eliminated using the seasonal filter Δ_{12} . Many time series, can be characterized by seasonal patterns which evolve over time, while each season sharing a common drift μ . In our domestic series for example this would translate into domestic

traffic in season $i = 1, 2, \dots, 12$ depending on the sum of $\epsilon_{t-i}, \epsilon_{t-12-i}$ plus an annual linear trend with slope μ .

As in model (1), the only practical decision to be made concerns selecting the autoregressive order p . The results from the application of this model for each series are presented in Appendix B (equations B.4, B.5 and B.6). For the domestic and transborder series, diagnostic tests suggest a well fit. For the international series the only problem is with serial correlation. Increasing the lag order for the international series does not solve this problem, therefore we don't further try to fix it.

3.3 A SARIMA Model

The Seasonal Autoregressive moving average model (*SARIMA*) is nothing but the combination of models (1) and (2) with the special case that the only relevant moving average term is the first year error terms. The model relies on moving average model with double differencing. This model reads as follows:

$$\phi(L)\Delta_1\Delta_{12}y_t = \mu + (1 - \theta_1L)(1 - \theta_{12}L^{12})\epsilon_t \quad (3)$$

where it is assumed that $|\theta_1| < 1, |\theta_{12}| < 1$. This is the well known "airline model" of Box and Jenkins (1970), where $\phi(L) = 1$ and $\mu = 0$. Here we also follow this restriction. Specification results for the SARIMA model are presented in appendix B (equations B.7, B.8 and B.9). For all models there is a problem of serial correlation. However since we strictly adopt the Box and Jenkins model we do not try to fix this problem by adding lagged dependent variables in the model.

3.4 A Periodic Autoregressive Model (PAR)

All the models above have had some type of restriction on either the constant, the *AR* or the *MA* terms. However, the so-called periodic autoregressive (*PAR*) model allows all the parameters to vary with the season. One may want to consider periodic *ARMA* models but this is very rarely done in practice. Due to their flexibility, the *PAR* models have been used widely in the literature, and it was first introduced by Osborn (1988) and Osborn and Smith (1989).

The model for an univariate time series y_t is specified as follows:

$$y_t = \mu_s + \tau_s T_t + \phi_{1,1}y_{t-1} + \dots + \phi_{1,12}y_{t-p} + \epsilon_t, \quad (4)$$

where the disturbance ϵ_t is assumed to be standard white noise process with constant variance σ^2 . One can observe that the number of parameters can be very large for a relatively small p value (the lag value).

One can proceed estimation of PAR by simple OLS¹⁵. The first step is to determine the lag order of the PAR model. Franses and Paap (1999) suggest the usage of BIC criterion in combination with LM tests for (periodic) serial correlation. In the case of domestic, transborder and international series the preferred lag order is five, five and four respectively. Results corresponding to the PAR model for each series are presented in appendix B (equations B.10, B.11 and B.12). The diagnostic checks show a well fit. For each series the R^2 is very high. The *LM*-statistics for first-to-seventh order serial correlation have high p -values for each series.

In the estimation of PAR model, first step involves the testing of the null hypothesis that the autoregressive parameters are the same over the seasons. Rejecting the null justifies using PAR models. The hypothesis boils down testing the null $H_0 : \phi_{i,s} = \phi_i$ for $s = 1, \dots, 12$ and $i = 1, \dots, p$ where p is the lag order chosen. Here F_{per} is the statistics to look for, see Franses (1999). The low p -values reject this null for all series, which agrees that the autoregressive terms vary with each season. The next step would be to test for the presence of periodic integration. However, since we are already using the first difference of each series, we skip this step (see Franses and Paap (1999)). Next we continue testing several other restrictions on the seasonal intercepts and trend parameters. The first test is the test for the absence of quadratic trends or more strongly $H_0 : \tau_1 = \dots = \tau_{12} = 0$ (an LR statistic). The second test involves testing for the presence of common (seasonal) linear trend (an F statistic). Low p -values suggest no reason for concern. Rest of the diagnostic tests for the PAR model also show no reasons of concern for misspecification serial correlation, autoregressive conditional heteroskedasticity or non-normality of errors.

3.5 A Structural Time Series Model (STSM)

Harvey et al (1993) argues that it is possible to model a time series with a trend, seasonal, cycle, and irregular components. The model may be written as:

$$y_t = \mu_t + \gamma_t + \psi_t + v_t + \epsilon_t, \quad (5)$$

¹⁵Under normality of the error process ϵ_t and with fixed starting values, the maximum likelihood estimators (MLE) are equivalent to the OLS estimates.

where $\epsilon_t \sim NID(0, \sigma_\epsilon^2)$, μ_t is the trend, γ_t is the seasonal, ψ_t is the cycle, v_t is first-order autoregressive term and ϵ_t is the irregular component. The trend component μ_t can be generalized by including a stochastic slope, β_t , thus

$$\begin{aligned}\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, & \eta_t &\sim NID(0, \sigma_\eta^2), \\ \beta_t &= \beta_{t-1} + \zeta_t, & \zeta_t &\sim NID(0, \sigma_\zeta^2),\end{aligned}\quad (6)$$

This generalized trend is called the *local linear trend*, where the irregular, level and slope disturbances, ϵ_t, η_t , and ζ_t respectively are mutually independent. It is clearly seen that when $\sigma_\zeta^2 = 0$ the slope is fixed and the trend reduces to a random walk with drift.

The seasonal component γ_t , is generated by

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{*j,t} \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{*j,t-1} \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega_{*j,t} \end{bmatrix}\quad (7)$$

where $j = 1, \dots, [s/2]$, $t = 1, \dots, T$, s is the number of seasons, $\lambda_j = 2\pi j/s$ is the frequency, in radians, and $\omega_{j,t}$ and $\omega_{*j,t}$ are two mutually uncorrelated white noise disturbances with zero means and common variance σ_ω^2 , which is the same for all j . The seasonality component in STSM is treated as usual.

The statistical specification of the cycle is given by

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho_\psi \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}\quad (8)$$

where $t = 1, \dots, T$, ρ_ψ is a damping factor, in the range of $0 < \rho_\psi \leq 1$; λ_c is the frequency, in radians, in the range $0 < \lambda_c \leq 1$; and κ_t, κ_t^* are two mutually uncorrelated NID disturbances with zero mean and common variances σ_κ^2 . Lastly, the autocorrelation term v_t , is treated as usual and used to treat any autocorrelation present in the residuals. The statistical treatment of the STSM is based in the state space form and the Kalman filter, the ultimate reference for the STSM model is Koopman, Harvey, Doornik and Shephard (1995). All the calculations reported in this paper for the STSM model were carried out using the STAMP package of Koopman *et al* (1995). Results for this model are presented in appendix B (equations B.13, B.14 and B.15). The diagnostic statistics show well behavior of the residuals and the estimation in general.

3.6 Testing for Seasonal Unit Roots

All the models above assumed different specifications regarding the stationarity of y_t . For example, model (1) assumed that y_t has non-seasonal unit roots and consequently a first-differences filter is used as the dependent variable. In model (2) the use of the seasonal filter (twelve differences) assumes that there are twelve unit roots in y_t . Finally, model (3), which is a combination of (1) and (2), assumed the existence of thirteen unit roots in y_t . As Franses (1996, 1999) argues, there are many cases where seasonal and/or first-differences filters are not relevant. Following this advise, we test for the presence of seasonal unit roots in y_t using the approach proposed by Hylleberg et al. (1990). Under the null hypothesis, y_t has twelve unit roots given the fact that the filter $(1 - L^{12})$ contains twelve unit roots which are $1, -1, \pm i, -0.5(1 \pm \sqrt{3}i), 0.5(1 \pm \sqrt{3}i), -0.5(\sqrt{3} \pm i), 0.5(\sqrt{3} \pm i)$, where the first root corresponds to the non-seasonal frequency (or frequency zero) and the other corresponds to different seasonal frequencies. Given that, the method for testing for seasonal unit roots consists of estimating (by OLS) the following regression

$$\begin{aligned} \phi(L)y_{8t} = & \mu_t + \pi_1 y_{1t-1} + \pi_2 y_{2t-2} + \pi_3 y_{3t-1} + \pi_4 y_{3t-2} & (9) \\ & + \pi_5 y_{4t-1} + \pi_6 y_{4t-2} + \pi_7 y_{5t-1} + \pi_8 y_{5t-2} \\ & + \pi_9 y_{6t-1} + \pi_{10} y_{6t-2} + \pi_{11} y_{7t-1} + \pi_{12} y_{7t-2} + \epsilon_t \end{aligned}$$

where $y_{1t} = (1 + L)(1 + L^2)(1 + L^4 + L^8)y_t$, $y_{2t} = -(1 - L)(1 - L^2)(1 + L^4 + L^8)y_t$, $y_{3t} = -(1 - L^2)(1 + L^4 + L^8)y_t$, $y_{4t} = -(1 - L^4)(1 - \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t$, $y_{5t} = -(1 - L^4)(1 + \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t$, $y_{6t} = -(1 - L^4)(1 - L^2 + L^4)(1 - L + L^2)y_t$, $y_{7t} = -(1 - L^4)(1 - L^2 + L^4)(1 + L + L^2)y_t$, and $y_{8t} = (1 - L^{12})y_t$. The critical values for the $t_{\hat{\pi}_1}, t_{\hat{\pi}_2}, F_{\hat{\pi}_3, \hat{\pi}_4}, F_{\hat{\pi}_5, \hat{\pi}_6}, F_{\hat{\pi}_7, \hat{\pi}_8}, F_{\hat{\pi}_9, \hat{\pi}_{10}},$ and $F_{\hat{\pi}_{11}, \hat{\pi}_{12}}$ have been tabulated by Hylleberg et al. (1990).

