Forecasting of New Car Sales in Canada

MARIA BEZYMYANNAYA

Contributors

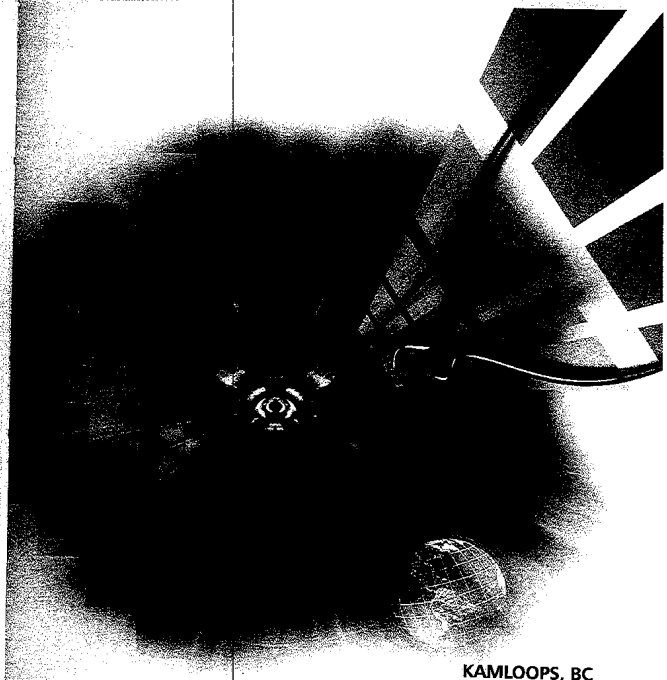
Nick Beauchesne
***Maria Bezymyannaya**
***Tara Chambers**
Alexander Dirksen
Sara Impey
John Kennedy
John McGregor
Ellen Ramsey
***Darryl Strange**

***Best Paper Award Winner**

ISSN 1915-2213
9 771915 221002 05

Printed in Canada

Proceedings of the 5th Annual
Thompson Rivers University Undergraduate Conference



KAMLOOPS, BC
April 2010

cars are substantially influenced by the consumer's confidence in the future state of the economy. Finally, new car sales by month can vary because of the seasonal nature of the industry.

The multiple regression model estimates how fluctuations in new car sales in Canada are predicted by a number of exogenous variables. The first independent variable is the state of the economy measured by the real GDP growth rate. The second variable is the price of money measured by the interest rate. The third variable is gasoline price, which is an unavoidable car expense. The fourth variable is seasonality; new car sales depend on the month in which they occur. Finally, to incorporate the impact of financial crises on the car industry, occurrence of such unexpected events is estimated by adding financial crises proxy variables to the model. Therefore, the regression model used in this study is as follows:

SALES=f(GDP, GAS, IR, WORLD, ASIA, MONTHS DUMMY)

Where:

IR - represents the overnight rate charged by the Bank of Canada

GDP - is the change in Real Gross Domestic Product

GAS - is the one-year lagged price in cents per liter of gasoline at the pump (adjusted for inflation)

WORLD - is the world financial crisis of 2008, a dummy variable equal to 1 if the crisis occurs and zero otherwise

ASIA - is the Asian financial crisis of 1997, a dummy variable equal to 1 if the crisis occurs and zero otherwise

MONTHS DUMMY - is a measure of seasonality where a dummy variable equal to 1 is assigned to each month except December.

Table 1 provides a summary of results of the regression model for new car demand. The results are obtained using ForecastX® software (Keating and Wilson, 2009). The data used in the model cover the period of January 1993 to April 2009.

Using a t-table, we are able to conclude that the coefficients are statistically significant except for October and November, which are seasonal coefficients. It can be seen from Table 1 that the explanatory power of the adjusted R^2 explains 91.72% of the variation in the sales of new cars in Canada.

