CLK OIL

Integration Analysis of CLK Oil and DCB Oil

Prepared by CLK Finance Division

Table of Contents

Executive Summary

The Finance Division of CLK Oil has prepared the following report on expected gas prices and demand. CLK Oil is a midsize petroleum company representing 1.5% of U.S. gasoline sales. CLK Oil is evaluating a potential acquisition—DCB Oil. CLK needs to anticipate combined value of gasoline sales for the two companies from March 2017 to February 2018. Traffic volume was forecasted to predict demand, whereas gas prices at the pump were forecasted to predict profit. Several forecasting models were evaluated, and the following were selected: Multiple Regression with Benchmark for traffic volume and Decomposition Method for gas price.

For the forecasted period, this model forecasted 3.154 trillion miles driven in the U.S. Based on expected traffic volume and gas prices at the pump, CLK's future 12-month net profit from gasoline sales after the merger is expected to be $90.1 ± $17.7M, a $5.63M increase from profits without the merger. The net present value (NPV) of gasoline supply chain integration is expected to be $4.34M based on a 5-year time horizon.

1.0 Introduction

CLK Oil sells gasoline at service stations throughout United States and represents 1.5% of US market share in gasoline. CLK sells fuel at over 2,300 service stations nation-wide.[1] While modest in the global market, CLK is a prominent, innovative firm in domestic shale oil production, which began booming since 2010.

Organic growth has become increasingly difficult in the commodity-priced gas industry. CLK has been investigating acquisition opportunities to increase its operational efficiencies. Currently, CLK is considering acquiring DCB Oil, which serves 0.1% of the US market. A small, family-owned company, DCB is known in the industry for sophisticated infrastructure investments. The combination of DCB's innovation in infrastructure and CLK's expertise in shale is expected to generate significant synergies.

[1] ("The U.S. Petroleum Industry: Statistics, Definitions | NACS Online ...." 2013)

The next generation of managers in the DCB family have jeopardized the company, and a merger will benefit DCB Oil's shareholders. Following a series of bad investments and a few public scandals, the market has lost confidence in DCB, as represented by a drop in market share from 0.5% to 0.1%. An injection of modernity into DCB corporate culture will help the brand recover from recent missteps.

Combining DCB infrastructure with CLK expertise in shale gas will lead to decreased duplication in the supply chain. These synergies are expected to increase revenues and reduce costs over the next 5 years, but they have yet to be evaluated fully. CLK's Operation Division expects that integration of gasoline operations between the two companies will have a one-time, upfront cost of $19.5 million.

To see if this cost will be offset by expected revenues, **CLK's Finance Division has forecasted demand, revenue, and cost expectations by combining CLK and DCB gasoline sales.** Traffic volume data was used as a proxy for demand using **benchmarking with multiple regression**. The traffic volume model also incorporated US Working Population, Total US Auto Sales, Long-term Government Bond Yield, and Consumer Price Index data. Gas Price Index timeseries data was used to predict profitability via the **Decomposition Method**.

Post-merger **net present value of gasoline operation integration is expected to be $4.34M** based on a 5-year time horizon. This positive NPV suggests that operational efficiencies will pay for themselves during this time period. This report presents rationale and methods for this conclusion. The report is organized as follows:

- ➔ **2.0 Data Characteristics:** description of key attributes of the traffic volume and gas price index data sets. Traffic volume is also described in relation to additional variables in the model (listed above).
- ➔ **3.0 Model Selection:** evaluation of models for suitability, diagnostics, internal forecasting, and future forecasting. Traffic Volume (demand) and Gas Price Index (profitability) separated for analysis. Net present value of the revenue synergies from combining gasoline sales operations are described.

- ➔ **4.0 Conclusion & Recommendations:** next steps for proceeding with the bid process and insights from the analysis.
- ➔ **Appendices:** additional information on the rationale and methodology of statistical analysis, including additional figures.

To best forecast traffic volume and gas prices, over the next year, CLK first examined important characteristics of both data sets to begin identification of suitable models.

2.0 Data Characteristics

CLK used two sets of data in this analysis which were reviewed for key characteristics: traffic volume and gas price index (GPI). All collected data came from the Federal Reserve Bank of St. Louis except auto sales, which came from the Bureau of Economic Analysis.

2.1 Traffic Volume and Independent Variables

Traffic volume data (miles driven each month) was used as a proxy for gasoline demand. The traffic data set contained monthly observations from January 1977 to February 2017.
traffic volume exhibits trend, seasonal, and cyclical effects **(Figure 1)**. Traffic volume peaks in summer and dips in February and March, with an overall upward trend. Growth in this trend decreased in the late 1970s in response to the oil crisis and the recession in the late 2000s, indicating two important cyclical effects. No visible outliers are present.