Equation (9) was estimated for the three time series using a maximum lag of 36 to obtain non-autocorrelated residuals. Unit root test results are presented in Table 2 in the appendix A. Given the results, we apply the corresponding filters to each time series to construct a new dependent variable. Specification tests for estimating an $ARMA(p, q)$ model using these ‘‘filtered’’ series is given in appendix B. Diagnostic checks show well fit (except for ARCH in tranborder and domestic series).

4 Evaluating Forecasts

In general, the fact that a model ‘best’ fits the data does not necessarily mean that this model also renders better out-of-sample forecasts. This is the case

since the model may be a good representation of the economy but the structure of the economy may change in the future. We clearly see the structural changes in the air industry. For example the introduction of low cost carriers may structurally change travellers preferences in choosing their mode of transportation, especially in the short haul, or demand may permanently shift down due to the security concerns of passengers. Therefore, when one is evaluating out-of-sample forecasts, one has to bear in mind the events that happened in the sample chosen for forecasting evaluation, in our case 2001:02 2003:03 (the last 24 months of our sample period). This period in air transportation was especially eventful. The events of september 11, 2001 reduced demand for air travel within Canada as well as demand for travel to US and overseas from Canada. Especially the transborder and international sectors were largely effected in this period. Furthermore, it is still not clear if the negative effects of these events are permanent and wheather the events will cause a permanent shift in demand for air transportation in the long run. In supply side, almost all the airlines have reduced capacity since September 2001, and some airlines have applied for bankruptcy protection which resulted in even more reduction in the supply of seats available. It is still not clear if supply of air services will be back to their historical levels¹⁶. In addition to the changes in the structure of the air industry, there is a new factor that needs to be included in forecasting of air demand which is named as the ‘fear factor’. Many people are still reluctant to travel by air as they are still scared away due to the ‘fear factor’. Furthermore, the security concerns have also raised the cost of air travel both in time (longer waiting in lines) and dollar cost of air ticket price (taxes, security charges, etc.)

Given the background above, and the uncertainties in the future of air transportation system, forecasting of air travel are largely unpredictable and depend mainly on the assumptions of trend and effects of the recent level shifts. Therefore, one must remember that, given the model that ‘best’ fits the data, one should use judgements that would improve forecasts such as judgements based on advance warnings of events to come (for example notice of future tax changes, the second Gulf War, SARS, etc.). Hendry and Clements (2001) further discuss the areas of our understanding in forecasting improvement. Having laid out some of the things we have to consider in evaluating short and long term forecasts, we choose a set of statistical criteria for the evaluation of out of sample forecasting performance of each of the six models discussed above.

¹⁶The long term outlook of air travel will further depend on the international ‘open skies’ agreements between countries and continents.

The first forecast evaluation criteria we use is the mean squared prediction error (*MPSE*). For simplicity, we assume h 1-step ahead static forecasts $\tilde{y}_{i,t+m|t}$, where in our case $m = 1, \dots, 24$. The *MPSE* is defined by $MPSE = \frac{1}{m} \sum_{j=1}^m \left| \frac{\tilde{y}_{i,t+m|t} - y_{i,t+m|t}}{\tilde{y}_{i,t+m|t}} \right|$. Therefore, the first step in forecast evaluation is to compare the *MSPE*'s of each model for each series. Models with smaller *MPSE* mean the model is performing better in forecasting the series. Table A.4 (appendix A) provides the relative *MPSE* of each model for each sector to the unweighted average of the six models for different horizons. Numbers in the parenthesis are the ranks. We define horizon of forecasts in four sections. These are before and after the events of September 2001, the recovery period and the longer term forecasts which brings us to 2003:02 the latest date of our data series. We choose this partitioning method to observe the performance of each model before and after a structural break and to observe if models are able to adopt to longer term adjustments in air demand. We also believe that judging the models based on simple lump sum *MPSE* could be misleading given the circumstances the air industry is going through. The ranking in Table A.4 justifies our concerns. None of the six model can claim 'best' performance, as a model's prediction rank depends on the horizon and sector chosen. For example, while SUR model is performing better for the short term for the domestic sector, it performs poorly for the transborder model. Nevertheless, when one carefully analyses forecast performances, simpler models tend to perform better in the short term while more flexible, adoptable, models perform better in the longer term. Furthermore, we observe that some models do not perform good at all for some series (for example the AR model for international data) therefore we eliminate these 'worst' performance models as candidates of forecasting long term forecasts for that series. The 'worst' models for each series are provided in Table A.4 (appendix A).

The second step in forecast evaluation is to perform a simple encompassing test for each of the remaining 'candidate' models for each series. In this step we simply regress the actual out-of-sample passenger numbers to the forecasts generated by each 'candidate' model forecasts. The aim here is to further eliminate models that do not have an 'value added' explanatory power in explaining the actuals. Therefore the null hypothesis tests the significance of a given 'candidate' model forecast. The null is rejected for small p-values. The first two columns of Table A.5 (appendix A) summarises the findings of encompassing tests.

We observe that for each series we still have more than one model that 'best' explains the actual series. Therefore next we apply the Diebold and

Mariano (1995) test which tests whether the models are statistically providing equally accurate forecasts. The test is applied as follows. Let $\tilde{y}_{i,t+m|t}$ and $\tilde{y}_{j,t+m|t}$ be the first and second ranking model for a series, the Diebold and Mariano test tests the null hypothesis of equal forecast accuracy for arbitrary loss function $g(e_{i,t|t})$, where $e_{i,t|t}$ is the corresponding 1 step ahead forecast error for model i , $e_{i,t|t} = y_t - \tilde{y}_{i,t|t-h}$. The loss function is defined as $d_t = g(e_{i,t|t}) - g(e_{j,t|t})$, so that the null of forecast accuracy entails an F statistic that $d_t = 0$. Given the vector of $\{d_t\}_{t=T}^{T+24}$, the Diebold-Mariano statistic for testing the null hypothesis of equal forecast accuracy is given by

$$\frac{\bar{d}}{\sqrt{\widehat{V}(\bar{d})}}, \quad (10)$$

where \bar{d} is the sample mean loss differential and $\widehat{V}(\bar{d})$ is the variance of \bar{d} . The third column in Table A.5 (appendix A), provides the p - values at 5% significance level of the null hypothesis that the first and second ranking model for each series have a model performance equal in favor of the one sided alternative that model i 's performance is better than model j 's. For all series we fail to reject the null hypothesis meaning that for each series the forecasting performances of the remaining 'candidate' models are equal. In this case we use the remaining 'best' models to first estimate and next forecast each series.

As a last step, we compare the unweighted average of our model forecasts with the forecasts of a naive model, where the naive model forecast is simply the actual of last period ($\hat{y}_{t+1naive} = y_t$). The fourth column in Table A.5 (appendix A) provides the ratio of the error by our forecasts to the forecast error of the naive forecasts, which are all below one. Convinced with the power of the remaining 'candidate' models for each series in explaining the actuals and that they perform better than naive forecasts, we proceed by estimating each 'remaining' model using the full sample (1984:01 - 2003:02) and forecasting for the long term (2015). The results of the estimates for each series are provided in Appendix C.

4.1 Forecast Comparisons

For each of our series we estimated and forecasted the 'best' models using the full sample. The 'best' models, selection criteria described above, provides us with more than one series of forecasts for each series which are used

as different ‘scenario’ forecasts. Next we discuss the different ‘scenario’ forecasts for each series.

For the domestic sector we estimated a SUR and STSM model. The SUR model forecasts are very pessimistic, while the STSM model forecasts are more optimistic. This is because the SUR model has a short memory adopting only to most recent data (2001 - 2003 which show consecutive negative growth rates), while the STSM model is more adoptable and flexible, therefore showing a positive trend for the forecast period. Having a pre-knowledge of events that occurred after our sample period 2003:02 (Second Gulf War, SARS epidemic in 2003, etc.), it is reasonable to adopt the more pessimistic forecasts for the short term (until 2004, SUR forecasts), and a more optimistic forecasts for the medium and long term (2005-2015, STSM forecasts).

For the transborder sector we estimate an AR and a STSM model. Similarly, AR model with a short memory have more pessimistic forecasts. To make an educated decision for our pool forecasts, we note that during 2003 both the US and the Canadian economies were still in a slow growth process. Moreover, during March 2003, the US was involved in the second Gulf war which further decreased demand for air travel between Canada and the US. Therefore, it is reasonable to adopt a more pessimistic scenario for the short term (until 2006, AR forecasts). For the medium and long term we assume that both economies will recover, and both demand and supply of air travel will slowly reestablish itself to its historical levels. Therefore we adopt the unweighted average of AR and STSM model forecasts.

For the international sector we estimate three models, AR, PAR and STSM. While AR portrays a very pessimistic scenario, STSM portrays a very optimistic scenario. In the short term, in the international sector, the SARS epidemic was the main reason for slowdown in demand for air travel. The sharp drop in demand for air travel for this sector forced airlines to cut back on capacity. However, the recent news of the lift of travel warnings by the World Health Organization (WHO) to the SARS affected cities is a sign that demand for international travel will recover slowly. Therefore we adopt the forecasts generated by the PAR model as they vary in between the AR model forecasts and the STSM model forecasts.

The forecasting results achieved for this exercise are provided in Table A.6 in appendix A. The second column provides the forecasts calculated by Transport Canada. We observe that the two forecasts are very close especially for the transborder and international sectors. For the domestic sector this paper estimates a slower recovery, while Transport Canada forecasts are slightly higher.

5 Concluding Remarks

In this paper we used six time series models to forecast seasonally unadjusted Canadian E/D air passenger data. The results show that forecast performance differs depending on the forecast horizon and the series estimated. We find that our forecasts are very similar to Transport Canada's forecasts for the transborder and international series while for the domestic sector our forecasts show a slower recovery than Transport Canada forecasts.