Dependent Variable SALES				
	Coefficient		t-Stat	p-Value
(Constant)		197,809	33.26	0
IR		-1,778	-3.4	0
GAS		-1,277	-3.62	0
GDP		5,378	2.58	0.01
ASIA		-10,995	-2.52	0.01
WORLD		-2,47	-2.03	0.01
JAN		-17,531	-5.5	0
FEB		-16,978	-5.79	0
MAR		30,746	10.5	0
APR		37,258	12.74	0
MAY		50,928	17.19	0
JUN		37,908	12.78	0
JUL		17,756	5.96	0
AUG		17,413	5.84	0
SEP		12,445	4.19	0
OCT		3,749	1.26	0.21
NOV		-311	-0.11	0.92
R^2	Adjusted R^2		MAPE	RMSE
91.72%		90.87%	5.63%	7,923.16

Table 1: Summary of Regression Results for New Car Sales, Jan. 1993–Dec. 2009

The economic interpretation of all significant estimated coefficients out of the muliple regression model can be explored. Gross Domestic Product growth is an economic indicator that reflects the state of the economy. Positive GDP growth is associated with an expanding economy where more people are employed, higher income is generated, and more business transactions occur. When GDP growth goes down, both employment and investment decrease as people expect a turndown in the economy. If consumers are not confident in future income, employment opportunities, and the general economic situation of a country, they are not likely to purchase a new car. The regression model predicts that a 1% increase in real GDP boosts monthly new car sales by 5,378 units, holding all other variables constant.

Introduction

World news headlines such as *A Car-Sales Indicator Suggests a Recession is Near* or *New Car Sales Continue to Dip* attract the attention of economists, government officials, consumers, and people employed in the auto industry. The performance of the automotive industry is closely watched by people involved in the industry, government officials, etc., because new car sales are sensitive to fluctuations in the overall state of an economy, and such fluctuations reflect consumer response and, more importantly, consumer confidence.

In 2009, the Canadian Automobile Dealers Association reported that there are 3,000 new car dealerships across Canada employing 140,000 people (Hatch, 2010). According to Blair Qualey, president and CEO of the New Car Dealers Association of B.C., in B.C. one in seven jobs is tied to the auto industry (Gorman, 2009). Moreover, in 2009 new car dealerships generated over 43 million dollars in annual sales, which is 20.3 percent of total retail sales in Canada (Hatch, 2010). In addition, 12 percent of the total Canadian payroll is collected from new car dealerships (Williams, 2009). According to the Canadian Vehicle Manufacturers' Association, the automotive industry accounts for roughly 12% of manufacturing GDP and over 2% of total industrial GDP. Therefore, the Canadian automotive industry contributes to development of a highly skilled workforce, government revenue, and an improvement to the overall welfare of Canadians. Accurate forecasts of new car sales will provide an improved outlook on the state of Canadian economy.

This paper focuses on building a model that produces accurate and reliable forecasts of new car sales in Canada. The organization of the paper is as follows. First, the research examines the application of three forecasting techniques. These include multiple-regression, time-series decomposition, and a combined model. The multiple regression analysis evaluates how changes in economic variables, such as GDP growth, interest rate, gasoline price, and occurrence of financial crisis, affect sales of new cars in Canada. Unlike multiple regression, time series decomposition is based on available observations and forecasts future values by breaking the series into possible cyclical, trend, seasonal, and irregular components. Second, multiple regression is combined with the time-series decomposition to test any possible improvement in the overall forecast accuracy. Third, the most accurate and reliable model is chosen to forecast new car sales in Canada for the twelve months of 2010. The resulting model can be considered as an alternative method for timely and cost-effective forecasting of new car sales in Canada.

Data Set

This research utilizes New Motor Vehicle Sales Survey monthly data from January 1993 to April 2009. The data are collected directly from Canadian automobile manufacturers and importers of vehicles produced overseas through completed surveys mailed out by the head office of Statistics Canada. The survey includes monthly data on retail sales (in units) of new motor vehicles in Canada. The response rate to the survey was 100% because responding to the survey was mandatory. The data are not adjusted for seasonality, and include the sales of both trucks and passenger cars in Canada.

Multiple Regression Model

A multiple regression is used to examine how a number of independent variables interact to determine the value of a dependent variable. The regression model estimates the degree to which each of the independent variables influences the behaviour of the dependent variable.

