CJRS (Online)/ RCSR (en ligne) ISSN : 1925-2218
Vol. 33 (2): 45-56

A NOTE ON FORECASTING CROSS BORDER SHOPPING: THE VIRTUE OF SIMPLICITY

Nicole S. C. Jones
Tax Policy Officer (Tax Policy Branch)
Department of Finance
Ottawa, Canada
Nicole.Jones@fin.gc.ca

Xavier de Vanssay*
Department of Economics
Glendon College, York University
devanssay@glendon.yorku.ca
*Corresponding author

Abstract.
We use data on cross border shopping by Canadians in the U.S. from January 1972 to December 2007 and estimate three different regression models: (1) a model with a time-trend, deterministic seasonality and ARMA(p, q) disturbances, (2) a model with a time-trend, deterministic seasonality, and ARMA(p,q) disturbances including seasonal ARMA $(P,Q)_{12}$, and (3) a naïve model which focuses only on deterministic seasonality. We test the forecasting performance of the models over one to five years. Our results show that as the time horizon grows, and as predicted by the theory, the naïve model performs best.

Key Words: Cross border shopping, forecasting, naïve model, seasonality.

JEL Codes: C22, C52, F47.

Résumé. Une note sur la prévision des achats transfrontaliers: Le mérite de la simplicité
Nous utilisons les données sur les visites transfrontalières effectuées par les Canadiens aux États-Unis à partir de Janvier 1972 à Décembre 2007 et estimons trois modèles de régression : (1) un modèle avec une tendance chronologique, une saisonnalité déterministe, et des perturbations ARMA (p, q), (2) un modèle avec une tendance chronologique, une saisonnalité déterministe, et des perturbations ARMA (p, q) et saisonnières ARMA $(P, Q)_{12}$, et (3) un modèle naïf qui se concentre uniquement sur la saisonnalité déterministe. Nous testons les performances prédictives de ces modèles sur un horizon d'un à cinq ans. Nos résultats montrent que, tel que prédit par la théorie, le modèle naïf donne les meilleurs résultats au fur et à mesure que l'on augmente l'horizon de la prévision.

Mots clés : Achats transfrontaliers, prévisions, modèle naïf, saisonnalité.

Codes JEL : C22, C52, F47.

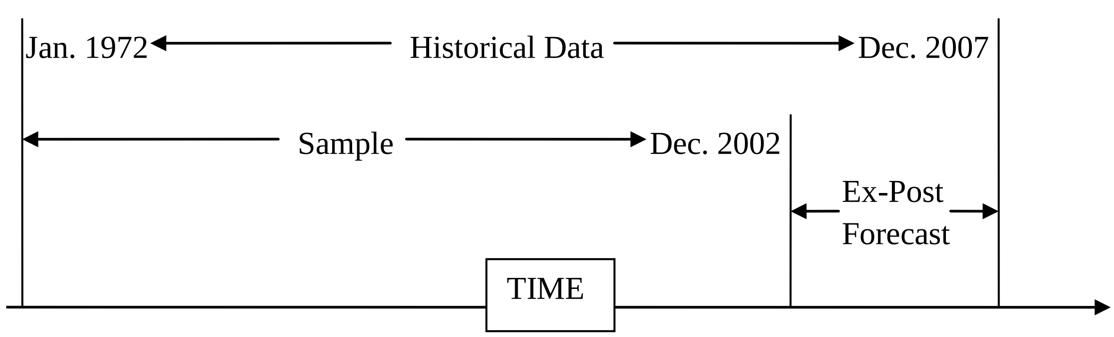
Introduction

The existence of price differences for many goods across countries is well known. This has and continues to lead people in neighbouring countries to engage in cross border shopping in order to take advantage of these price differences. Many papers and studies have been published examining the determinants and economic impact of the cross border shopping phenomenon (Di Matteo and Di Matteo, 1993; Ferris, 2000; Dmitrovic and Vida, 2007; Roy, 2007). It has received particular attention in Canada, which shares the world's longest ‘undefended’ border with the US. Much time and energy has been devoted to the subject due to its importance for both governments and retailers in Canada. When customers cross the border to shop, this represents lost tax revenue for the government. For retailers, cross border shopping represents lost business.