For future research, we suggest two further applications of time series techniques which would investigate the possibility of nonlinear characteristics of Canadian E/D passenger data¹⁷. If justified that air demand cycles in Canada have asymmetric characteristics, modelling the air data using nonlinear time series techniques would further improve the fit and forecasts of air data. The two methods that we propose are the nonlinear error correction model as used in Ericsson et al. (1998) and the Seasonal Smooth Transition Autoregressive model (SEASTAR) which has been used in modelling asymmetric behaviour of unemployment rates in US by Franses et al. (2000).

References

- [1] Box, G. E. P. and G. M. Jenkins (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- [2] Brons, M., E. Pels, P. Nijkamp, P. Rietveld (2002), "Price Elasticities of Demand for Passenger Air Travel: A Meta-analysis," *Journal of Air Transport Management* **8**, 165-175.
- [3] Dielbold, F.X. and R.S. Mariano, (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics* **13**, 253-263.
- [4] Durbin, J. and G. S. Watson (1951), "Testing for Serial Correlation in Least Squares Regression II," *Biometrika* **38**, 159-178.
- [5] Davidson, R and Mackinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

¹⁷Already the Air Forecasting Department of Transport Canada uses Box-Cox transformation in estimating airport specific forecasts.

- [6] Dickey, D.A. and W.A. Fuller (1979): "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* **74**, 427-431.
- [7] Dijk, V. D. (2002), "Smooth Transition Autoregressive Models-A Survey of Recent Developments," *Econometric Reviews*. **21**(1), 1-47.
- [8] Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity, with estimates of the United Kingdom Inflation," *Econometrica* **50**, 987-1007.
- [9] Ericsson, N. R., D. F. Hendry and K. M. Prestwich (1998), "The Demand for Broad Money in the United Kingdom, 1878-1993", *Scandinavian Journal of Economics* **100**, 289-324
- [10] Franses, P. H. (1996), *Periodicity and Stochastic Trends in Economic Time Series*, Oxford University Press.
- [11] Franses, P. H. (1999), *Time Series Models for Business and Economic Forecasting*, Cambridge University Press.
- [12] Franses, P. H., Bruinm De P., and V. D. Dijk (2000), "Seasonal Smooth Transition Autoregression," Econometric Institute Report 2000-06/A, Erasmus University Rotterdam.
- [13] Franses, H. P. and R. Paap (1999), "Forecasting with Periodic Autoregressive Time Series Models," Econometric Institute Report 9927/A, Erasmus University Rotterdam.
- [14] Gillen, W. D., W. G. Morrison, and C. Stewart (2002), "Air Travel Demand Elasticities: Concepts, Issues, and Measurement," Interim Draft Final Report for Finance Canada.
- [15] Harvey, A. C. (1990), *The Econometric Analysis of Time Series*, Second Edition. Hemel Hempstead: Philip Allan.
- [16] Harvey, A. C. and N. Shephard (1993), "Structural Time Series Models," in Maddala, G. S., C. R., Rao and H. D. Vinod (eds), *Handbook of Statistics, Volume 11*, Amsterdam: Elsevier Science Publishers B. V.
- [17] Hendry, D. F, and M. P. Clements (2001), "Economic Forecasting: Some Lessons from Recent Research," European Central Bank Working Paper Series No. 82.

- [18] Hylleberg, S., R. F. Engle, C. W. J. Granger, and B. S. Yoo (1990), "Seasonal Integration and Cointegration," *Journal of Econometrics* **44**, 215-238.
- [19] Jarque, C. M. and A. K. Bera (1987), "A Test for Normality of Observations and Regression Residuals," *International Statistical Review* **55**, 163-172.
- [20] Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (1995), *STAMP, Time Series Analyzer, Modeller and Predictor*. Timberlake Consultants LTD.
- [21] Ljung, G. M. and G. E. P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika* **65**, 297-303.
- [22] Nigel, D. P. S. (2002), "Long-term Route Traffic Forecasts and Flight Schedule Pattern for a Medium-sized European Airport," *Journal of Air Transport Management* **8**, 313-324.
- [23] Osborn, D. R. (1988), "Seasonality and Habit Persistence in a Life Cycle Model of Consumption," *Journal of Applied Econometrics* **3**, 255-266.
- [24] Osborn, D. R., and J. P. Smith (1989), "The Performance of Periodic Autoregressive Models in Forecasting Seasonal U.K. Consumption," *Journal of Business and Economic Statistics* **7**, 117-127.
- [25] Ramsey, J. B. (1969), "Test for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Statistical Royal Society B* **31**, 350-371.
- [26] Said, S.E. and D.A. Dickey (1984): "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika* **71**, 599-608.

6 Appendix A

In this appendix we present definitions of variables used in the paper. We are also presenting details of the sources of the data. The definitions are:

1. Enplaned (E) are passengers which embark and take off from an airport in Canada.
2. Deplaned (D) are passengers which land and disembark at an airport in Canada.
3. Domestic is defined as the traffic between an airport in Canada and an airport in Canada.
4. Transborder is defined as the traffic between an airport in the United States (including Alaska, Hawaii and Puerto Rico) and an airport in Canada.
5. International considers the traffic between an airport in a foreign country (excluding the United States) and an airport in Canada.

Data used in this paper was obtained from the following sources:

1. Data for the period 1984-1988 is from Catalogue 51-005 of Statistics Canada, Table 7 (Total All Airports: International and Domestic detail for scheduled services). The period 1984-1985 also considers charter data from confidential airport data base (Transport Canada).
2. Data for the period 1988-2001 is from:
 - (a) Catalogue 51-5005 (Statement 2), Table 2.2 (Enplaned and Deplaned passengers for major scheduled services, by sector);
 - (b) Catalogue 51-5005 (Statement 4), Table 3.2 (Enplaned and Deplaned passengers for regional and local scheduled services, by sector);
 - (c) Catalogue 51-5005 (Statement 6), Table 4.2 (Enplaned and Deplaned passengers for major charter services, by sector).
3. Data for the period 2001-2003 is from top 25 airport (93% of all traffic and adjusted to 100 percent). Confidential information.

Table A.1. ADF Unit Root Test

	Log(DOM)	Log(TRANS)	Log(INT)
<u>Series in Levels</u>			
t-statistic	-1.30	-2.12	-1.90
(p-value)	(0.88)	(0.53)	(0.65)
<u>Series in First Differences</u>			
t-statistic	-4.03	-3.84	-6.38
(p-value)	(0.001)	(0.02)	(0.00)

Figures in parantesis correspond to one-sided p-values of the t-stat value calculated using Davidson and MacKinnon (1993). The lag order with the minimum BIC was selected with a maximum of 14 lags. All samples correspond to 1985:12-2001:02 (207 observations) which is the sample used in forecast evaluation.

Table A.2. Seasonal Unit Root Tests

	Domestic $k_{AIC} = 24$	Transborder $k_{AIC} = 26$	International $k_{AIC} = 14$
$t_{\hat{\pi}_1}$	-1.59	-2.52	-3.02
$t_{\hat{\pi}_2}$	-0.79	-0.83	-0.24
$F_{\hat{\pi}_3, \hat{\pi}_4}$	2.40	0.01	2.56
$F_{\hat{\pi}_5, \hat{\pi}_6}$	0.55	4.92	6.67**
$F_{\hat{\pi}_7, \hat{\pi}_8}$	6.54**	1.03	0.29
$F_{\hat{\pi}_9, \hat{\pi}_{10}}$	3.45	6.94**	10.29***
$F_{\hat{\pi}_{11}, \hat{\pi}_{12}}$	0.39	0.39	4.44

** and *** denotes significance at 5% and 1% respectively. The filter applied to each series is: a) Domestic:

$$(1-L)(1+L)(1+L^2)(1+L+L^2)(1+\sqrt{3}+L^2)(1-\sqrt{3}L+L^2);$$

b) Transborder: $(1-L)(1+L)(1+L^2)(1+L+L^2)(1-L+L^2)(1-\sqrt{3}L+L^2);$

$$c) \text{ International: } (1-L)(1+L)(1+L^2)(1-L+L^2)(1-\sqrt{3}L+L^2).$$

Table A.3. Canadian Air Passenger Market Distribution

		1997	1999	2001
<u>Domestic Market</u>	<u>% Passengers Captured by:</u>			
		%	%	%
	Ontario	40	43	46
	British Columbia	17	15	14
	Quebec	12	14	14
	Alberta	15	12	10
	Others	17	16	15
<u>Transborder Sector</u>	<u>% Passengers Captured by:</u>			
	Ontario	49	51	50
	British Columbia	17	15	15
	Quebec	16	17	19
	Alberta	11	10	11
	Others	7	6	6
<u>International Sector</u>	<u>% Passengers Captured by:</u>			
	Ontario	44	45	47
	British Columbia	22	21	20
	Quebec	18	19	20
	Alberta	9	9	8
	Others	7	6	6

Data is based on the Billing and Settlement Plan (BSP) which is collected by the IATA. It captures all the airline tickets sold by Canadian travel agents.

Table A.4. Out-of-Sample Forecasting Performance (Static Forecasting)

		Model					
	Horizon	<i>AR</i>	<i>SUR</i>	<i>SARIMA</i>	<i>PAR</i>	<i>STSM</i>	<i>Filter</i>
Domestic	2001:09	0.03 (3)	0.03 (2)	0.06 (5)	0.04 (6)	0.03 (3)	0
	2002:02	0.06 (3)	0.05 (1)	0.09 (5)	0.15 (6)	0.06 (4)	0
	2002:08	0.03 (5)	0.02 (1)	0.04 (6)	0.03 (6)	0.03 (3)	0
	2003:02	0.02 (2)	0.06 (6)	0.06 (5)	0.04 (3)	0.01 (1)	0
Transborder	2001:09	0.02 (3)	0.02 (4)	0.08 (6)	0.02 (1)	0.02 (2)	0
	2002:02	0.09 (1)	0.09 (2)	0.14 (5)	0.23 (6)	0.12 (3)	0
	2002:08	0.04 (5)	0.03 (2)	0.06 (6)	0.04 (3)	0.02 (1)	0
	2003:02	0.05 (1)	0.10 (6)	0.13 (5)	0.04 (3)	0.06 (2)	0
International	2001:09	0.02 (4)	0.02 (2)	0.05 (6)	0.02 (3)	0.03 (5)	0
	2002:02	0.34 (6)	0.09 (3)	0.14 (5)	0.08 (1)	0.12 (4)	0
	2002:08	0.47 (6)	0.04 (3)	0.07 (5)	0.02 (1)	0.04 (4)	0
	2003:02	0.43 (6)	0.05 (3)	0.16 (5)	0.03 (2)	0.03 (1)	0

The table contains ratios of the different models to the MAPE of an un-weighted average forecast at horizon $h = 1$ in individual quarters, with ranks of the models in brackets. The Diebold-Mariano test statistic fails to reject the null hypothesis of forecast accuracy at 5% level for all series.