The number of new vehicles sold in Canada is considered to be the dependent variable. New car sales account for a significant part of all retail sales and reflect the spending decisions of households and businesses. In order to understand what causes fluctuations in the automotive market, it is necessary to determine what factors affect consumer purchases of new cars. When making a "buy" decision, a typical consumer looks not only at the price of a car, but also at the costs of operating it. The consumer estimates associated car financing costs and average car expenses. In addition, sales of new

The sales of new cars are sensitive to interest rates. A new motor vehicle is a large-scale purchase which is typically acquired by taking a loan. An increase in interest rate depresses demand for new motor vehicles by raising the cost of borrowing money. Moreover, high interest rates increase the opportunity cost of holding inventory by dealerships. Conversely, lower interest rates make credit more available for consumers in all income brackets. Lower interest rates also contribute to inventory accumulation by new car dealerships. Therefore, it is expected that a decrease in interest rate would stimulate the sales of new cars. In a model that predicts the sales of new cars in Canada, we might like to have a measure of the prime interest rate. However, a more readily available series, the overnight interest rate, may be a reasonable proxy for what we want to measure, since all interest rates tend to be closely related to each other. Figure 1 illustrates the interest rates for the period January 1993 to April 2009.

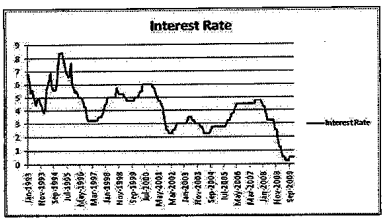


Figure 1. Bank Rate in Canada, Jan. 1993–Dec. 2009 (Source: Statistics Canada)

The overnight rate since 1993 tends to decrease. Such a pattern suggests that consumers are more likely to make large scale purchases such as cars because loans are more easily available. Higher interest rates have a negative effect on consumer confidence; lower rates have the reverse effect. In fact, the estimated model predicts a 1% increase in interest rate causes a reduction of 1,778 car sales in each month.

The price of gasoline is assumed to be another determinant of demand for new cars because gasoline is a complement to a car. A complement is a commodity which is regularly used in conjunction with some other commodity. The decision to purchase a new motor vehicle involves a long-run decision-making process. Rational consumers do not rush through this process; they estimate the future costs of operating a car. The past trend in gasoline prices influences consumers' expectations about future prices. The consumers' decision whether or not to buy a vehicle is based on their expectations about future costs they will have to budget. In his 2008 research on the link between gasoline prices and vehicle sales, Walter McManus concludes that consumers are not responsive to changes in the price of gasoline in the short run, but that a sustained increase of 10 percent in price eventually would reduce new car sales by about 4 percent (McManus, 2008). In order to estimate the influence of past gasoline prices on current demand for new car sales, a one year lagged gasoline variable is used in the model. The estimated coefficient for a gasoline price variable shows that, on average, a 1% increase in the price of gasoline reduces next year's monthly sales by 1,277 units of cars.

Sales of many commodities are subject to seasonality. Sales of new motor vehicles start to rise during spring and summer months. The sales throughout the year are typically lower in winter. This pattern is reasonably consistent, although there is variability in the degree of seasonality and some deviation from the overall pattern. To account for and measure seasonality in a regression model, we will use 11 dummy variables. In order to prevent the so-called dummy trap, the month of December is considered as the reference month for the regression model. The seasonal dummy variables are expected to have signs representing their relationships to the omitted month (i.e. December). The coefficients of dummy variables show how much the sales of new cars decrease or increase because of the effect of the season, given that all other factors are constant.