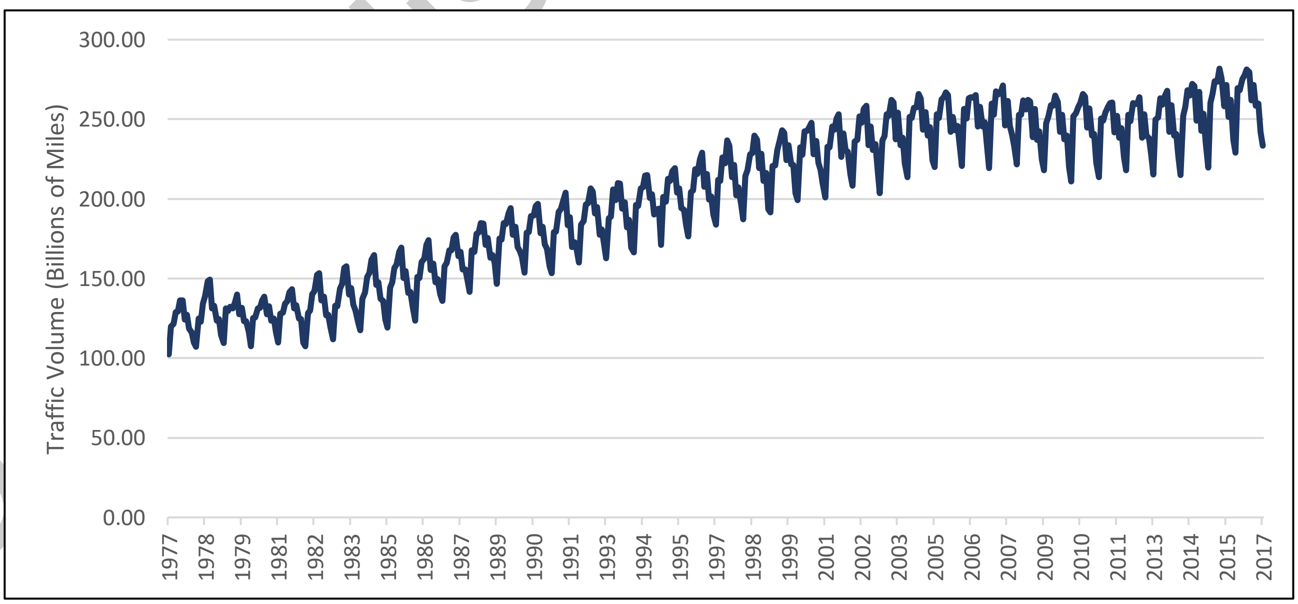

Figure 1: Time series plot of US traffic volume from 1977 to 2017

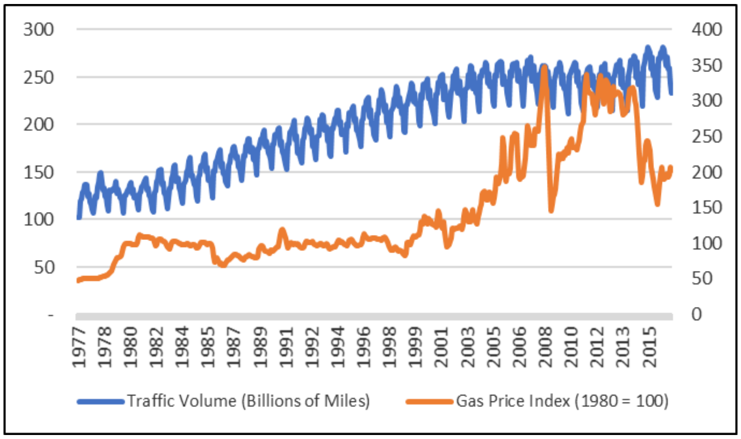

Figure 2: Traffic volume with the corresponding gas price index overlaid

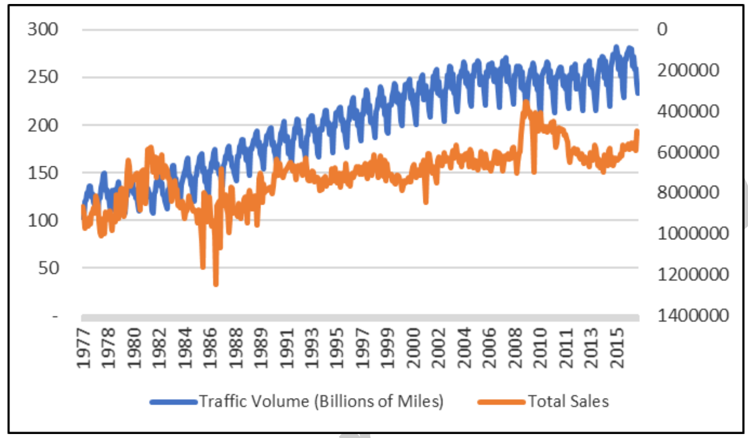

Figure 3: Traffic volume with corresponding total vehicle sales overlaid (inverted axis)

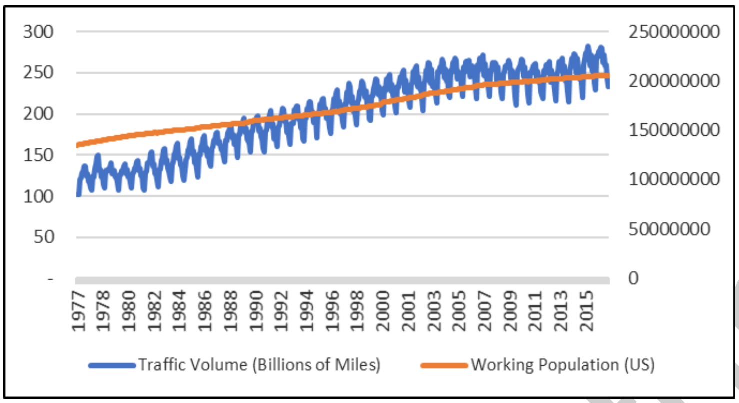

Figure 4: Traffic volume with corresponding working population overlaid

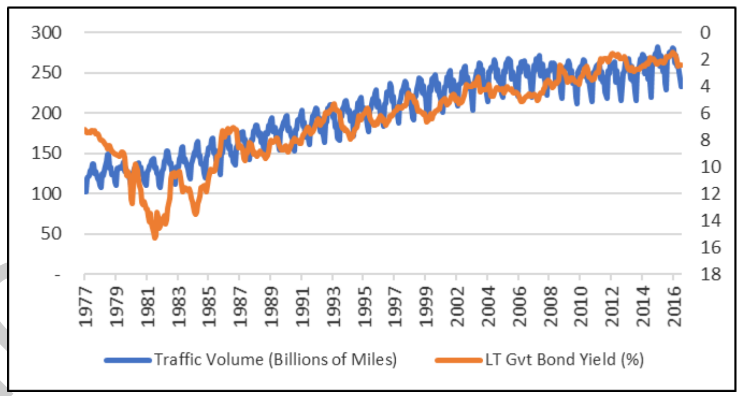

Figure 5: Traffic volume with corresponding long term gov. bond yields overlaid (inverted axis)

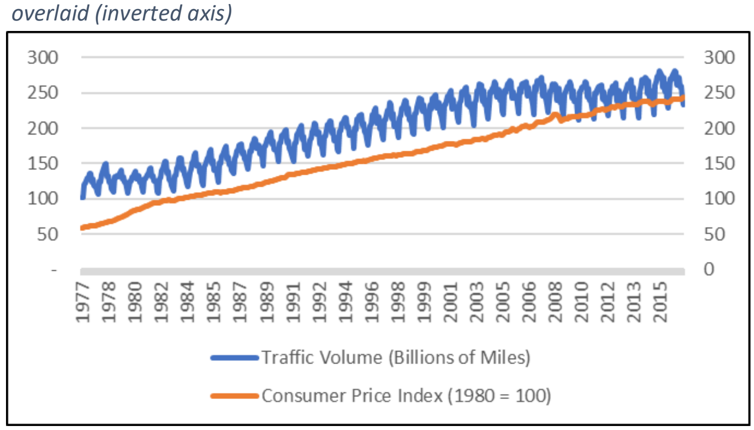

Figure 6: Traffic volume with corresponding consumer price index overlaid

Gas price index (top left), US total auto sales (top right), US working population (bottom left), long-term government bond yield (center right), and consumer price index (bottom right) were used in the model to forecast traffic volume as independent variables. Time series plots of all independent variables **(Figures 2-6)** indicate a relationship with traffic volume. Gas price index had the most pronounced cyclical disturbances (described in detail below). Long-term government bond yield and total US auto sales have a inverse relationship with traffic volume. These two variables also had the most evidence of cyclical effect, which leads to the expectation that some degree of uncertainty exists that needs to be controlled for in the model to get the best forecasts.