The cross border shopping phenomenon also falls within the broader area of travel and tourism. While much study has been devoted to forecasting tourism demand in general (Song et al, 2003; Pulina, 2003; Cho, 2003, 2009), not much work has been done in forecasting cross border shopping specifically. Due to its aforementioned importance for both governments and retailers in Canada, this appears to be a noticeable void. This paper aims to address the issue of forecasting Canada-US cross border shopping and attempts to fill some of this void. In addition, special attention is paid to the forecast time-horizon. It turns out that, in line with the theory, univariate Box-Jenkins models are not really suited for long-run forecasting, but naïve models perform rather well.

We examine the issue of modeling and forecasting cross border shopping in the Canada-US case by proposing a set of models and studying their ex-post forecasting performance over various time-horizons. To illustrate, we use data on cross border shopping by Canadians in the US from January 1972 to December 2007 and estimate models for the period of January 1972 to December 2002 (see Figure 1 for timeline). We use three different ordinary least squares (OLS) models: one with a time-trend, deterministic seasonality, and ARMA (autoregressive and moving average) disturbances, another with a time-trend, deterministic seasonality, and ARMA and seasonal ARMA (SARMA) disturbances and lastly, a naïve model focusing only on the deterministic seasonality. Then, ex-post forecasts are generated for each model for various time-horizons over the period 2003-2007, and the accuracy of the forecasting models are ranked according to Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Finally, we discuss the results of the forecasts and present our conclusions.

FIGURE 1 Timeline



CJRS (Online)/ RCSR (en ligne) ISSN : 1925-2218
Vol. 33 (2): 45-56

Estimating the Models

Cross border shopping activity is difficult to measure as there is no "shopping" classification to denote the purpose of travel. Thus, there is not any specific data on cross border shopping itself. However, in Canada two measures of cross border shopping are used by Statistics Canada. One is a count of travelers (frontier counts) which is conducted by Customs officers for Statistics Canada at border crossing points and classifies returning travelers by category (length of stay) and type of transportation (Kemp, 1992: 5.3-5.4). The second is an anonymous questionnaire (International Travel Survey of Canadian Residents) which is distributed by customs officials to returning Canadians and includes questions of expenditures. The frontier counts (e.g. same-day trips) are a more reliable measure because they are easy to collect while the surveys depend on the cooperation of both Customs officers and returning travelers (Kemp, 1992: 5.5). Although the survey is anonymous, the possibility of understatement exists, making the data less reliable.

Although same-day trips could be for business purposes or to visit friends and relatives, Statistics Canada uses the number of same-day trips (and same-day auto trips) to represent cross border shopping. This has become the barometer for cross border shopping. A same-day traveler is defined as being one who enters and leaves the country in less than 24 hours. Trips lasting one or more nights in duration are not a reliable measure because they are likely to contain vacationers as well as cross-border shoppers (Di Matteo, 1993: 52).

To measure cross border shopping, we use same-day trips data obtained from Statistics Canada and not trip expenditures due to the aforementioned unreliability of this data.[1] Figure 2 shows a graph of same-day trips by Canadians to the US from January 1972 to December 2007.

It is obvious from the graph that the series is not linear and there is a trend. To keep the model as simple as possible, we shall detrend the data using time variables (time and time^2) to capture the deterministic (non) linearity. As an added benefit, the use of a time-varying trend also renders our time-series (after removal of the trend) mean-stationary.[2] In addition, the repetitive "peak and valley" pattern in the same-day trips data suggest a strong seasonal component to the data, which are measured using monthly intercept dummy variables.[3] For example, the January dummy variable takes a value of 1 in January and 0 otherwise.