Table A.5. Forecast comparisons and combining Forecasts

Series	(1)	(2)	(3)	(4)
Domestic	0.97	0.00	0.94	0.20
Transborder	0.00	0.00	0.88	0.21
International	0.40	0.05	0.20 ^a 0.14 ^b 0.49 ^c	0.93

Column (1) shows the p-values of the following three null hypothesis H_0 : $\beta_2 = 0, \beta_1 = 0$ and $\beta_1 = 0$ corresponding to the following three equations: a) $ldom_{t+h} = \beta_1 \widehat{ldom}_{SUR,t+h} + \beta_2 \widehat{ldom}_{FilterARMA,t+h} + \beta_3 \widehat{ldom}_{STSM,t+h} + \epsilon_{t+h}$; b) $ltrans_{t+h} = \beta_1 \widehat{ltrans}_{AR,t+h} + \beta_2 \widehat{ltrans}_{STSM,t+h} + \epsilon_{t+h}$; c) $lint_{t+h} = \beta_1 \widehat{lint}_{FilterAR,t+h} + \beta_2 \widehat{lint}_{SUR,t+h} + \beta_3 \widehat{lint}_{PAR,t+h} + \beta_4 \widehat{lint}_{STSM,t+h} + \epsilon_{t+h}$. Column (2) shows the p-values of the following three null hypothesis H_0 : $\beta_1 = 0, \beta_2 = 0$ and $\beta_2 = 0$ corresponding to the following three equations: a) $ldom_{t+h} = \beta_1 \widehat{ldom}_{SUR,t+h} + \beta_3 \widehat{ldom}_{STSM,t+h} + \epsilon_{t+h}$; b) $ltrans_{t+h} = \beta_1 \widehat{ltrans}_{AR,t+h} + \beta_2 \widehat{ltrans}_{STSM,t+h} + \epsilon_{t+h}$; c) $lint_{t+h} = \beta_1 \widehat{lint}_{FilterAR,t+h} + \beta_2 \widehat{lint}_{SUR,t+h} + \beta_3 \widehat{lint}_{PAR,t+h} + \beta_4 \widehat{lint}_{STSM,t+h} + \epsilon_{t+h}$. Column (3) shows the p-value corresponding to the Diebold-Mariano statistic for comparing predictive accuracy of forecasts where $H_0 : E(d_t) = 0$. For the Domestic sector, we compare the STSM and the SUR models, for the Transborder sector, we compare the AR and the STSM models, and for the International sector, we compare: ^a the AR with the PAR, ^b the AR with the STSM; and ^c the PAR compared with the STSM. Column (4) shows the U-statistic obtained as the ratio of the 1-step ahead MAPE for the pool forecasts to that of a random walk forecast $\widehat{y}_{t+1naive} = y_t$, and $U = \frac{\sum_{t=1}^T (y_{t+1} - \widehat{y}_{t+1naive})^2}{\sum_{t=1}^T (y_{t+1} - y_t)^2}$.

Table A.6. Forecasting Results

Series	<i>Horizon</i>	<i>Forecasts</i>	<i>T.Canada*</i>
Domestic	2001	51,843	51,843
	2002	49,470	49,900
	2003	50,090	48,400
	2004	51,169	51,200
	2005	53,487	54,000
	2006	54,978	56,600
	2010	61,372	65,200
	2015	70,419	72,600
	<i>Horizon</i>	<i>Forecasts</i>	<i>T.Canada</i>
Transborder	2001	19,371	19,371
	2002	18,097	18,000
	2003	17,921	17,100
	2004	18,708	18,500
	2005	19,624	19,800
	2006	20,557	21,000
	2010	25,385	26,700
	2015	32,457	32,000
	<i>Horizon</i>	<i>Forecasts</i>	<i>T.Canada</i>
International	2001	13,043	13,043
	2002	12,440	12,400
	2003	12,879	11,800
	2004	13,463	12,800
	2005	14,200	13,800
	2006	14,958	14,700
	2010	18,191	18,900
	2015	22,598	22,700

* Source: Transport Canada. The values correspond to the latest unpublished forecasts (June 2003). Earlier forecasts are available at <http://www.tc.gc.ca/POL/EN/airforecasting/summary2002final.htm>.

7 Appendix B

In this appendix, we present the models estimated for each of the series for forecasting evaluation. All equations present the following statistics and diagnostic tests. The R^2 is the coefficient of determination measuring the fit of the data. The F-statistic tests the null hypothesis that all coefficients of the regression are null (exception of the intercept). The DW is the Durbin-Watson statistic testing for the presence of first-order autocorrelation, see Durbin and Watson (1951). The $\tilde{\sigma}$ is the estimated standard deviation of the regression. The AIC is the Akaike information criteria. The BIC is the Bayesian information criteria. The LJB is the statistic testing for non-normal residuals proposed by Doornik and Hansen (1994), which considers a small sample correction based on the original statistic proposed by Jarque and Bera (1987). The $LM_{AR}(p)$ is a Lagrange Multiplier statistic testing for the presence of correlation of order p in the residuals of the regression, see Harvey (1990). The F_{ARCH} is an F-statistic testing for the presence of autoregressive conditional heteroskedasticity, see Engle (1982). The RESET statistic test for misspecification in the estimated equation, see Ramsey (1969). In the case of the PAR models, two more statistics are presented. The F_{PER} is an F-statistic testing for periodicity in the coefficients of the estimated PAR models where the null hypothesis is constancy of the coefficients, see Franses (1996). The $LR_{\tau=0}$ statistic tests for the null hypothesis of constancy of the coefficients associated to the time trend. The $Q(i, j)$ is the statistic of Ljung and Box (1978) based on the first i residuals autocorrelations. The value PEV is the prediction error variance of one step ahead predictor errors. Further we included seasonal dummy variables and outlier dummy variables in some of our regressions. The seasonal dummies are defined as $D_{t,s} = 1$ if $s = t$ and 0 otherwise. An example of an outlier dummy used for all series for obvious reason is $D_{2001:09}$ representing September, 2001. Values in parenthesis below coefficient estimates are the standard errors.

Equation B.1. Estimated AR Model for the Domestic Sector

$$\begin{aligned}
 \Delta Ldom_t = & -0.2754 \Delta Ldom_{t-1} - 0.1352 \Delta Ldom_{t-3} + 0.1434 \Delta Ldom_{t-6} \\
 & (0.0576) \qquad\qquad\qquad (0.0591) \qquad\qquad\qquad (0.0602) \\
 & -0.2299 \Delta Ldom_{t-8} - 0.1506 \quad + 0.1397 D01_t \\
 & (0.0462) \qquad\qquad\qquad (0.0113) \qquad\qquad\qquad (0.0138) \\
 & + 0.1029 D02_t + 0.3029 D03_t + 0.1450 D04_t \\
 & (0.0188) \qquad\qquad\qquad (0.0201) \qquad\qquad\qquad (0.0153) \\
 & + 0.2445 D05_t + 0.2532 D06_t + 0.1306 D07_t \\
 & (0.0247) \qquad\qquad\qquad (0.0129) \qquad\qquad\qquad (0.0191) \\
 & + 0.2564 D08_t + 0.1187 D10_t + 0.1051 D12_t \\
 & (0.0138) \qquad\qquad\qquad (0.0197) \qquad\qquad\qquad (0.0221) \\
 & + 0.1962 D1988 : 01 + 0.1268 D1998 : 10 + 0.1581 D1999 : 05 \\
 & (0.0341) \qquad\qquad\qquad (0.0343) \qquad\qquad\qquad (0.0341)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.90, \tilde{\sigma} = 0.033, AIC = -3.91, \\
 BIC = -3.61, LJB = 4.57 (0.10), LM_{AR}(1 - 7) = 0.63 (0.67), \\
 F_{ARCH}(1 - 7) = 1.60 (0.14), RESET = 0.40 (0.53)
 \end{aligned}$$

Equation B.2. Estimated AR Model for the Transborder Sector

$$\begin{aligned}
 \Delta Ltrans_t = & -0.2649 \Delta Ltrans_{t-1} - 0.1091 \Delta Ltrans_{t-4} \\
 & (0.0286) \qquad\qquad\qquad (0.0457) \\
 & -0.0967 \Delta Ltrans_{t-5} - 0.097 \Delta Ltrans_{t-10} \\
 & (0.0439) \qquad\qquad\qquad (0.0453) \\
 & + 0.1754 \Delta Ltrans_{t-11} + 0.2396 \Delta Ltrans_{t-12} \\
 & (0.0406) \qquad\qquad\qquad (0.0437) \\
 & - 0.0407 \quad + 0.0983 D01_t + 0.1879 D03_t \\
 & (0.00523) \qquad\qquad\qquad (0.0172) \qquad\qquad\qquad (0.0155) \\
 & + 0.0346 D06_t + 0.1183 D07_t + 0.1624 D08_t \\
 & (0.0127) \qquad\qquad\qquad (0.0175) \qquad\qquad\qquad (0.0171) \\
 & - 0.1075 D09 + 0.0613 D12 + 0.1923 D1990 : 11 \\
 & (0.0128) \qquad\qquad\qquad (0.0127) \qquad\qquad\qquad (0.0338) \\
 & - 0.3887 D1990 : 12 \\
 & (0.0349)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.91, \tilde{\sigma} = 0.033, AIC = -3.89, \\
 BIC = -3.61, LJB = 7.03 (0.03), LM_{AR}(1 - 7) = 1.30 (0.25), \\
 F_{ARCH}(1 - 7) = 0.23 (0.97), RESET = 0.99 (0.32)
 \end{aligned}$$