2.2 Gas Price

CLK evaluated GPI data to determine its expected profitability, examining it independently of traffic volume **(Figure 7)**. The GPI data also shows trend, seasonal, and cyclical effects. Impact of the trend effect between 2004 and 2016 is difficult to clearly understand because the cyclical effect is most prominent and volatile during this period. This cyclical effect reflects the recession leading up to 2009. The impact of this effect may make time series forecasting beyond a couple of periods ill-advised.

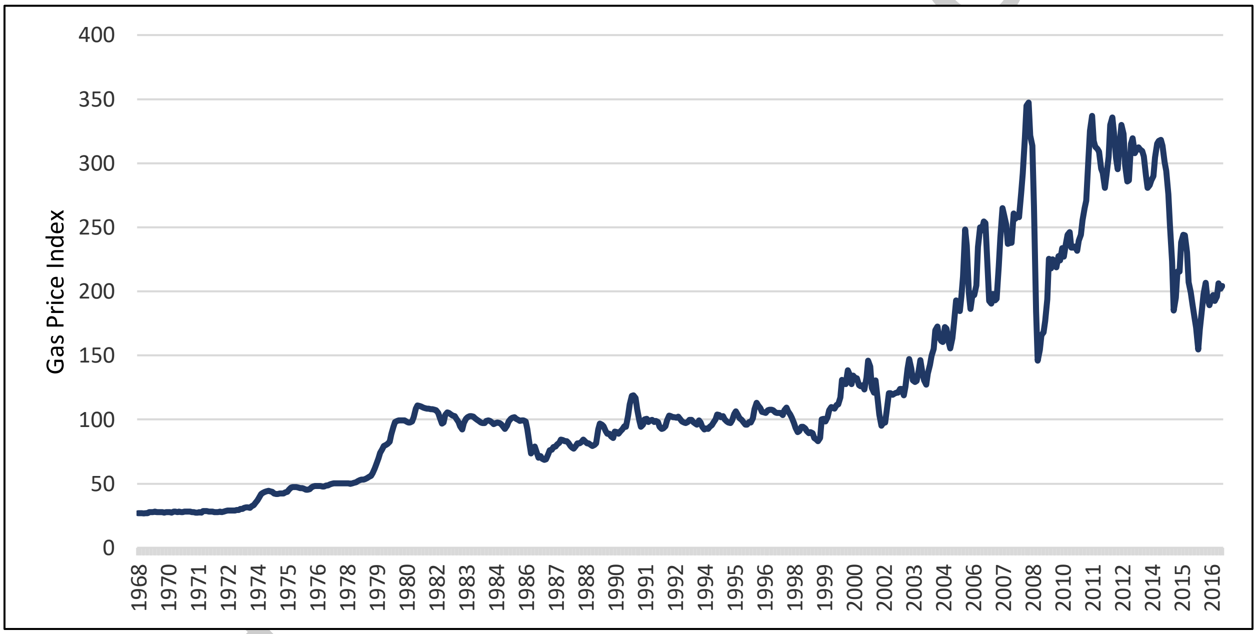

Figure 7: Time series plot of the gas price index from 1968 to 2017 (1980 = 100 [$1.19 per gallon])

While both timeseries data sets have shared effects, other characteristics, including availability of independent variables, suggest different modeling approaches may be more appropriate for each data set.

→ Availability of correlated variables indicate that **multiple regression with benchmark forecasting** is a good model candidate for traffic volume. The prominence of trend, cyclical, and seasonal effects also lead to the consideration of the decomposition method.

→ For GPI, **decomposition method** is appropriate because of trend, seasonal, and cyclical effects. Double exponential smoothing places more weight on recent events, such as the recession, which may have had a more dramatic impact on the behavior of and variation in gas price index. Thus, double exponential smoothing is another candidate.

3.0 Model Selection, Assessment, Forecasting, & Interpretations

This section details four proposed statistical models—two for demand and two for profitability[2]—to determine which can best fit data and make accurate predictions. Subsection 3.1 evaluates models for traffic volume and subsection 3.2 evaluates models for GPI. Each subsection presents a rigorous set of criteria used during the evaluation including error measurements, model diagnostics, and internal forecasting. Interpretations, including Net Present Value estimates can be found in subsection 3.3.

3.1 Traffic Volume

The two models considered to fit the traffic volume data were **multiple linear regression,**[3] later to be used for **benchmark forecasting,[4]** and **decomposition.**[5] Multiple regression—which ultimately provided the best fit—incorporates a variety of correlated economic indicators (independent variables) to produce accurate estimates. Decomposition produces accurate forecasts based on seasonal, cyclical, and trend effects. While these approaches serve different purposes, comparing both best acknowledges how both time and variable dependence factor into traffic volume predictions. Note that both models assume residuals[6] follow a normal distribution with an average value of zero and a constant variance of σ^2.[7]

---

[2] Refers to the profitability of gasoline sales only. Does not take into account DCB purchase price.
[3] (DeLurgio 1998, page 24)
[4] Benchmark forecasting is a method used to predict a response variable with a multiple regression model. Each independent variables' values are chosen independently and carefully to represent the data.
[5] (DeLurgio 1998, page 600)
[6] (DeLurgio 1998, page 99)
[7] A normal distribution follows a typical bell curve where the majority of data are closest to the average of the set and as you move away from the average (either increasing or decreasing), there are fewer and fewer data points. σ^2 represents the variance, or a measure of spread within the distribution.