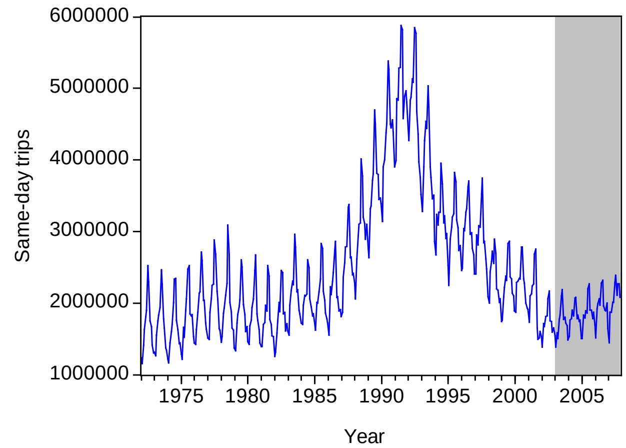
---

[1] Using annual data from 1986-2006, the correlation coefficient between same-day trips and same-day trip expenditures by Canadians in the US is about 0.46.

[2] This is why we can concentrate on ARMA disturbances rather than on ARIMA models.

[3] Ghysels and Osborn (2001: 20) note that a disadvantage of this dummy variable representation is that "it mixes seasonality and the overall mean when the latter is nonzero." For us, however, this is not an issue because we are not estimating seasonality itself but are forecasting cross border shopping.

Figure 2 Same-Day Trips by Canadians to the U.S. (monthly data), 1972-2007



Model 1: OLS with deterministic time-trend and seasonality, and ARMA disturbances

We estimate a linear model using OLS on monthly data from January 1972 to December 2002. The model uses same-day trips as the dependent variable to measure cross border shopping. For simplicity, we want to use the least amount of external data and thus, we estimate our model represented by Equation 1:

$$Daytrips_t = \sum_{i=1}^{2} \beta_i Time^i{}_t + \sum_{i=1}^{12} \gamma_i D_{it} + e_t$$ **Equation 1**

"*Time*" represents the time variables Time and Time^2, capturing the deterministic trend. D_i is a seasonal dummy variable, representing the ith month of the year (D_1 is January for example). e_t is a weakly stationary process with mean zero. Column 1 of Table 1 shows the results of the first regression.[4] The partial autocorrelation (PACF) and autocorrelation (ACF) functions of the error terms call for a correction using ARMA disturbances (making the modeling of our disturbances a univariate Box-Jenkins model, Box et al, 1994), which contain a dynamic within themselves incorporating the effect of other omitted variables, and this is important for the model, in that additional variables (and the forecast issues attached to them) are not needed.

[4] In addition, an F-test is performed to test the equality of the coefficients of the 12 dummy variables and it concludes that we reject the null hypothesis of equality and should keep the dummy variables in the model.

Table 1 Regression Results

Dependent Variable: Same-Day Trips
Monthly Data from 1972:01 to 2002:12
Method: OLS
Number of observations: 372