The estimated seasonal coefficients are statistically significant except for October and November. Figure 2 shows by how many units

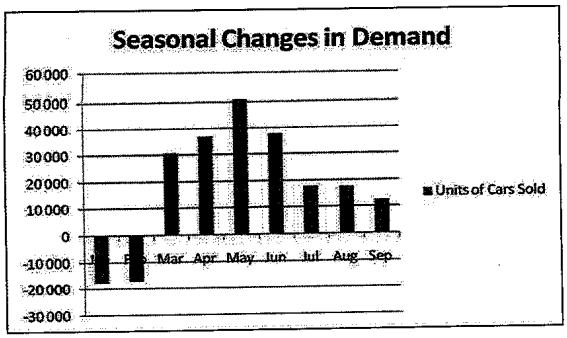


Figure 2. Seasonal Changes in Demand for New Cars in Canada (Source: ForecastX)

of cars sales decrease/increase due to the effect of a month.

If an unexpected event influences the overall state of an economy, either in a positive or negative way, it will immediately be reflected in the sales of new motor vehicles. The Asian Financial Crisis was a crisis that gripped much of Asia beginning in July 1997, and raised fears of a worldwide economic meltdown due to financial contagion (Battelle, 2001). The World Financial Crisis of 2008 is also one of the worst financial crises since the Great Depression in the 1930s. It caused a strong decline in economic activities, wealth of consumers, and serious financial losses by businesses and governments. As indicated in the article *The 2007-08 Financial Crisis in Review*, "with the onset of the global credit crunch and the fall of Northern Rock, August 2007 turned out to be just the starting point for big financial landslides" (Singh, 2009). To incorporate the effect of the two major financial crises into the model, two qualitative variables are defined, one for the Asian crisis, and one for the World crisis of 2008. The model estimates that during the occurrence of the Asian Crisis, monthly sales decreased by 10,995 units of cars. Similarly, the effect of the Financial Crisis of 2008 is responsible for an estimated decrease of 2,470 units in each month of its occurrence.

Time-Series Decomposition Model

Utilizing alternative forecasting techniques is a common practice in real world examples. Time-series decomposition is applied to forecast the sales of new motor vehicles in Canada. The technique assumes that some aspects of the past pattern of the data will remain in the future. The model is based on available observations from a time series and forecasts future values by breaking the series into cyclical, trend, and seasonal components, assuming there is no irregularity pattern in the data. Although the model acts like a black box because it tells us only 'what' will happen but not 'why' it happens, it has been widely used in reality. This model is expected to provide additional information about the patterns in the underlying data. The plot of new car sales depicted in Figure 3 shows the volatility in the data.

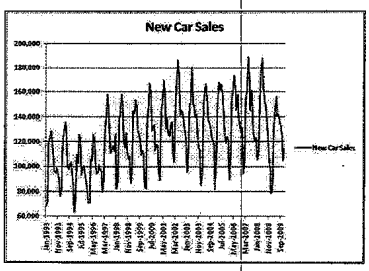


Figure 3: New Car Sales Measured in Units of Cars

There is also a successive pattern in the movement of the sales of new cars in Canada, with sharp increases and decreases following each other. This pattern suggests that there is seasonality in the data. In addition, the data appear to have some long-term wavelike movement and probably a slight positive trend. To better understand these underlying patterns, it is necessary to isolate and examine them individually.

Month	Seasonal Index
January	0.70
February	0.73
March	1.13
April	1.15
May	1.28
June	1.21
July	1.02
August	1.01
September	1.01
October	0.93
November	0.90

Table 2: Seasonal Indices for Canada's New Car Sales

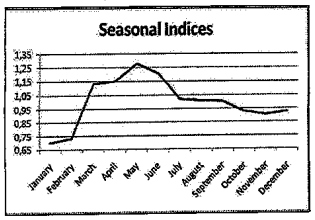


Figure 4: Seasonal Indices for New Car Sales (Graphical Representation)

The first step is to remove seasonality from the data. After applying the time series decomposition process, the following seasonal indices then can be summarized in Table 2 and Figure 4.

As shown above, the seasonal index of new car sales for January is 0.70. Compared to other monthly sales, this means that typical January sales of new cars in Canada are 30 percent below the average monthly value for the year. The seasonal indices for the months of March through September are all above unity, indicating that sales during these months are generally higher.

The estimated time-series decomposition model also indicates a long-term positive trend. A rising trend implies that the average of new car sales increases through time. This can be due to an overall increase in population, inflation, and/or general economic changes. Figure 5 depicts the series in which there is an obvious upward trend over time.