Equation B.3. Estimated AR Model for the International Sector

$$\begin{aligned}
 \Delta Lint_t = & -0.5730 \Delta Lint_{t-1} - 0.3337 \Delta Lint_{t-2} - 0.3809 \Delta Lint_{t-3} \\
 & (0.052) \quad (0.0636) \quad (0.0672) \\
 & -0.4015 \Delta Lint_{t-4} - 0.3094 \Delta Lint_{t-5} - 0.2353 \Delta Lint_{t-6} \\
 & (0.07) \quad (0.0725) \quad (0.074) \\
 & -0.2656 \Delta Lint_{t-7} - 0.2194 \Delta Lint_{t-8} - 0.1591 \Delta Lint_{t-9} \\
 & (0.0722) \quad (0.0687) \quad (0.0637) \\
 & -0.1638 \Delta Lint_{t-10} + 0.3593 \Delta Lint_{t-12} + 0.1144 \Delta Lint_{t-13} \\
 & (0.0494) \quad (0.0585) \quad (0.0447) \\
 & -0.1945 + 0.2962 D01_t + 0.1222 D02_t \\
 & (0.0388) \quad (0.0621) \quad (0.066) \\
 & +0.2657 D03_t + 0.1528 D04_t + 0.1900 D05_t \\
 & (0.0633) \quad (0.0525) \quad (0.0499) \\
 & +0.2467 D06_t + 0.3596 D07_t + 0.3741 D08_t \\
 & (0.0589) \quad (0.0667) \quad (0.0639) \\
 & +0.2369 D09_t + 0.1447 D10_t + 0.1701 D12_t \\
 & (0.0509) \quad (0.0346) \quad (0.0454) \\
 & +0.3443 D1990 : 11 - 0.6399 D1990 : 12 \\
 & (0.0443) \quad (0.0475) \\
 & +0.2924 D1991 : 12 \\
 & (0.0615)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.97, \tilde{\sigma} = 0.042, AIC = -3.35, \\
 BIC = -3.15, LJB = 0.042 (0.98), LM_{AR}(1-7) = 2.14 (0.04), \\
 F_{ARCH}(1-7) = 1.3 (0.25), RESET = 0.04 (0.86)
 \end{aligned}$$

Equation B.4. Estimated SUR Model for the Domestic Sector

$$\begin{aligned}
 \Delta_{12}Ldom_t &= 0.5014 \Delta_{12}Ldom_{t-1} + 0.1556 \Delta_{12}Ldom_{t-2} \\
 &\quad (0.0681) \qquad\qquad\qquad (0.0731) \\
 &\quad + 0.1884 \Delta_{12}Ldom_{t-4} + 0.1524 D1988 : 1 \\
 &\quad\quad\quad (0.0625) \qquad\quad\quad (0.0457) \\
 &\quad - 0.1605 D2000 : 05 \\
 &\quad\quad\quad (0.0454)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] &= 207, R^2 = 0.58, \tilde{\sigma} = 0.045, AIC = -3.30, \\
 BIC &= -3.20, LJB = 2.53 (0.28), LM_{AR}(1 - 7) = 1.13 (0.34), \\
 F_{ARCH}(1 - 7) &= 3.30 (0.0025), RESET = 0.01 (0.93)
 \end{aligned}$$

Equation B.5. Estimated SUR Model for the Transborder Sector

$$\begin{aligned}
 \Delta_{12}Ltrans_t &= 0.5911 \Delta_{12}Ltrans_{t-1} + 0.2535 \Delta_{12}Ltrans_{t-2} \\
 &\quad (0.049) \qquad\qquad\qquad (0.0474) \\
 &\quad - 0.1007 \Delta_{12}Ltrans_{t-10} + 0.013 \\
 &\quad\quad\quad (0.0377) \qquad\quad\quad (0.00385) \\
 &\quad + 0.153 D1990 : 11 - 0.3585 D1990 : 12 \\
 &\quad\quad\quad (0.0371) \qquad\quad\quad (0.0378) \\
 &\quad - 0.2339 D1991 : 11 + 0.3468 D1991 : 12 \\
 &\quad\quad\quad (0.0378) \qquad\quad\quad (0.04)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] &= 207, R^2 = 0.77, \tilde{\sigma} = 0.037, AIC = -3.71, \\
 BIC &= -3.66, LJB = 1.56 (0.46), LM_{AR}(1 - 7) = 0.59 (0.76), \\
 F_{ARCH}(1 - 7) &= 0.74 (0.64), RESET = 0.47 (0.49)
 \end{aligned}$$

Equation B.6. Estimated SUR Model for the International Sector

$$\begin{aligned}
 \Delta_{12}Lint_t = & \underset{(0.0447)}{0.3851} \Delta_{12}Lint_{t-1} + \underset{(0.0406)}{0.2411} \Delta_{12}Lint_{t-2} + \underset{(0.00512)}{0.0238} \\
 & + \underset{(0.0482)}{0.3485} D1990 : 11 - \underset{(0.0502)}{0.6589} D1990 : 12 \\
 & - \underset{(0.0482)}{0.3651} D1991 : 11 + \underset{(0.0513)}{0.5853} D1991 : 12
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.72, \tilde{\sigma} = 0.047, AIC = -3.20, \\
 BIC = -3.15, LJB = 7.37 (0.025), LM_{AR}(1-7) = 5.53 (0.00), \\
 F_{ARCH}(1-7) = 1.88 (0.08), RESET = 5.70 (0.02)
 \end{aligned}$$

Equation B.7. Estimated SARIMA Model for the Domestic Sector

$$\begin{aligned}
 \Delta_{13}Ldom_t = & \underset{(0.058)}{0.81} \Delta_{13}Ldom_{t-1} - \underset{(0.07)}{0.42} MA(1) + \underset{(0.03)}{0.67} MA(12) \\
 & - \underset{(0.065)}{0.66} MA(13) + \underset{(0.057)}{0.11} D1995 : 05
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.58, \tilde{\sigma} = 0.077, AIC = -2.26, \\
 BIC = -2.18, LJB = 2.91 (0.23), LM_{AR}(1-7) = 10.97 (0.00), \\
 F_{ARCH}(1-7) = 2.34 (0.026), RESET = 3.12 (0.08)
 \end{aligned}$$

Equation B.8. Estimated SARIMA Model for the Transborder Model

$$\begin{aligned} \Delta_{13}Ltrans_t = & \underset{(0.03)}{0.94} \Delta_{13}Ltrans_{t-1} - \underset{(0.004)}{0.62} MA(1) + \underset{(0.028)}{0.70} MA(12) \\ & - \underset{(0.003)}{0.83} MA(13) + \underset{(0.005)}{0.13} D1990 : 11 - \underset{(0.005)}{0.25} D1990 : 12 \end{aligned}$$

$$\begin{aligned} T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.57, \tilde{\sigma} = 0.076, AIC = -2.26, \\ BIC = -2.18, LJB = 2.91 (0.23), LM_{AR}(1 - 7) = 10.97 (0.00), \\ F_{ARCH}(1 - 7) = 2.34 (0.03), RESET = 3.10 (0.08) \end{aligned}$$

Equation B.9. Estimated SARIMA Model for the International Sector

$$\begin{aligned} \Delta_{13}Lint_t = & \underset{(0.15)}{0.46} \Delta_{13}Lint_{t-1} - \underset{(-0.48)}{0.08} MA(1) + \underset{(0.04)}{0.79} MA(12) \\ & - \underset{(0.14)}{0.24} MA(13) + \underset{(0.1)}{0.32} D1990 : 11 - \underset{(0.1)}{0.21} D1990 : 12 \end{aligned}$$

$$\begin{aligned} T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.57, \tilde{\sigma} = 0.076, AIC = -2.26, \\ BIC = -2.18, LJB = 2.91 (0.23), LM_{AR}(1 - 7) = 10.97 (0.00), \\ F_{ARCH}(1 - 7) = 2.34 (0.03), RESET = 3.10 (0.08) \end{aligned}$$

Equation .B.10. Estimated PAR Model for the Domestic Sector

$$\begin{aligned}
\Delta Ldom_t = & + \frac{0.19}{(0.083)} D01_t - \frac{0.01}{(0.008)} D02_t + \frac{0.13}{(0.01)} D03_t + \frac{0.12}{(0.0086)} D05_t \\
& + \frac{0.16}{(0.025)} D06_t + \frac{0.07}{(0.008)} D08_t - \frac{0.14}{(0.008)} D09_t - \frac{0.15}{(0.012)} D10_t \\
& - \frac{0.01}{(0.002)} D01_t * Trend + \frac{0.01}{(0.0009)} D07_t * Trend \\
& + \frac{0.01}{(0.002)} D10_t * Trend - \frac{0.72}{(0.2)} (D01 * \Delta Ldom)_{t-1} \\
& - \frac{0.67}{(0.25)} (D03 * \Delta Ldom)_{t-1} - \frac{0.54}{(0.07)} (D04 * \Delta Ldom)_{t-1} \\
& - \frac{0.82}{(0.18)} (D06 * \Delta Ldom)_{t-1} - \frac{0.27}{(0.14)} (D07 * \Delta Ldom)_{t-1} \\
& - \frac{0.50}{(0.22)} (D10 * \Delta Ldom)_{t-1} - \frac{0.43}{(0.19)} (D08 * \Delta Ldom)_{t-1} \\
& - \frac{0.57}{(0.17)} (D12 * \Delta Ldom)_{t-1} - \frac{0.40}{(0.24)} (D01 * \Delta Ldom)_{t-2} \\
& + \frac{0.25}{(0.14)} (D03 * \Delta Ldom)_{t-2} - \frac{0.43}{(0.13)} (D04 * \Delta Ldom)_{t-3} \\
& + \frac{0.38}{(0.21)} (D12 * \Delta Ldom)_{t-3} + \frac{1.21}{(0.43)} (D01 * \Delta Ldom)_{t-4} \\
& - \frac{0.50}{(0.19)} (D10 * \Delta Ldom)_{t-4} - \frac{0.72}{(0.23)} (D10 * \Delta Ldom)_{t-5} \\
& - \frac{0.43}{(0.16)} (D11 * \Delta Ldom)_{t-5} + \frac{0.16}{(0.033)} D1999 : 05 \\
& - \frac{0.08}{(0.034)} D1987 : 12
\end{aligned}$$

$$\begin{aligned}
T[1985 : 12 - 2001 : 02] &= 207, R^2 = 0.91, \tilde{\sigma} = 0.03, AIC = -3.9, \\
BIC &= -3.4, LJB = 5.75 (0.06), LM_{AR}(1-7) = 1.003 (0.43), \\
F_{ARCH}(1-7) &= 1.85 (0.08), RESET = 1.25 (0.26), F_{PER} = 4.17 (0.00), \\
LR_{\tau=0} &= 11.06 (0.00), F = 130.7 (0.00)
\end{aligned}$$