Variable	"OLS"	"ARMA" (Model 1)	"SARMA" (Model 2)	"Naïve" (Model 3)
TIME	19719.78 (14.145)	50600.58 (1.411)	38620.14 (2.441)	-
TIME^2	-43.267 (-11.894)	-97.5406 (-1.702)	-83.225 (-2.465)	-
January	388649.8 (2.332)	-4011275 (-0.666)	-1530500 (-0.929)	2037909 (12.05)
February	310512.1 (1.862)	-4086743 (-0.678)	-1622957 (-0.984)	1963872 (11.61)
March	754601.3 (4.520)	-3634098 (-0.603)	-1121525 (-0.681)	2411975 (14.26)
April	839124.6 (5.022)	-3549541 (-0.589)	-1040423 (-0.632)	2500425 (14.78)
May	1032799 (6.176)	-3351224 (-0.556)	-830808.6 (-0.505)	2697940 (15.95)
June	1066081 (6.370)	-3315524 (-0.550)	-806369.2 (-0.49)	2734976 (16.17)
July	1616562 (9.652)	-2761880 (-0.458)	-227389.2 (-0.138)	3289125 (19.45)
August	1533854 (9.151)	-2841659 (-0.472)	-309846.8 (-0.188)	3209998 (18.98)
September	865504.2 (5.159)	-3507179 (-0.582)	-1016580 (-0.617)	2545143 (15.05)
October	809097.4 (4.819)	-3560716 (-0.591)	-1063115 (-0.646)	2492144 (14.74)
November	597216.5 (3.555)	-3769851 (-0.626)	-1268213 (-0.77)	2283585 (13.50)
December	565657.4 (3.365)	-3798666 (-0.630)	-1291484 (-0.785)	2255260 (13.33)
AR(1)	-	0.0352 (0.262135)	0.958037 (51.208)	-
AR(2)	-	0.671 (5.346613)	-	-
AR(3)	-	0.267 (5.202492)	-	-
MA(1)	-	0.745 (5.615883)	-0.201663 (-3.471)	-
MA(2)	-	-	-0.026054 (-.447)	-
MA(3)	-	-	0.106298 (1.845)	-
SAR(12)	-	-	1.786381 (53.029)	-
SAR(24)	-	-	-0.870813 (-27.258)	-
SMA(12)	-	-	-1.806516 (-111.234)	-
SMA(24)	-	-	0.821897 (56.017)	-
Adjusted R-squared	0.483	0.982	0.986	0.125
S.E. of regression	723655.5	134037.3	119276	941643
Akaike info criterion	29.859	26.497	26.278	30.380
Schwarz criterion	30.006	26.688	26.522	30.507
Durbin-Watson stat	0.037	1.980	2.03	0.022

Note: t-statistics in parentheses.

As a guide for our model selection, we use three different Information Criteria. Specifically, we try various (parsimonious, Box et al, 1994: 329) ARMA specifications, and use the method of comparing the Akaike Information Criterion (AIC) (Akaike, 1981), Schwarz Criterion (SC) (Schwarz, 1978) and Hannan-Quinn Criterion (HQ) (Hannan and Quinn, 1979) to find the best model.[5] Table 2 shows the values of these criteria for different ARMA specifications:

Table 2 Akaike, Schwarz, and Hannan-Quinn Criteria Values for Model 1

	AR(1)	AR(2)	ARMA(2,1)	AR(3)	AR(4)	**ARMA(3,1)**	ARMA(3,2)
AIC	26.554	26.520	26.520	26.514	26.513	*26.497*	26.500
SC	26.712	26.689	26.699	26.694	26.704	*26.688*	26.702
HQ	26.617	26.587	26.591	26.585	26.588	*26.573*	26.580

Since the AIC, SC and HQ values are the lowest in the table for a model with ARMA (3,1) disturbances, this structure is chosen for our model. Thus, the ARMA (3,1) specification is added to our earlier model to correct for the positive autocorrelation problem and the regression is run again.[6] Column 2 of Table 1 presents the results of this regression.[7] When the ARMA (3,1) specification is added, the time variables and especially the 12 dummy variables become insignificant.

Model 2: OLS with deterministic time-trend and seasonality, and SARMA disturbances

For the second model, we use a model with SARMA disturbances which includes stochastic seasonal disturbances. In general, this model is denoted by SARMA(p,q) $(P,Q)_{12}$ where the orders P and Q represent the autoregressive and moving average parameters of the seasonal part of the model, while the orders p and q represent the non-seasonal portion. Different SARMA models are tried and compared using the AIC, SC and HQ information criteria. To be precise, 128 models with SARMA (p,q) $(P,Q)_{12}$ disturbances are estimated, ranging from SARMA$(0,0)(1,0)_{12}$ to SARMA$(3,3)(2,2)_{12}$. Table 3 summarizes these results, using the AIC, SC and HQ.[8]

---

[5] The Akaike, Schwarz and Hannan-Quinn measures are estimates of the out-of-sample forecast error variance that penalize for loss of degrees of freedom (due to additional independent variables). Thus, the lower these values are, the better the model is.