3.1.1 Model Fitting

Multiple linear regression for traffic volume had four continuous predictors (described in section 2.1).[8] Indicator variables[9] were also used to represent months (with February as the benchmark)**.** Initial attempts at the multiple regression resulted in non-normally distributed residuals with non-constant variance, and non-statistically significant independent variables[10]. **A Box Cox transformation**[11] of the traffic volume data and the addition a single period lag of the transformed traffic volume data[12] resulted in a model that met assumptions and had exceptional fit—the resulting **mean squared error** (**MSE)[13] being 0.014 and mean absolute percentage error (MAPE)[14] being 0.66%.** Note that lower error is more desirable. Below is the final model structure used for the multiple regression.[15]

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \hat{\beta}_4 x_{i4} + \hat{\beta}_5 x_{i5} + \hat{\beta}_6 x_{i6} + \hat{\beta}_7 x_{i7} + \hat{\beta}_8 x_{i8} + \hat{\beta}_9 x_{i9} \ldots + \hat{\beta}_{19} x_{i19}$$

$$x_{i9} - x_{i19} \begin{cases} 1: the\ i-th\ obs\ is\ in\ the\ 1st\ category, \boldsymbol{March - January} \\ \quad 0: otherwise, \qquad \boldsymbol{February} \end{cases}$$

Model Rewrite:

$$\hat{y}_i = \begin{cases} \left(\hat{\beta}_0 + \hat{\beta}_9\right) + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \hat{\beta}_4 x_{i4} + \hat{\beta}_5 x_{i5} + \hat{\beta}_6 x_{i6} + \hat{\beta}_7 x_{i7} + \hat{\beta}_8 x_{i8}, \qquad \boldsymbol{March} \\ \left(\hat{\beta}_0 + \hat{\beta}_{10}\right) + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \hat{\beta}_4 x_{i4} + \hat{\beta}_5 x_{i5} + \hat{\beta}_6 x_{i6} + \hat{\beta}_7 x_{i7} + \hat{\beta}_8 x_{i8}, \ \boldsymbol{April} \\ \qquad \ldots \ldots \ldots \ldots \\ \beta_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \hat{\beta}_4 x_{i4} + \hat{\beta}_5 x_{i5} + \hat{\beta}_6 x_{i6} + \hat{\beta}_7 x_{i7} + \hat{\beta}_8 x_{i8}, \qquad \boldsymbol{February} \end{cases}$$

The **decomposition model was also considered.** This modeling technique estimates three components: trend (T_t), seasonal (S_t), and cyclical (C_t). Using a multiplicative structure, the decomposition model takes the below forecasting structure.[16]

$$\hat{Y}_{t+k} = \hat{T}_{t+k} * \hat{S}_{t+k} * \hat{C}_{t+k}$$

---

[8] An independent variable whose value is not discrete, i.e. 0 or 1.
[9] (DeLurgio 1998, page 412)
[10] These issues, if left unaddressed, can lead to inaccurate forecasts and/or wide prediction intervals.
[11] Data is exponentially transformed by an exponent represented by λ. Transformations are performed from λ -5 to 5 until the data is normally distributed.
[12] (DeLurgio 1998, page 478)
[13] (DeLurgio 1998, page 55)
[14] (DeLurgio 1998, page 55)
[15] Full model with numeric estimated parameters are found in Appendix D.
[16] The decomposition model's component estimators are found in Appendix D.

Each component was optimized to fit the traffic volume data, achieving a fit with a **MSE of 0.018 and a MAPE of 0.74%,** which was not as good as the multiple regression model.[17] **Figure 8** depicts a time series plot of both models' fits and the original traffic volume data. The fits were converted back from the box cox transformed units to their original units (billions of miles).

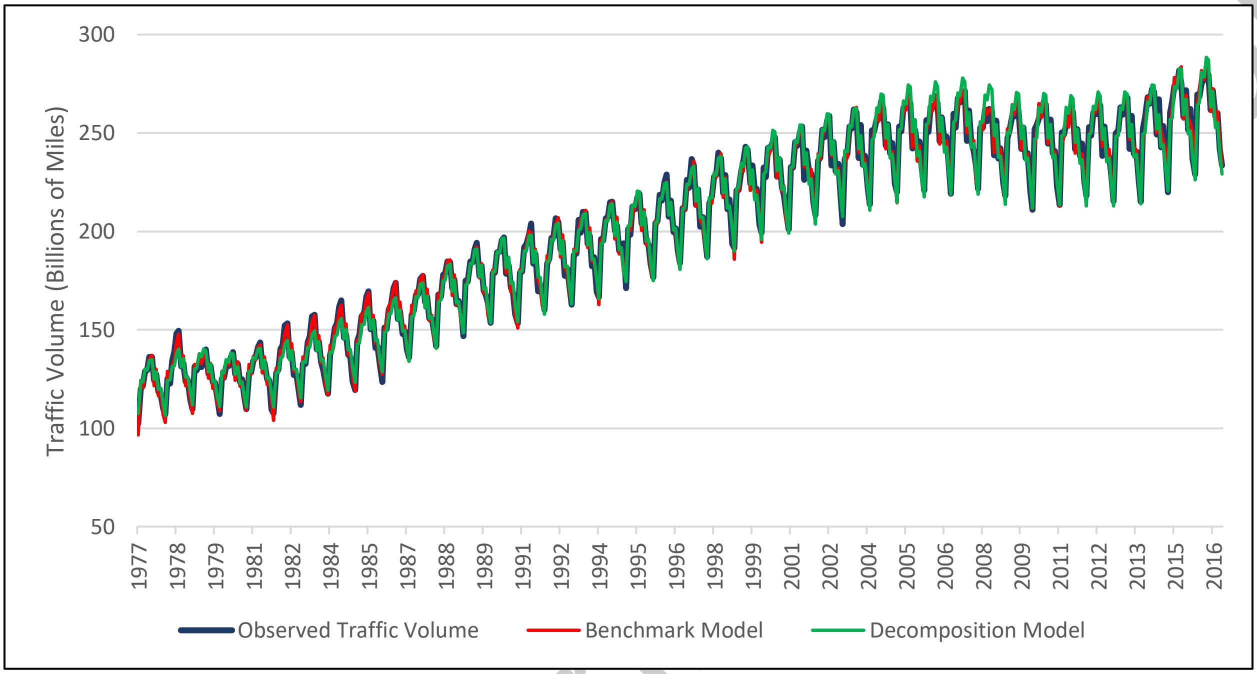

Figure 8: Time series plot of observed traffic volume, the multiple regression model fits, and the decomposition model fits, 1977-2017

3.1.2 Model Assessment

To ensure the models used were valid, critical assumptions[18] about the models were tested through diagnostics. Diagnostics revealed that the benchmark model passed all assumptions and decomposition model passed all except normality.[19]

Multi-collinearity[20] is also a factor that should be considered when performing model diagnostics. Certain independent variables in the benchmark model were found to have high amounts of interdependency, but this can be explained by the nature of the variables and their relation to time.