Equation B.11. Estimated PAR Model for the Transborder Sector

$$\begin{aligned}
\Delta Ltrans_t = & 0.122 D01_t + 0.321 D03_t - 0.177 D04_t - 0.115 D06_t \\
& (0.01) \quad (0.044) \quad (0.018) \quad (0.055) \\
& + 0.114 D07_t - 0.194 D09_t - 0.203 D10_t + 0.154 D12_t \\
& (0.07) \quad (0.007) \quad (0.05) \quad (0.07) \\
& - 0.009 D03_t * Trend_t + 0.005 D04_t * Trend_t \\
& (0.002) \quad (0.0016) \\
& + 0.002 D05_t * Trend_t + 0.006 D06_t * Trend_t \\
& (0.0009) \quad (0.002) \\
& + 0.004 D08_t * Trend_t - 0.003 D11_t * Trend_t \\
& (0.002) \quad (0.0015) \\
& - 0.532 (D03 * \Delta Ltrans)_{t-1} - 0.891 (D10 * \Delta Ltrans)_{t-1} \\
& (0.2) \quad (0.24) \\
& - 0.362 (D11 * \Delta Ltrans)_{t-1} - 1.313 (D12 * \Delta Ltrans)_{t-1} \\
& (0.16) \quad (0.31) \\
& + 0.889 (D01 * \Delta Ltrans)_{t-2} - 0.646 (D03 * \Delta Ltrans)_{t-2} \\
& (0.12) \quad (0.33) \\
& - 0.260 (D03 * \Delta Ltrans)_{t-3} + 0.355 (D06 * \Delta Ltrans)_{t-3} \\
& (0.13) \quad (0.22) \\
& + 0.956 (D12 * \Delta Ltrans)_{t-3} - 0.943 (D05 * \Delta Ltrans)_{t-4} \\
& (0.3) \quad (0.13) \\
& - 0.582 (D08 * \Delta Ltrans)_{t-4} - 0.721 (D10 * \Delta Ltrans)_{t-4} \\
& (0.05) \quad (0.29) \\
& - 0.382 (D11 * \Delta Ltrans)_{t-4} + 0.726 (D12 * \Delta Ltrans)_{t-4} \\
& (0.16) \quad (0.32) \\
& - 0.360 (D05 * \Delta Ltrans)_{t-5} - 0.767 (D12 * \Delta Ltrans)_{t-5} \\
& (0.08) \quad (0.44) \\
& + 0.197 D1990 : 11 - 0.171 D1990 : 12 \\
& (0.03) \quad (0.07)
\end{aligned}$$

$$\begin{aligned}
T[1985 : 12 - 2001 : 02] &= 207, R^2 = 0.95, \tilde{\sigma} = 0.0267, AIC = -4.25, \\
BIC &= -3.69, LJB = 4.14 (0.13), LM_{AR}(1 - 7) = 0.98 (0.45), \\
F_{ARCH}(1 - 7) &= 0.79 (0.59), RESET = 0.41 (0.52), F_{PER} = 4.78 (0.00), \\
LR_{\tau=0} &= 7.68 (0.00), F = 187.7 (0.00)
\end{aligned}$$

Equation B.12. Estimated PAR Model for the International Sector

$$\begin{aligned}
\Delta Lint_t = & \quad 0.15 D01_t + \quad 0.32 D03_t - \quad 0.16 D04_t + \quad 0.28 D07_t \\
& \quad (0.06) \quad \quad (0.07) \quad \quad (0.04) \quad \quad (0.028) \\
& + 0.19 D08_t - \quad 0.64 D10_t + \quad 0.19 D12_t \\
& \quad (0.06) \quad \quad (0.098) \quad \quad (0.05) \\
& + 0.005 D04_t * Trend_t - 0.0064 D07_t * Trend_t \\
& \quad (0.002) \quad \quad \quad (0.002) \\
& + 0.0015 D10_t * Trend_t - \quad 0.01 D11 * Trend_t \\
& \quad (0.003) \quad \quad \quad (0.002) \\
& - 0.005 D12 * Trend_t - \quad 0.43 (D01 * \Delta Lint)_{t-1} \\
& \quad (0.002) \quad \quad \quad (0.05) \\
& - 0.42 (D02 * \Delta Lint)_{t-1} - \quad 0.54 (D03 * \Delta Lint)_{t-1} \\
& \quad (0.04) \quad \quad \quad (0.25) \\
& - 0.63 (D04 * \Delta Lint)_{t-1} - \quad 0.39 (D06 * \Delta Lint)_{t-1} \\
& \quad (0.2) \quad \quad \quad (0.18) \\
& - \quad 0.8 (D08 * \Delta Lint)_{t-1} - \quad 1.07 (D09 * \Delta Lint)_{t-1} \\
& \quad (0.27) \quad \quad \quad (0.23) \\
& - 1.16 (D10 * \Delta Lint)_{t-1} - \quad 0.56 (D12 * \Delta Lint)_{t-1} \\
& \quad (0.38) \quad \quad \quad (0.23) \\
& - 0.68 (D03 * \Delta Lint)_{t-2} - \quad 0.87 (D06 * \Delta Lint)_{t-2} \\
& \quad (0.23) \quad \quad \quad (0.09) \\
& - 0.77 (D08 * \Delta Lint)_{t-2} - \quad 1.06 (D10 * \Delta Lint)_{t-2} \\
& \quad (0.06) \quad \quad \quad (0.38) \\
& + 0.31 (D03 * \Delta Lint)_{t-3} - \quad 0.74 (D05_t * \Delta Lint)_{t-3} \\
& \quad (0.1) \quad \quad \quad (0.09) \\
& + \quad 0.37 (D08 * \Delta Lint)_{t-3} + - \quad 0.87 (D01 * \Delta Lint)_{t-4} \\
& \quad (0.17) \quad \quad \quad (0.25) \\
& + - \quad 0.66 (D11 * \Delta Lint)_{t-4} + \quad 0.37 D1990 : 11 \\
& \quad (0.1) \quad \quad \quad (0.04) \\
& - \quad 0.62 D1990 : 12 - \quad 0.14 D1999 : 12 \\
& \quad (0.1) \quad \quad \quad (0.05)
\end{aligned}$$

$$\begin{aligned}
T[1985 : 12 - 2001 : 02] &= 207, R^2 = 0.97, \tilde{\sigma} = 0.041, AIC = -3.37, \\
BIC &= -2.83, LJB = 6.52 (0.04), LM_{AR}(1-7) = 1.94 (0.07), \\
F_{ARCH}(1-7) &= 0.24 (0.97), RESET = 0.62 (0.43), F_{PER} = 3.83 (0.00), \\
LR_{\tau=0} &= 12.81 (0.00), F = 24.00 (0.00)
\end{aligned}$$

Equation B.13. Estimated Filtered ARMA Model for the Domestic Sector

$$\begin{aligned}
 Filt(Ldom)_t = & \quad 1.521 Filt(Ldom)_{t-1} - 1.323 Filt(Ldom)_{t-2} \\
 & \quad (0.0709) \qquad \qquad \qquad (0.122) \\
 & + 0.336 Filt(Ldom)_{t-3} + 0.383 Filt(Ldom)_{t-4} \\
 & \quad (0.146) \qquad \qquad \qquad (0.145) \\
 & - 0.275 Filt(Ldom)_{t-5} + 0.409 Filt(Ldom)_{t-6} \\
 & \quad (0.15) \qquad \qquad \qquad (0.147) \\
 & - 0.288 Filt(Ldom)_{t-7} + 0.348 Filt(Ldom)_{t-9} \\
 & \quad (0.105) \qquad \qquad \qquad (0.0959) \\
 & - 0.446 Filt(Ldom)_{t-10} + 0.168 Filt(Ldom)_{t-11} \\
 & \quad (0.112) \qquad \qquad \qquad (0.0708) \\
 & + 0.156 D1988 : 01 - 0.162 D2000 : 05 \\
 & \quad (0.0441) \qquad \qquad \qquad (0.0435)
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.87, \tilde{\sigma} = 0.043, AIC = -3.39, \\
 BIC = -3.35, LJB = 0.94 (0.62), LM_{AR}(1 - 7) = 0.27 (0.97), \\
 F_{ARCH}(1 - 7) = 2.96 (0.006), RESET = 2.33 (0.13)
 \end{aligned}$$