[6] The actual **Equation 1** becomes $Daytrips_t = \sum_{i=1}^{2} \beta_i Time^i{}_t + \sum_{s=1}^{12} \gamma_s D_{st} + Z_t$

where $\varphi(L)\, Z_t = \theta(L)\, e_t$ and e_t ~ i.i.d. $(0, \sigma^2)$ while θ(L) is the MA operator, a 1st order polynomial, and φ(L) is the AR operator, a 3rd order polynomial, with roots outside the unit circle.

[7] Once again, an F-test is performed to test the equality of the coefficients of the 12 dummy variables and it concludes that we reject the null hypothesis of equality and should keep the dummy variables in the model.

[8] The entire Table, with the 128 AIC, SC and HQ values (not presented here) is available upon request.

Table 3 Akaike, Schwarz, and Hannan-Quinn Criteria Values for Model 2: Summary

Information criterion	Minimum values	Maximum Values
AIC	26.27758	28.94726
SC	26.52163	29.10528
HQ	26.37476	29.01001
Model	**$(1,3)(2,2)_{12}$**	$(0,0)(0,1)_{12}$

Note: Number of models considered: 128, number of acceptable models: 126 (two were rejected due of the presence of singular covariance matrix, and non-unique coefficients).

Since the model with SARMA (1,3) $(2,2)_{12}$ disturbances has the lowest AIC, SC and HQ values in the table, this specification is selected for our model. Thus, our seasonal ARMA model is SARMA$(1,3)(2,2)_{12}$:

$$Daytrips_t = \sum_{i=1}^{2} \beta_i Time^i{}_t + \sum_{s=1}^{12} \gamma_s D_{st} + Z_t$$ **<u>Equation 2</u>**

where $\varphi(L)\Phi(L^s)Z_t = \theta(L)\Theta(L^s)e_t$ and $e_t \sim$ i.i.d. $(0, \sigma^2)$ while $\theta(L)$ is the non-seasonal MA operator (a 3rd order polynomial), $\varphi(L)$ is the non-seasonal AR operator (a 1st order polynomial), $\Theta(L^s)$ is the seasonal MA operator (a 2nd order polynomial), and $\Phi(L^s)$ is the seasonal AR operator (a 2nd order polynomial).

Column 3 of Table 1 shows the regression of this model. In this regression, all of the variables are statistically significant, except the MA (2) and the 12 dummy variables, which were already not significant in the ARMA (3,1) model.

Model 3: The "Naïve" model with deterministic seasonality

For the third model, we use a naïve model which captures only the basic deterministic seasonality through the use of 12 intercept dummy variables. Equation 3 represents this model.

$$Daytrips_t = \sum_{i=1}^{12} \gamma_i D_{it} + e_t$$ **<u>Equation 3</u>**

Column 4 of Table 1 shows the results of this regression. All of the dummy variables are statistically significant. An F-test to test the equality of the coefficients of the 12 dummy variables concludes that we reject this null hypothesis and should keep the dummy variables in our model.

Forecasting

In the previous section, three cross border shopping models were estimated based on monthly data for the period January 1972 to December 2002. In this section, ex-post forecasts are generated for the period 2003-2007. To fully assess the forecasting performance of the various models, one-year-ahead, two, three, four and five-years ahead forecasts will be generated for each model. The accuracy of the forecasting models will be compared using the RMSE, MAE and the MAPE. The results are presented in Table 4:

Table 4 Ex-Post Forecasting Accuracy of Models

Forecast Horizon		Model		
	Measure	**ARMA(3,1)**	**SARMA(1,3)(2,2)$_{12}$**	**Naïve**
1-year	RMSE	233374	245224	824265
	MAE	210098	208622	796564
	MAPE	12.04	11.75	45.65
2-year	RMSE	338426	329353	801710
	MAE	286428	276910.0	769197
	MAPE	17	15.76	43.169
3-year	RMSE	520546	494876	771871
	MAE	433655	406354	739027
	MAPE	24	22.85	40.826
4-year	RMSE	751962	714559	736363
	MAE	620354	581252	699313
	MAPE	34	31.65	37.928
5-year	RMSE	1001683	956073	708479
	MAE	821294	773551	662759
	MAPE	43	40.86	35.55

As is evident in the above table, the forecasting accuracy of the ARMA and SARMA models worsens the longer the time horizon in the future. This is not the case, however, for the naïve model: as the time horizon lengthens, its predicative ability increases.

For the one-year-ahead forecast, the ARMA (3,1) model is the most accurate forecasting model in terms of the RMSE, while the SARMA (1,3) $(2,2)_{12}$ is slightly better in terms of the MAE and MAPE (Figure 3 shows graphs of the one-year-ahead forecast using all three models compared to the actual data and the 95% confidence interval). The naïve model is clearly the least accurate forecasting model for the one-year-ahead forecast.

For the two-years-ahead forecast, the SARMA (1,3) $(2,2)_{12}$ model is the most accurate forecasting model according to all three measures of forecasting accuracy, followed by the ARMA (3,1) model and the naïve model is still the least accurate.

The SARMA (1,3) $(2,2)_{12}$ model still outperforms the other two models for the three-years-ahead forecast. The second-most accurate model is the ARMA (3,1) in terms of both the MAE and MAPE.

For the four-years-ahead forecast, the SARMA (1,3) $(2,2)_{12}$ model is the most accurate, however, the naïve model is now second in terms of the RMSE.

Finally for the five-years-ahead forecast, the naïve model ranks as the best with the SARMA (1,3) $(2,2)_{12}$ model in second place and the ARMA (3,1) a close third.