---

[17] Note that the Box Cox transformation of the traffic volume data was utilized in the decomposition model as well.

[18] Critical assumptions include: normality, homoscedasticity, independence & linearity.

[19] See Appendix C for detailed results of diagnostics.

[20] Highly correlated independent variables, which leads to more complex regression interpretations (DeLurgio 1998, page 114)

Based on the diagnostics, the multiple regression model passed all tests to meet the assumptions. The decomposition model failed to meet the normality assumption, which means the model could potentially produce less accurate forecasts.

3.1.3 Internal Forecasting

Using the multiple regression model, CLK conducted internal forecasting using the benchmark method to determine how well the model predicted observed data. The model was fit using the first 469 observations (February, 1977 – February, 2016). The model was then used to forecast the last 12 months of observations (March, 2016 – February, 2017). The error between the internally forecasted and observed values was used to validate the forecasting model.

The multiple regression model with chosen benchmarks[21] yielded a MSE of 0.009 and a MAPE of 0.50%. Figure 9 shows traffic volume for the past four years, the model's fit from March, 2013 – February, 2016, and the model's internal forecast for the last 12 months. It is clear from the figure that the expected value of the internal forecast is closest to the observed value (versus the model's conservative or aggressive estimates).

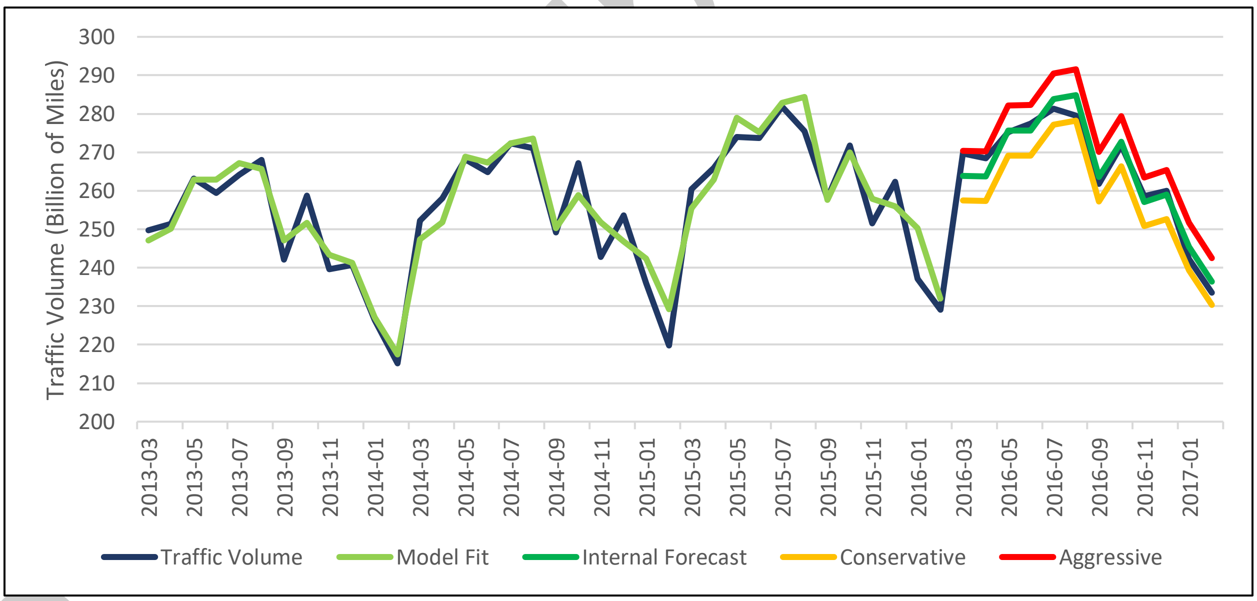

Figure 9: Observed traffic volume with the benchmark model fit and internally forecasted values, March 2013-Jan. 2017

The decomposition model was validated in the same fashion as the multiple regression model. The same time period was used to fit the model to forecast the same last 12 months of the data.

[21] Benchmarks used: Gas Price Index: 273.6 (average value of last 60 observations), Working Population: 205,208,000 (max value of the data set), Total Sales: 710,817 (average of all observations), Long-Term Government Bond Yield: 2.268 (average of last 24 observations), CPI: 238.65 (max value of the data set).

The decomposition model yielded a MSE of 0.019 and a MAPE of 0.74%, which is inferior to the multiple regression model. Figure 10 shows the same data as Figure 9 but for the decomposition's fit and internal forecast against the observed values for traffic volume. As with the multiple regression model, the expected value of the decomposition's internal forecast is closest to the observed values (versus the model's conservative or aggressive estimates).

The multiple regression model achieved better fit than the decomposition model, met all critical model assumptions, and achieved less error during internal forecasting. The multiple regression model was selected based on superior fit, meeting all assumptions, and achieving lower error measurements during internal forecasting.