The cyclical component of a time-series is the extended wavelike movement throughout the long-term trend. The estimated cyclical factor for the sales of new cars is plotted in Figure 6.

The cyclical component for new motor vehicle sales in Canada does not have constant amplitude. The distances from peak to trough before 2001 were much longer than the recent distances from peak to trough. This means that it takes less time for economic activity to move from peak to trough; additionally, the size of both expansion

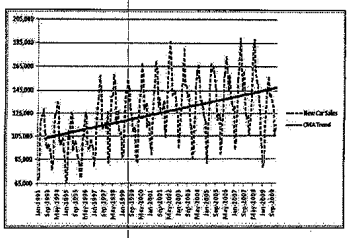


Figure 5: New Car Sales and Centred Moving-Average Trend

and contraction decreased. The accuracy measures for the time-series decomposition model are summarized in Table 3.

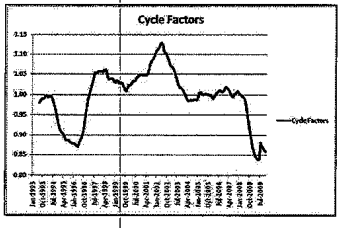


Figure 6: Cycle Factor for New Car Sales

MAPE	3.86%
R-Square	94.65%
R-Square Adjusted	94.62%
RMSE	5,876.41

Table 3: Summary of Accuracy Measures for Decomposition Forecast

Using the estimated time-series decomposition, actual and forecast values for January 1993 through April 2009 are shown in Figure 7.

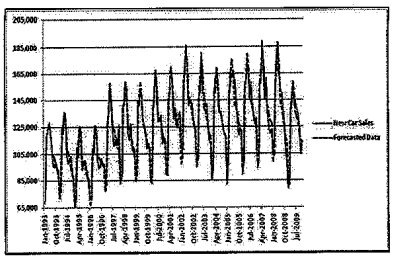


Figure 7: New Car Sales and its Time-series Decomposition Forecast from Jan. 1993–April 2009

Combination of Forecasts

When different forecasts are made of the same event, a decision maker might choose the best forecast and discard the rest. This procedure is not reasonable if the objective is to obtain the most accurate forecast. The discarded forecast may contain some useful and independent information. One forecast may be based on the variables that are not included in the other forecast. In addition, two forecasts can be based on different assumptions about the relationship between variables. Forecasts are combined to reduce likely errors and improve the predictive power of a forecast. In order to obtain an improved model to forecast the sales of new motor vehicles in Canada, the estimated multiple-regression and time-series decomposition models are combined.

The two separate forecasts of new car sales in Canada are combined on the basis of error-minimization criteria such as Root Mean Square Error (RMSE). Compared on the basis of RMSE, the decomposition model significantly outperforms the estimated multiple regression model. The RMSE of the decomposition model is almost 34 percent less than the RMSE of the multiple regression. However, Bates and Granger suggested that the combination of forecasts will do better than any of the individual forecasts (Bates & Granger, 1989). Research from over 200 studies demonstrates that combining forecasts produces consistent but modest gains in accuracy (Armstrong, 1989). Even though the time-series decomposition model shows lower RMSE, it is not logical to base the forecast solely on the decomposition model. It is likely that valuable independent information contained in the multiple-regression model may be lost. The multiple-regression model of the new car sales in Canada uses GDP growth, gasoline prices, interest rates, seasonality, and occurrence of financial crises to explain the variation in sales. It is possible, however, that some of these variables are not taken into account by the decomposition model. Additional information obtained from the multiple-regression model may improve the accuracy of the time-series decomposition. Moreover, the two models may test for different forms of the relationship. To prevent such loss of useful information, and to improve the accuracy of the forecast, a combined model is constructed. The hypothesis, yet to be tested, is that the combined forecast has a smaller error, as measured by RMSE, unless individual forecasting models are almost equally good.