Figure 3a One-Year-Ahead Forecast Using ARMA(3,1) Model

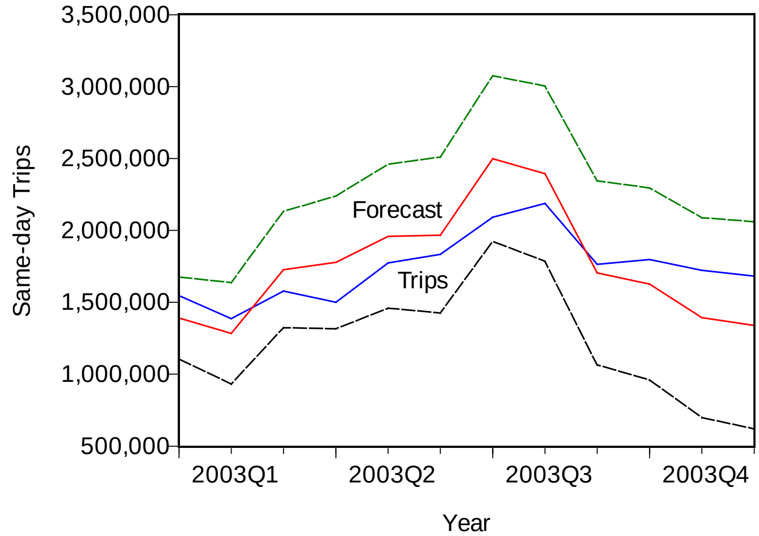
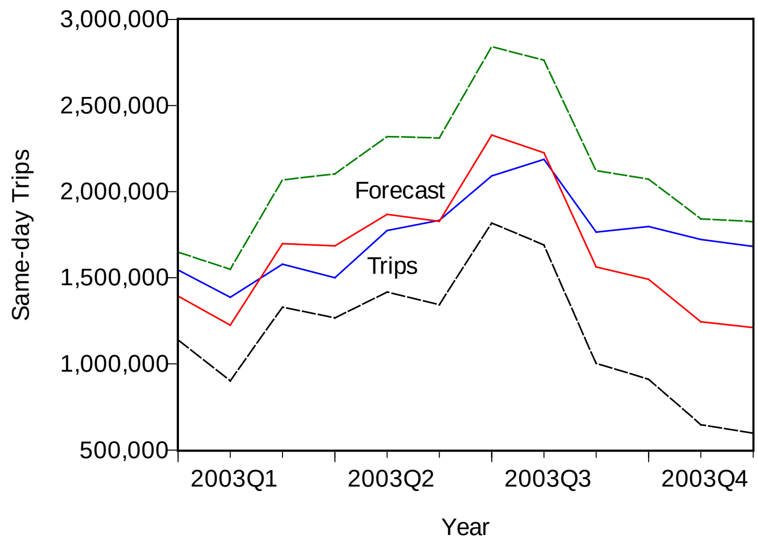
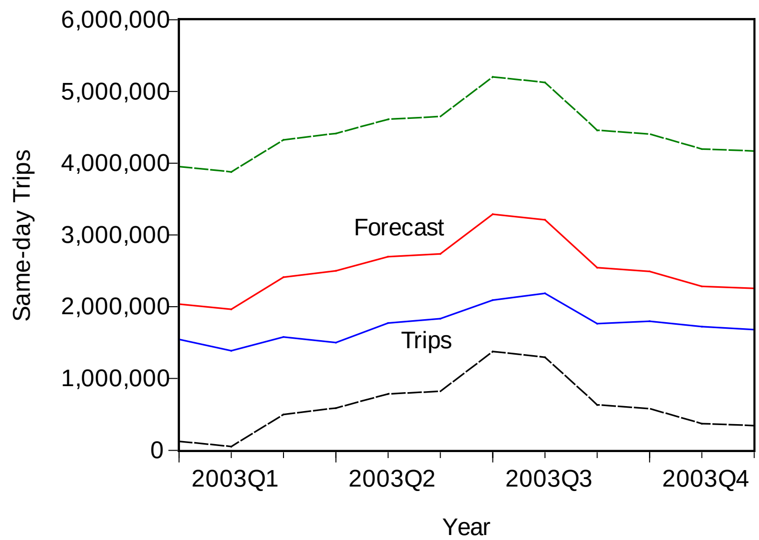


Figure 3b One-Year-Ahead Forecast Using SARMA(1,3)(2,2)$_{12}$ Model



CJRS (Online)/ RCSR (en ligne) ISSN : 1925-2218
Vol. 33 (2): 45-56

Figure 3c One-Year-Ahead Forecast Using Naïve Model



This clearly demonstrates that the forecasting capabilities of a model with deterministic trend and seasonality, and ARMA disturbances can be improved by adding a stochastic seasonality component (as recommended by Box et al, 1994: Chapter 9). Even more impressively, the forecasting accuracy of the naïve model substantially improves as the forecasting horizon increases, becoming the best model for the five-years-ahead forecast. This demonstrates that when it comes to forecasting cross border shopping, simplicity is the key. The naïve model is the simplest model of the three, yet it outperformed the other two for the five-years-ahead forecast. Lastly, the ARMA (3,1) model (which is less complex than the seasonal ARMA model) performed *reasonably* well in the short term (when compared to the sophisticated SARMA (1,3) $(2,2)_{12}$).

Conclusion

Three different OLS models have been estimated using data on Canadians cross border shopping in the US from January 1972 to December 2002: a model with ARMA (p,q) disturbances (ARMA (3,1)), a model with SARMA (p,q) $(P,Q)_{12}$ disturbances (SARMA (1,3) $(2,2)_{12}$), and a naïve model which focuses only on deterministic seasonality. Ex-post forecasts have been generated for each model for the period 2003-2007. The forecasting performance of the various models has been compared for one, two, three, four and five-years-ahead forecasts using RMSE, MAE and MAPE. The empirical results show that the ARMA (p,q) and SARMA (p,q) $(P,Q)_{12}$ models generate the most accurate one-year ahead forecasts. The SARMA (p,q) $(P,Q)_{12}$ model generates the most accurate two, three and four-years-ahead forecasts. Finally, the naïve model clearly outperforms the other two models for the five-years-ahead forecast.