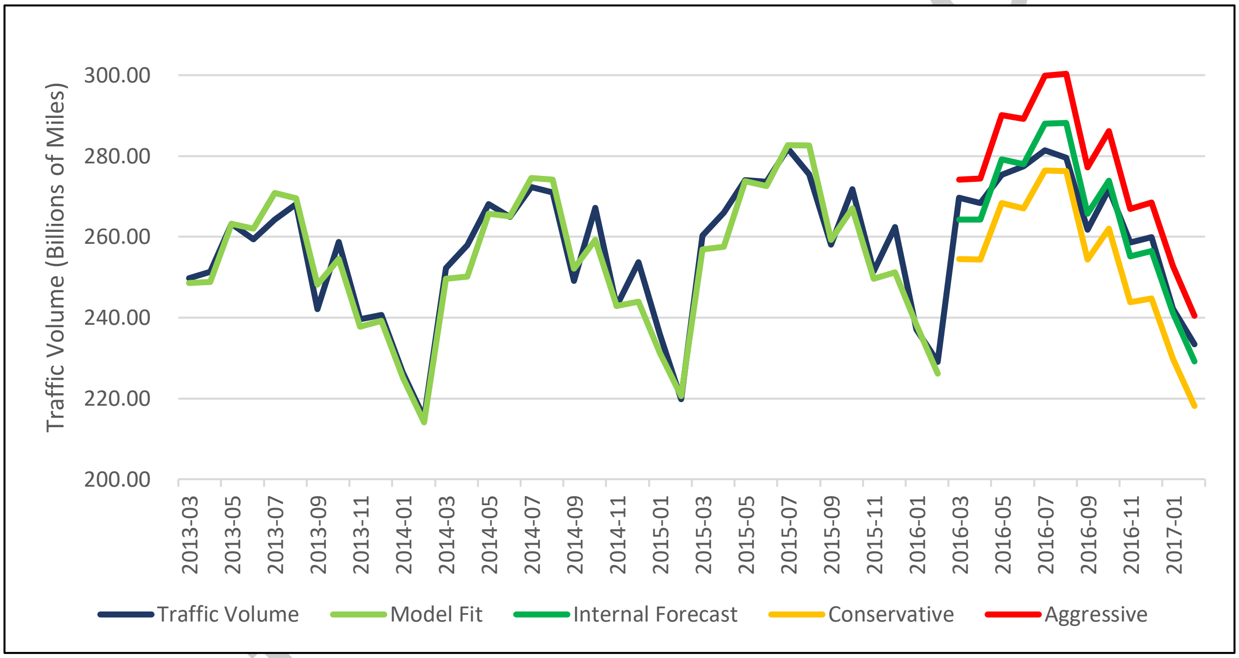

Figure 10: Observed traffic volume with the decomposition model fit and internally forecasted values, March 2013-Jan. 2017

3.1.4 Future Forecasting

The multiple regression model using benchmark forecasting was used to forecast the next 12 months (March, 2017 – February, 2018) of traffic volume. This time, the model was fit from the entire data set (February, 1977 – February, 2017). The benchmarks used for future forecasting were selected on the same basis as internal forecasting.[22] Figure 11 shows the future forecast

[22] Benchmarks used: Gas Price Index: 257.78 (average value of last 60 observations), Working Population: 205,903,000 (max value of the data set), Total Sales: 707,086 (average of all observations), Long-Term Government Bond Yield: 2.03 (average of last 24 observations), CPI: 243.60 (max value of the data set).

using the benchmark method for the multiple regression model. For the forecasted period (March, 2017 – February, 2018), this **model forecasted 3.154 trillion miles driven in the U.S.**

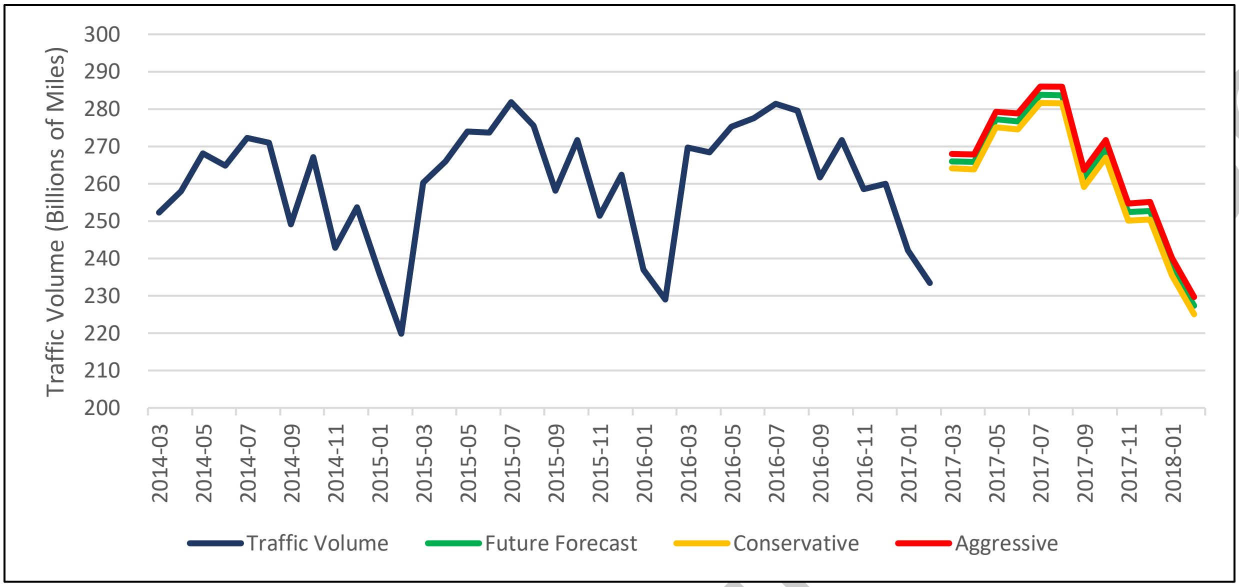

Figure 11: Observed traffic volume and benchmark forecasted traffic volume, March 2014 – Jan. 2018

The forecasts of gas prices for the same period will allow CLK to analyze the financial viability of the purchase of DCB Oil.

3.2 Gas Prices

Decomposition and **double exponential smoothing**[23] were the two best options for modeling GPI. Time series models were selected for GPI forecasting rather than benchmarking with multiple regression to focus on the timeseries effects instead of relationships between GPI and other variables. Further, using the same continuous variables as traffic volume would create dependency between the GPI variable and the benchmark forecasting of traffic volume.

3.2.1 Model Fitting

The **decomposition model** was used to fit **GPI**, after undergoing a **Box Cox transformation** (similar to traffic volume)**.** Decomposition was ultimately selected for future forecasting. This model structure is identical to the one used to predict traffic volume in section 3.1. The model achieved a **MSE of 0.362 MAPE of 3.15%.** See Figure 12 for fits.