To combine forecasts, one needs to assign weights to each forecast. It has been suggested that regression can be used to estimate such weights. New forecasts based on combined forecasting models using weights out of regression were found to be more accurate than combined forecasts with equal weights (Krishnamurti, 1999). We begin by regressing actual sales of new cars in Canada on multiple regression and time-series decomposition model outputs. It was determined that the intercept of the combined model is essentially equal to zero. The results are summarized in Table 4.

Conclusion

Forecasting of new motor vehicle sales is extremely important for the Canadian economy to make fundamental economic decisions, and to analyze the ups and downs in the economy. Increasing demand for cars reflects increasing income, growing purchasing power, improving of financial conditions and positive potential in the labor market. Thus, the growth in new car sales can be considered as an indicator of the overall health of the economy.

The sale of new cars contributes to the Canadian GDP, retail sales, employment, and federal and provincial revenues. It is critical to predict the consumer and producer behaviours of the new car market. New car dealers need to know how much inventory to hold. Producers adjust the production of cars for ups or downs in sales. Governments plan their budgets based on potential tax revenue collected from the automotive industry.

The purpose of this research was threefold. First was to construct an accurate forecasting model of new car sales in Canada. Second was to generate a more accurate forecast by combining two forecasting models, multiple regression and decomposition. Third was to apply the procedure to predict the sales of new cars in Canada for 2010.

The three models, multiple regression, decomposition, and a combined model, were tested for new car sales predictions in Canada. For multiple regression, the sales were regressed on six independent variables, namely real GDP growth rate (GDP), interest rate (IR), price of gasoline (GAS), Asian Financial Crisis of 1997 (ASIA), World Financial Crisis of 2008 (WORLD), and seasonality (MONTHS DUMMY). The results indicate a positive impact of GDP growth on sales of new cars. However, an increase in the price of gasoline and interest rates depresses the monthly sales of new cars. Additionally, occurrences of financial crises significantly decrease demand for new motor vehicles in Canada. In time-series decomposition, the forecast is based on available time-series observations. Time-series decomposition significantly outperforms the multiple regression model. However, the combination of the two forecasts provides the most accurate model. In order to test the accuracy of the sample

Dependent Variable: SALES			
	Coefficient	Standard Error	t-test
(Constant)	-1,911.81	1,695.06	-1.13
Multiple Regression	0.16	0.05	5.25
Decomposition	0.84	0.05	16.51

Table 4: Summary of Regression Results for Combined Model

Based on the estimated weights, the new forecasts can be expressed as follows. Given a t-value of -1.13 for the intercept, we would conclude that it is not statistically different from zero at any meaningful significance level. The next step is to redo the same regression but force it without an intercept. The new estimated model is as follows:

Sales = (0.151397 * Forecasts out of Multiple Regression Model) + (0.848819 * Forecasts out of Decomposition Model)

The new estimated model indicates that the greatest weight should be assigned to the forecast out of the decomposition model. Table 5 summarizes accuracy measures for the three models.

Method	MAPE	RMSE
Multiple Regression	5.63%	7,923
Decomposition	3.86%	5,876
Combined	3.32%	5,683

Table 5: Summary of Accuracy for 3 Tested Models

The forecast for 24 months of 2009 and 2010 is constructed. The values for sales and combined forecasts are shown in Table 6.

Date	Actual Sales	Forecast
Jan-09	78,817	85,517
Feb-09	82,039	90,772
Mar-09	129,831	136,651
Apr-09	146,026	143,958
May-09	156,025	158,563
Jun-09	140,442	148,380
Jul-09	141,621	132,968
Aug-09	136,971	130,759
Sep-09	132,595	125,826
Oct-09	123,607	118,365
Nov-09	104,926	109,515
Dec-09	114,443	111,575
Jan-10		87,037
Feb-10		92,226
Mar-10		138,246
Apr-10		145,509
May-10		160,217
Jun-10		150,168
Jul-10		145,670
Aug-10		141,411
Sep-10		135,438
Oct-10		126,198
Nov-10		110,360
Dec-10		115,615

Table 6: Actual and Forecasted Values for New Car Sales in Canada (Units of Cars)

Figure 8 shows the combined forecasts and the actual available sales values. The forecast values appear reasonable.