The superiority of the SARMA (p,q) $(P,Q)_{12}$ model for the shorter time horizons reinforces the view that SARMA modeling works well for the short term. The deteriorating performance of the SARMA (p,q) $(P,Q)_{12}$ model as the forecasting horizon increases suggests that although including seasonal ARMA disturbances is good in theory, in practice it does not bring much benefit in terms of accurately forecasting a cross border shopping model over the long term. The improved performance of the naïve model as the forecasting horizon increases and its superiority for the five-years-ahead forecast emphasizes the importance of simplicity when forecasting cross border shopping. In the short term, an SARMA (p,q) $(P,Q)_{12}$ model will forecast cross border shopping well, but in the long term, a simple naïve model which only captures basic seasonality works the best. Thus, governments and the business community who are interested in predicting the impact of cross border shopping on their revenue might take note that when it comes to forecasting cross border shopping, simplicity rules.

References

Akaike, H. 1981. "Likelihood of a Model and Information Criteria". *Journal of Econometrics* 16: 3-14.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control. Third Edition.* Englewood Cliffs (NJ): Prentice Hall.

Cho V. 2003. "A Comparison of Three Different Approaches to Tourist Arrival Forecasting". *Tourism Management* 24: 323–330.

Cho, V. 2009. "A Study on the Temporal Dynamics of Tourism Demand in the Asia Pacific Region". *International Journal of Tourism Research* 11: 465–485.

Di Matteo, L. 1993. "Determinants of Cross-border Trips and Spending by Canadians in the United States: 1979-1991". *Canadian Business Economics* 1: 51-61.

Di Matteo L and R. Di Matteo. 1993. "The Determinants of Expenditures by Canadian Visitors to the United States". *Journal of Travel Research* 31: 34-42.

Dmitrovic, T. and I. Vida. 2007. "An Examination of Cross-Border Shopping Behavior in South-East Europe". *European Journal of Marketing* 41: 382-395.

Ferris, J. S. 2000. "The Determinants of Cross Border Shopping: Implications for Tax Revenues and Institutional Change". *National Tax Journal* 53: 801-824.

Ghysels, E. and D. R. Osborn. 2001. *The Econometric Analysis of Seasonal Time Series.* Cambridge: Cambridge University Press.

Hannan, E. J. and B. G. Quinn. 1979. "The Determination of the Order of an Autoregression". *Journal of the Royal Statistical Society,* Series B (Methodological), 41: 190–195.

Kemp, K. 1992. "Cross-Border Shopping - Trends and Measurement Issues". *Canadian Economic Observer* 5: 5.1-5.13.

Pulina, M. 2003. "Quantitative Forecasting for Tourism OLS and ARIMAX Approaches". Cagliari: Working Paper CRENoS 200303, Centre for North South Economic Research, University of Cagliari and Sassari: Sardinia.

Roy, F. 2007. "Cross-Border Shopping and the Loony: Not What It Used To Be". *Canadian Economic Observer* 20: 3.1-3.13.

Schwarz, G. 1978. "Estimating the Dimension of a Model". *The Annals of Statistics* 60: 461-464.

Song, H., Witt, S. F. and T. C. Jensen. 2003. "Tourism Forecasting: Accuracy of Alternative Econometric Models". *International Journal of Forecasting* 19: 123-141.

Statistics Canada. *International Travel.* Statistics Canada Catalogue no. 66-201-XWE. http://www.statcan.ca/bsolc/english/bsolc?catno=66-201-X (accessed July 31, 2008).

Statistics Canada. *Table 4270001 Number of International Travelers Entering or Returning to Canada, By Type of Transport, Canada Same Day, Monthly* (table), Series V125230. CANSIM (database). (accessed June 25, 2008).