[23] (DeLurgio 1998, page 218-222)

Brown's Double exponential smoothing was also evaluated, and the model structure can be observed below. Using an optimized alpha of 0.2 and beta of 0.01, a **MSE of 0.407 and MAPE of 3.29%** was achieved with the double exponential smoothing model.

$$\boldsymbol{S_t'} = aY_t + (1-a)S_{t-1}' \qquad \boldsymbol{S_t''} = aS_t' + (1-a)S_{t-1}''$$

$$\boldsymbol{a_t} = 2S_t' - S_t'' \qquad \boldsymbol{b_t} = \frac{a}{1-a}(S_t' - S_t'')$$

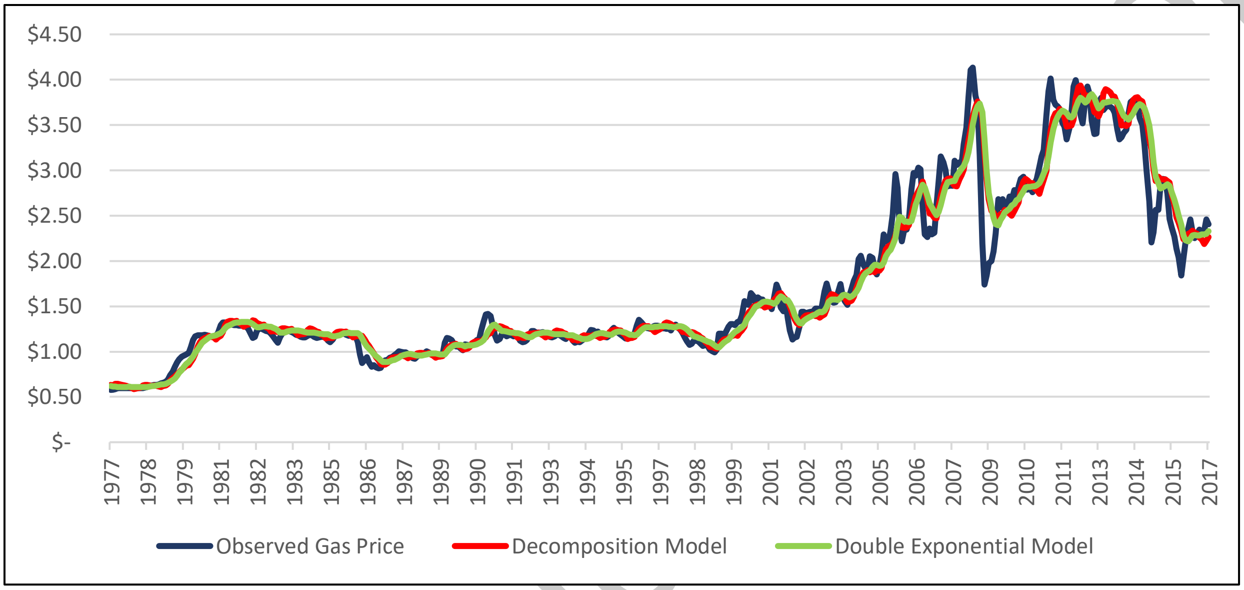

Figure 12: Observed gas prices with decomposition and double exponentially smoothed model fits, 1977-2017

3.2.2 Model Assessment

As before, four primary assumptions must be checked to validate that the models' predictions will be accurate. Although neither model met all the assumptions, they came closer to meeting them than any other model vetted.[24] Predictions made by either of these models should be accompanied by wide prediction intervals, and multiple scenarios should be produced and vetted using said prediction intervals.

3.2.3 Internal Forecasting

Internal forecasting was used to determine the accuracy of each model. Each model was fitted using the first 469 observations, and then each models' forecasts were compared against the last observed 12 months (in the same fashion as section 3.1.3).

[24] See Appendix C for detailed diagnostics results.

The decomposition method yielded a MSE of 0.162 and MAPE of 2.30% for its internal forecast of gas prices. Figure 14 shows gas prices for the past four years, the model's fit from March, 2013 – February, 2016, and the model's internal forecast for the last 12 months. It is clear from the figure that the expected value of the internal forecast is closest to the observed value (versus the model's conservative or aggressive estimates).

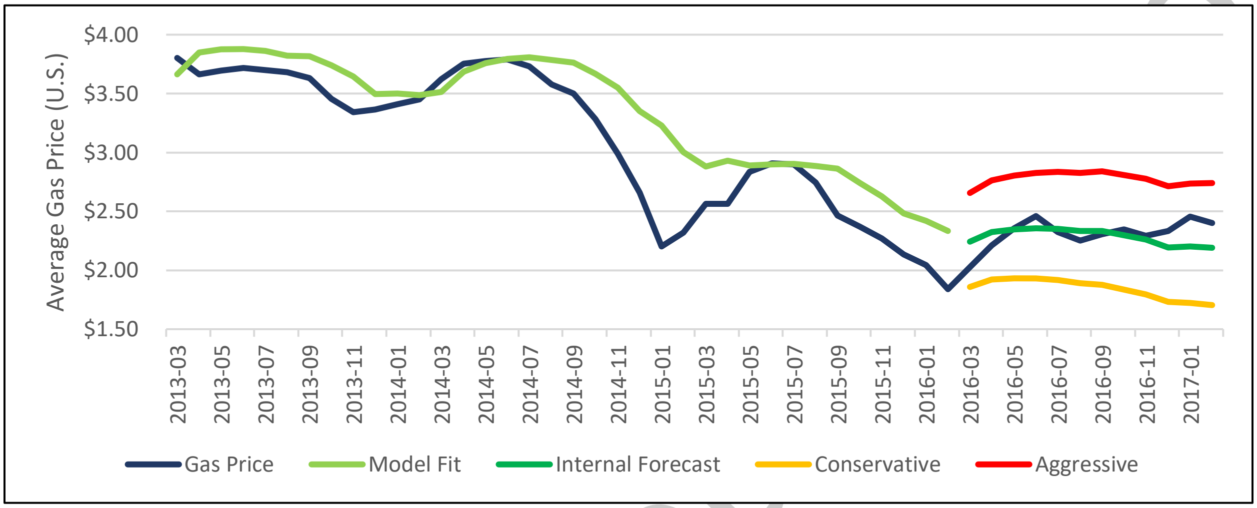

Figure 13: Internal forecast of gas price indices using the decomposition model, March 2013 – Jan. 2017

The double exponential smoothing method yielded a MSE of 0.149 and MAPE of 2.26% for its internal forecast of gas prices. Figure 13 shows gas prices for the past four years, the model's fit from March, 2013 – February, 2016, and the model's internal forecast for the last 12 months. In Figure 14, expected value of the internal forecast is closest to the observed value (versus the model's conservative or aggressive estimates).