Equation B.14. Estimated Filtered ARMA Model for the Transborder Sector

$$\begin{aligned}
Filt(Ltrans)_t = & - 1.291 Filt(Ltrans)_{t-1} + 0.855 Filt(Ltrans)_{t-3} \\
& (0.0304) \qquad \qquad \qquad (0.0603) \\
& + 0.252 Filt(Ltrans)_{t-4} + 0.017 - 0.346 D1990 : 12 \\
& (0.0515) \qquad \qquad \qquad (0.00321) \qquad \qquad (0.041) \\
& - 0.224 D1991 : 11 + 0.318 D1991 : 12 \\
& (0.0413) \qquad \qquad \qquad (0.0415)
\end{aligned}$$

$$\begin{aligned}
T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.99, \tilde{\sigma} = 0.04, AIC = -3.54, \\
BIC = -3.49, LJB = 5.41 (0.07), LM_{AR}(1 - 7) = 1.79 (0.09), \\
F_{ARCH}(1 - 7) = 2.42 (0.02), RESET = 3.83 (0.05)
\end{aligned}$$

Equation B.15. Estimated Filtered ARMA Model for the International Sector

$$\begin{aligned}
Filt(Lint) = & - 2.337 Filt(Lint)_{t-1} - 2.413 Filt(Lint)_{t-2} \\
& (0.0457) \qquad \qquad \qquad (0.126) \\
& - 0.619 Filt(Lint)_{t-3} + 0.929 Filt(Lint)_{t-4} \\
& (0.191) \qquad \qquad \qquad (0.185) \\
& + 1.001 Filt(Lint)_{t-5} + 0.227 Filt(Lint)_{t-6} \\
& (0.118) \qquad \qquad \qquad (0.0418) \\
& + 0.029 - 0.661 D1990 : 12 - 0.347 D1991 : 11 \\
& (0.00951) \qquad (0.0504) \qquad \qquad (0.0494) \\
& + 0.568 D1991 : 12 + 0.333 D1990 : 11 \\
& (0.052) \qquad \qquad \qquad (0.0488) \\
& - 4.181e - 005 * Trend \\
& (6.73e-005)
\end{aligned}$$

$$\begin{aligned}
T[1985 : 12 - 2001 : 02] = 207, R^2 = 0.99, \tilde{\sigma} = 0.047, AIC = -3.17, \\
BIC = -3.09, LJB = 7.45 (0.02), LM_{AR}(1 - 7) = 5.41 (0.00), \\
F_{ARCH}(1 - 7) = 1.03 (0.42), RESET = 5.06 (0.03)
\end{aligned}$$

Equation B.16. Estimated STSM Model for the Domestic Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	18.050	18.0
<i>slope</i>	0.0026	0.8
<i>Ldom</i> _{<i>t</i>-3}	-0.180	-2.8
(<i>D</i> [*])1987 : 12	-0.060	-1.8
(<i>D</i>)1988 : 01	0.160	4.6
(<i>L</i> ^{**})1999 : 05	0.160	5.7
<i>Jan</i>	0.030	-22.0
<i>Feb</i>	0.015	3.0
<i>Mar</i>	-0.0007	3.7
<i>Apr</i>	-0.0018	-0.2
<i>May</i>	0.005	-0.5
<i>Jun</i>	-0.008	1.2
<i>Jul</i>	0.012	-2.1
<i>Aug</i>	-0.041	3.4
<i>Sep</i>	0.029	-9.7
<i>Oct</i>	0.030	6.8
<i>Nov</i>	0.010	3.2

$$T[1985 : 12 - 2001 : 02] = 183, R^2 = 0.40, \tilde{\sigma} = 0.032$$

$$LJB = 7.52 (0.02), DW = 1.97, Q(12, 9) = 17.06 (0.05)$$

$$F_{ARCH} = 1.47 (0.08), PEV = 0.00102$$

notes: *D*^{*} is a regular dummy as defined before and *L*^{**} is a level dummy suggesting a change in level at the given date
is

Equation B.17. Estimated STSM Model for the Transborder Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	3.80	5.26
<i>slope</i>	0.0007	2.75
<i>Ltrans</i> _{<i>t</i>-1}	-0.18	-2.84
<i>Ltrans</i> _{<i>t</i>-2}	0.21	0.05
(<i>L</i>)1986 : 05	0.05	1.65
(<i>D</i>)1986 : 08	0.05	1.63
(<i>D</i>)1988 : 02	0.09	3.62
(<i>L</i>)1990 : 11	0.206	7.66
(<i>D</i>)1990 : 12	-0.34	-12.44
(<i>D</i>)1995 : 05	0.05	3.60
(<i>D</i>)1998 : 10	0.09	3.37
<i>Jan</i>	0.003	0.32
<i>Feb</i>	0.06	7.60
<i>Mar</i>	0.05	7.26
<i>Apr</i>	-0.32	-4.15
<i>May</i>	-0.01	-1.34
<i>Jun</i>	0.04	6.13
<i>Jul</i>	-0.01	-1.12
<i>Aug</i>	-0.01	-1.14
<i>Sep</i>	-0.05	-6.35
<i>Oct</i>	0.02	3.5
<i>Nov</i>	-0.006	-1.0

$T[1985 : 12 - 2001 : 02] = 183$, $R^2 = 0.68$, $\tilde{\sigma} = 0.028$
 $LJB = 0.37$ (0.83), $DW = 1.91$, $Q(12, 9) = 12.65$ (0.21)
 $F_{ARCH} = 1.07$ (0.40), $PEV = 0.00079$

Equation B.18. Estimated STSM Model for the International Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	9.843	9.2
<i>slope</i>	0.004	3.7
<i>Lint</i> _{<i>t</i>-1}	0.200	3.5
<i>Lint</i> _{<i>t</i>-3}	0.103	1.8
<i>AR</i> (1)	-0.013	
(L)1986 : 08	0.031	0.017
(L)1990 : 08	-0.05	-2.4
(D)1990 : 11	0.361	8.9
(D)1990 : 12	-0.551	-13.7
(D)1991 : 02	-0.182	-4.0
(L)1991 : 01	-0.043	-1.8
<i>Jan</i>	-0.046	-3.3
<i>Feb</i>	0.110	9.5
<i>Mar</i>	0.176	10.0
<i>Apr</i>	-0.043	-3.3.3
<i>May</i>	-0.033	-2.52
<i>Jun</i>	-0.012	-1.04
<i>Jul</i>	-0.054	-4.52
<i>Aug</i>	0.007	0.63
<i>Sep</i>	-0.033	-3.31
<i>Oct</i>	0.035	2.73
<i>Nov</i>	-0.003	-0.34

$$T[1985 : 12 - 2001 : 02] = 183, R^2 = 0.78, \tilde{\sigma} = 0.041$$

$$LJB = 0.95 (0.62), DW = 1.93, Q(12, 7) = 29.54 (0.001)$$

$$F_{ARCH} = 0.88 (0.68), PEV = 0.00167$$

8 Appendix C

In this appendix, the final models for each of the series for long term forecasting are presented. As in the estimations presented in the Appendix B, each equation presents a set of diagnostic statistics. See Appendix B for brief information about these statistics.

Equation C.1. Estimated SUR Model for the Domestic Sector

$$\begin{aligned}
 \Delta_{12}Ldom_t = & \underset{(0.0552)}{0.5623} \Delta_{12}Ldom_{t-1} + \underset{(0.0576)}{0.1485} \Delta_{12}Ldom_{t-2} \\
 & + \underset{(0.0492)}{0.1567} \Delta_{12}Ldom_{t-4} + \underset{(0.0407)}{0.1561} D1988 : 01 \\
 & + \underset{(0.0405)}{0.1245} D1998 : 10 + \underset{(0.0408)}{0.1373} D1999 : 05 \\
 & + \underset{(0.0407)}{0.2070} D2002 : 09 - \underset{(0.0406)}{0.1633} D2000 : 05 \\
 & - \underset{(0.0406)}{0.2398} D2001 : 09 - \underset{(0.041)}{0.1331} D1989 : 01
 \end{aligned}$$

$$\begin{aligned}
 T[1985 : 06 - 2003 : 03] = 214, R^2 = 0.70, \tilde{\sigma} = 0.0405, AIC = -3.53, \\
 BIC = -3.47, LJB = 2.85 (0.24), LM_{AR}(1 - 7) = 1.16 (0.33), \\
 F_{ARCH}(1 - 7) = 1.85 (0.08), RESET = 0.02 (0.89)
 \end{aligned}$$

Equation C.2. Estimated STSM Model for the Domestic Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	15.22	968.00
<i>slope</i>	0.0023	2.27
(<i>D</i>)1987 : 12	-0.08	-2.07
(<i>L</i>)1988 : 01	0.12	3.84
(<i>D</i>)1988 : 03	0.08	3.35
(<i>L</i>)1990 : 12	-0.11	-4.23
(<i>D</i>)1999 : 05	0.16	5.86
(<i>L</i>)2001 : 09	-0.18	-6.73
<i>Jan</i>	-0.13	-21.23
<i>Feb</i>	0.05	7.94
<i>Mar</i>	0.02	3.12
<i>Apr</i>	-0.002	-0.41
<i>May</i>	-0.007	-1.56
<i>Jun</i>	0.006	1.41
<i>Jul</i>	-0.003	-0.65
<i>Aug</i>	0.001	0.28
<i>Sep</i>	-0.033	-6.76
<i>Oct</i>	0.041	8.09
<i>Nov</i>	0.013	3.40