The forecast for 2010 indicates that new car purchases are rebounding sharply in February. The global economy and auto sector are in the early stages of a recovery which will lift new car sales in Canada above 1.65 million units in 2010, up from 1.5 million last year.

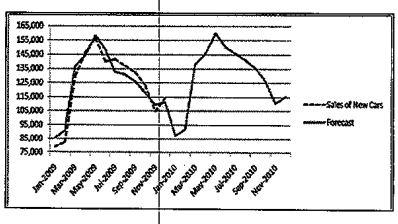


Figure 8: New Car Sales in Canada and Forecasted Sales by the Combined Model (Units of Cars)

forecast, the combined model was tested to predict the sales of new cars for 2009. The forecasted values turned out to be reasonably close to the actual values. Finally, new car sales in Canada for the 12 months of 2010 were predicted. Further research can be conducted by combining quantitative forecasting methods with qualitative and judgmental information provided by automotive industry experts.

Works Cited

"About CADA" (2010). *Canadian Automobile Dealership Association* http://www.cada.ca/cmsDisplay.asp?CONTAINER_GUID=EB8CD3E7-679B-4470-AE81-9F51C4229133&ARTICLE_GUID=&intAction=5000

Armstrong, J. (1989). "Combining Forecasts: The End of the Beginning or the Beginning of the End?" *International Journal of Forecasting*, Vol. 5, No. 4, p. 585.

Armstrong, J. (1984). "Forecasting by Extrapolation: Conclusions from Twenty-five Years of Research." Marketing Papers. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1083&context=marketing_papers

Batelle, J. (2001). "The Asian Financial Crisis." *Economic Journal*. www.hoover.org/multimedia/economic_crisis

Bates, A. and Granger, C. (1989). "Invited Review: Combining Forecasts - Twenty years Later." *Journal of Forecasting* Vol. 8, p. 167–173.

Gorman, T. (2008). "New Vehicle Sales Rise." *Nanaimo News Bulletin*. http://intranet.newcardealers.ca/Uploads/Public/NewsAndEvents/sept1809ar.pdf

Hatch, M. (2010). "Canada's Dealers and Economy at a Glance." *Annual Economic Review 2009*. http://www.cada.ca/Uploads/PDFS/public/era_EN-2010.pdf

Keating, B. and Wilson, K. (2009). Business Forecasting with ForecastX, 6/e [Software]. Available from http://www.forecastxperttoolkit.com/

Krishnamurti, T. N. *et al.* (1999). "Improved weather and seasonal climate forecasts from multimodal ensemble." *Science*, 285 (Sept. 3), 1548-1550.

McManus, W. (2008). "Effects of Gasoline Prices on Driving Behavior and Vehicle Markets." *A CBO Study*. http://www.cbo.gov/ftpdocs/88xx/doc8893/01-14-GasolinePrices.pdf

Singh, M. (2009). "The 2007-08 Financial Crisis in Review." *Investing Notes*. http://myinvestingnotes.blogspot.com/2009/12/2007-08-financial-crisis-in-review.html

Statistics Canada, "New Motor Vehicle Sales, Canada, Provinces." 5 December 2009, http://estat.statcan.gc.ca.ezproxy.tru.ca/cgi-win/CNSMCGLEXE

Statistics Canada, "Financial Market Statistics, Last Wednesday, Monthly." 4 December 2009, http://estat.statcan.gc.ca.ezproxy.tru.ca/cgi-win/CNSMCGLEXE

Statistics Canada, "Average Retail Prices for Gasoline and Fuel Oil." 4 December 2009, http://estat.statcan.gc.ca.ezproxy.tru.ca/cgi-win/CNSMCGLEXE

Statistics Canada, "Gross Domestic Product at Basic Prices." 4 December 2009, http://estat.statcan.gc.ca.ezproxy.tru.ca/cgi-win/CNSMCGLEXE

Williams, J. (2009). In *The Chief Economist*. http://chandan.com/globest/?p=167