$T[1984 : 01 - 2003 : 02] = 230, (n = 217), R^2 = 0.47, \tilde{\sigma} = 0.0306$

$LJB = 12.67 (0.002), DW = 1.86, Q(14, 11) = 18.13 (0.08)$

$F_{ARCH} = 1.54 (0.05), PEV = 0.00094$

Equation C.3. Estimated AR Model for the Transborder Sector

$$\begin{aligned}
 \Delta Ltrans_t = & - 0.2762 \Delta Ltrans_{t-1} - 0.0806 \Delta Ltrans_{t-2} \\
 & \quad (0.0359) \quad \quad \quad (0.0436) \\
 & - 0.1595 \Delta Ltrans_{t-3} - 0.1439 \Delta Ltrans_{t-4} \\
 & \quad (0.0346) \quad \quad \quad (0.046) \\
 & - 0.0730 \Delta Ltrans_{t-5} - 0.1279 \Delta Ltrans_{t-6} \\
 & \quad (0.0379) \quad \quad \quad (0.048) \\
 & - 0.1290 \Delta Ltrans_{t-7} - 0.2212 \Delta Ltrans_{t-8} \\
 & \quad (0.0401) \quad \quad \quad (0.0432) \\
 & - 0.1683 \Delta Ltrans_{t-9} - 0.1649 \Delta Ltrans_{t-10} \\
 & \quad (0.0339) \quad \quad \quad (0.0408) \\
 & + 0.054 D01_t + 0.1798 D03_t \\
 & \quad (0.0156) \quad \quad \quad (0.0128) \\
 & - 0.0553 D04_t - 0.0722 D05_t + 0.0862 D07_t \\
 & \quad (0.0139) \quad \quad \quad (0.0145) \quad \quad \quad (0.0138) \\
 & + 0.0973 D08_t - 0.1463 D09_t - 0.0305 D11_t \\
 & \quad (0.0129) \quad \quad \quad (0.0131) \quad \quad \quad (0.0157) \\
 & + 0.1991 D1990 : 11 - 0.3901 D1990 : 12 \\
 & \quad (0.0357) \quad \quad \quad (0.0355) \\
 & - 0.4463 D2001 : 09 \\
 & \quad (0.0355)
 \end{aligned}$$

$$\begin{aligned}
 T[1984 : 12 - 2003 : 03] = 219, R^2 = 0.92, \tilde{\sigma} = 0.034, AIC = -3.81, \\
 BIC = -3.69, LJB = 1.53 (0.45), LM_{AR}(1 - 7) = 0.75 (0.63), \\
 F_{ARCH}(1 - 7) = 1.11 (0.36), RESET = 0.33 (0.56)
 \end{aligned}$$

Equation C.4. Estimated STSM Model for the Transborder Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	14.202	790.03
<i>slope</i>	0.0043	2.75
(<i>D</i>)1990 : 11	-0.062	-1.81
(<i>D</i>)1990 : 12	0.164	4.62
(<i>L</i>)2001 : 09	0.167	5.72
<i>Jan</i>	-0.058	-5.45
<i>Feb</i>	0.056	5.57
<i>Mar</i>	0.069	8.39
<i>Apr</i>	-0.009	-1.18
<i>May</i>	-0.034	-4.67
<i>Jun</i>	0.035	4.80
<i>Jul</i>	-0.008	-1.24
<i>Aug</i>	-0.001	-1.49
<i>Sep</i>	-0.044	-6.70
<i>Oct</i>	0.032	5.56
<i>Nov</i>	0.004	0.72

$T[1984 : 01 - 2003 : 02] = 230$ ($n = 217$), $R^2 = 0.61$, $\tilde{\sigma} = 0.0356$

$LJB = 1.05$ (0.59), $DW = 1.85$, $Q(14, 11) = 16.76$ (0.12)

$F_{ARCH} = 1.27$ (0.22), $PEV = 0.00127$

Equation C.5. Estimated AR Model for the International Sector

$$\begin{aligned}
 \Delta LInt_t = & - 0.5531 \Delta Lint_{t-1} - 0.3128 \Delta Lint_{t-2} - 0.2459 \Delta Lint_{t-3} \\
 & (0.0362) \qquad (0.0281) \qquad (0.0287) \\
 & - 0.1797 \Delta Lint_{t-4} - 0.1082 \Delta Lint_{t-5} - 0.0901 \Delta Lint_{t-7} \\
 & (0.0378) \qquad (0.0382) \qquad (0.0371) \\
 & - 0.1033 \Delta Lint_{t-8} + 0.1968 D01_t + 0.1801 D03_t \\
 & (0.0382) \qquad (0.0185) \qquad (0.0202) \\
 & - 0.0887 D04_t + 0.1236 D06_t + 0.2771 D07_t \\
 & (0.0176) \qquad (0.0208) \qquad (0.0181) \\
 & + 0.2531 D08_t - 0.0862 D09_t - 0.2987 D10_t \\
 & (0.0209) \qquad (0.0151) \qquad (0.019) \\
 & - 0.4078 D11_t + 0.3561 D1990 : 11 - 0.6278 D1990 : 12 \\
 & (0.0181) \qquad (0.0494) \qquad (0.0505) \\
 & + 0.1571 D1998 : 01 \\
 & (0.0495)
 \end{aligned}$$

$$\begin{aligned}
 T[1984 : 12 - 2003 : 03] = 219, R^2 = 0.95, \tilde{\sigma} = 0.048, AIC = -3.15, \\
 BIC = -3.03, LJB = 1.38 (0.50), LM_{AR}(1 - 7) = 1.56 (0.15), \\
 F_{ARCH}(1 - 7) = 1.14 (0.34), RESET = 0.05 (0.83)
 \end{aligned}$$

Equation C.6. Estimated PAR Model for the International Sector

$$\begin{aligned}
 Lint_t = & + \frac{0.53}{(0.06)} D01_t - \frac{0.16}{(0.037)} D02_t + \frac{0.21}{(0.064)} D03_t - \frac{0.17}{(0.038)} D04_t \\
 & + \frac{0.77}{(0.034)} D06_t + \frac{0.34}{(0.038)} D07_t + \frac{0.15}{(0.047)} D08_t - \frac{0.23}{(0.023)} D09_t \\
 & - \frac{0.40}{(0.02)} D10_t - \frac{0.007}{(0.0018)} D01_t * Trend + \frac{0.004}{(0.0018)} D02_t * Trend + \\
 & \frac{0.005}{(0.0019)} D04_t * Trend - \frac{0.006}{(0.0017)} D06_t * Trend + \frac{0.004}{(0.0018)} D09_t * Trend \\
 & + \frac{0.015}{(0.0016)} D10_t * Trend - \frac{0.009}{(0.0014)} D11_t * Trend \\
 & - \frac{0.863}{(0.18)} D01_t * Lint_{t-1} - \frac{0.224}{(0.1)} D02_t * Lint_{t-1} \\
 & - \frac{0.552}{(0.17)} D04_t * Lint_{t-1} - \frac{0.684}{(0.25)} D05_t * Lint_{t-1} \\
 & - \frac{0.310}{(0.166)} D06_t * Lint_{t-1} - \frac{0.640}{(0.23)} D08_t * Lint_{t-1} \\
 & - \frac{0.630}{(0.27)} D09_t * Lint_{t-1} - \frac{0.546}{(0.09)} D12_t * Lint_{t-1} \\
 & - \frac{0.574}{(0.17)} D03_t * Lint_{t-2} - \frac{0.425}{(0.23)} D05_t * Lint_{t-2} \\
 & - \frac{0.482}{(0.19)} D06_t * Lint_{t-2} - \frac{0.553}{(0.08)} D12_t * Lint_{t-2} \\
 & - \frac{0.280}{(0.087)} D03_t * Lint_{t-3} - \frac{0.919}{(0.24)} D05_t * Lint_{t-3} \\
 & + \frac{0.364}{(0.15)} D08_t * Lint_{t-3} - \frac{0.312}{(0.12)} D05_t * Lint_{t-4} \\
 & - \frac{0.473}{(0.16)} D07_t * Lint_{t-4} - \frac{0.704}{(0.08)} D10_t * Lint_{t-4} \\
 & - \frac{0.511}{(0.27)} D12_t * Lint_{t-4} - \frac{0.483}{(0.13)} D10_t * Lint_{t-5} \\
 & - \frac{0.117}{(0.06)} D12_t * Lint_{t-5} + \frac{0.376}{(0.04)} D1990 : 11 - \frac{0.670}{(0.06)} D1990 : 12 \\
 & - \frac{0.434}{(0.16)} D1991 : 01 - \frac{0.093}{(0.04)} D2001 : 09
 \end{aligned}$$

$$\begin{aligned}
 T[1984 : 07 - 2003 : 02] &= 224, R^2 = 0.96, \tilde{\sigma} = 0.040, AIC = -3.42, \\
 BIC &= -2.80, LJB = 13.55 (0.001), LM_{AR}(1-7) = 1.07 (0.05), \\
 F_{ARCH}(1-7) &= 1.61 (0.13), RESET = 0.039 (0.98)
 \end{aligned}$$

Equation C.7. Estimated STSM Model for the International Sector

	<i>Coefficient</i>	<i>t - stat</i>
<i>level</i>	10.69	9.10
<i>slope</i>	0.0045	6.88
<i>Lint</i> _{<i>t</i>-1}	0.24	4.35
<i>Lint</i> _{<i>t</i>-2}	0.08	1.78
<i>Lint</i> _{<i>t</i>-4}	-0.09	-2.02
<i>AR</i> (1)	-0.01	
1990 : 11	0.33	8.73
1990 : 12	-0.61	-14.36
1991 : 01	-0.12	-3.98
2001 : 09	-0.13	-5.70
<i>Jan</i>	-0.05	-0.597
<i>Feb</i>	0.09	6.72
<i>Mar</i>	0.13	10.65
<i>Apr</i>	-0.08	-5.20
<i>May</i>	-0.04	-3.67
<i>Jun</i>	-0.01	-1.29
<i>Jul</i>	-0.06	-5.34
<i>Aug</i>	0.01	0.95
<i>Sep</i>	-0.04	-3.35
<i>Oct</i>	0.04	3.35
<i>Nov</i>	-0.01	-0.46

$T[1984 : 01 - 2003 : 02] = 230, (n = 212), R^2 = 0.74, \tilde{\sigma} = 0.04$

$LJB = 5.51 (0.06), DW = 2.02, Q(14, 9) = 48.54 (0.00)$

$F_{ARCH} = 0.67 (0.95), PEV = 0.042$