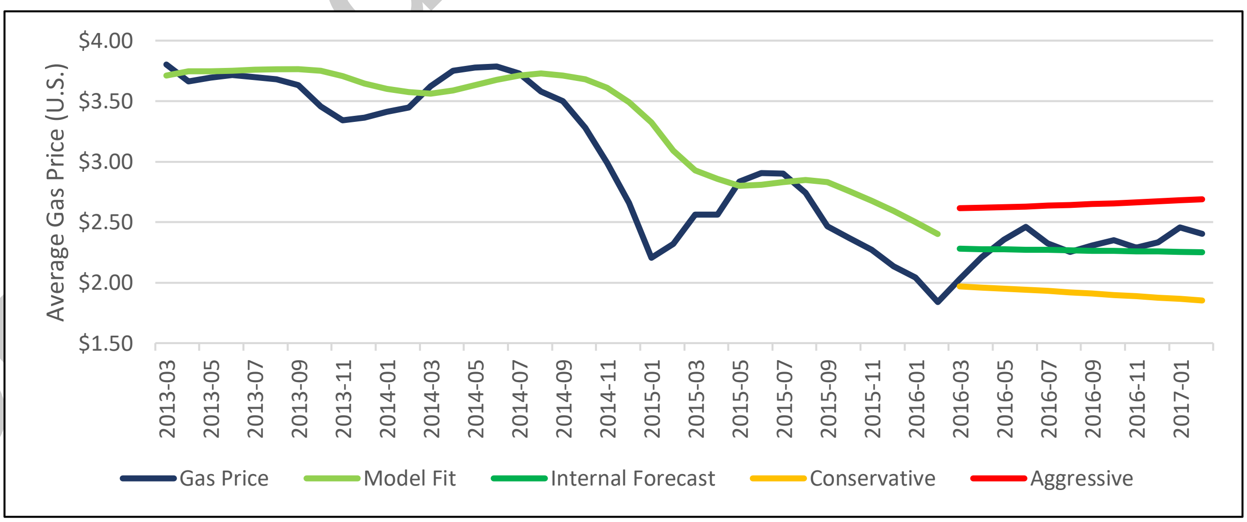

Figure 14: Internal forecast of gas price indices using the double exponential smoothing model, March 2013 – Jan. 2017

The two models yielded similar results. The double exponential smoothing method achieved slightly better accuracy during internal forecasting. However, the decomposition method achieved better overall fit of the data. Despite both models failing critical model assumptions, the decomposition model had slightly less heteroscedasticity[25] than the double exponential smoothing model. The decomposition's forecast followed a seasonal structure, as opposed to a relatively flat line produced by the double exponential smoothing method. Knowing this, the decomposition method was chosen for future forecasting acknowledging that a larger prediction interval will be applied to our forecast to compensate for the model's error and failure to meet critical assumptions.

3.2.4 Future Forecasting

The decomposition model was used to forecast the next 12 months (March, 2017 – February, 2018) of gas prices. This time, the model was fit from the entire data set (February, 1977 – February, 2017).

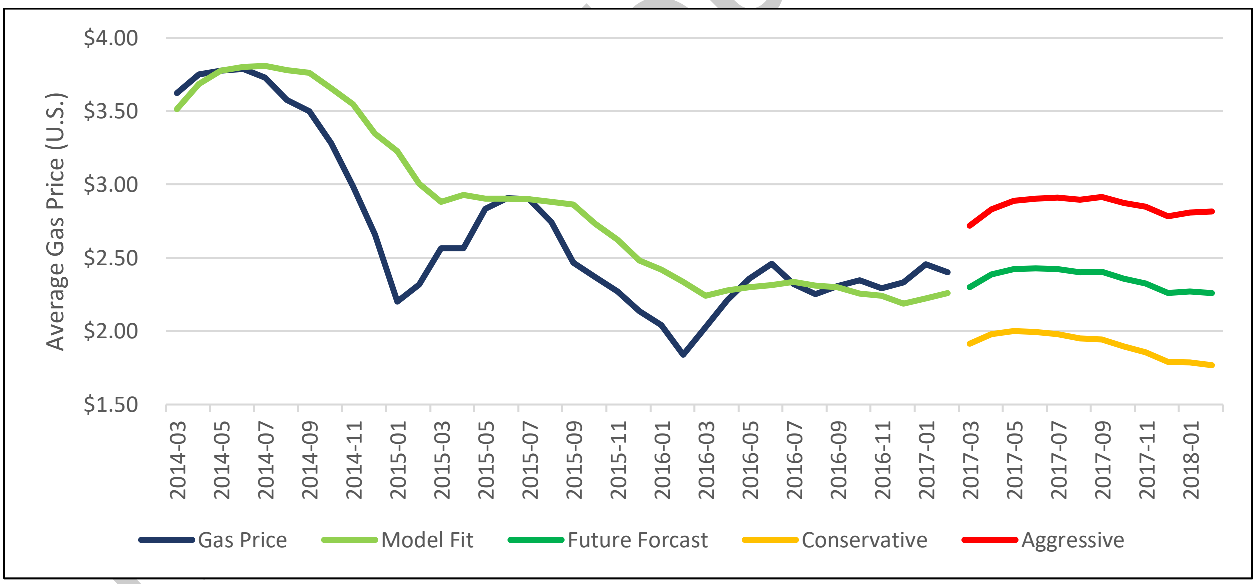

Figure 15: Decomposition model future forecast of gas price indices, March 2014 – Jan. 2018

Figure 15 shows the future forecasts using the decomposition method. For the forecasted period (March, 2017 – February, 2018), the values predicted for gas prices take on the seasonal structure of years past. It is important to note that because gas prices were more difficult to

[25] Non-constant variance (critical assumption)

forecast than traffic volume, the error in our final model resulted in a wider spread using prediction intervals for each months' forecast (approximately $0.50 above and below).

With the forecast of traffic volume (demand) and gas prices (profitability) for the next 12 months, a financial analysis of the acquisition of DCB could be conducted.

3.3 Interpretations

The goal of this analysis aims to evaluate whether the costs associated with gasoline operations integration between the two companies is a viable investment. To assure post-merger CLK holds market share at 1.6%,[26] CLK must invest $19.5M upfront to assure operation efficiencies are achieved. Net present value of this investment is calculated using discounted cash flows over a 5-year horizon.

Without the purchase of DCB, CLK Oil's future 12-month net profit from fuel sales will be $84.5M ± $16.6M.[27] Figure 16 shows the profit by month for CLK Oil for the next 12 months.

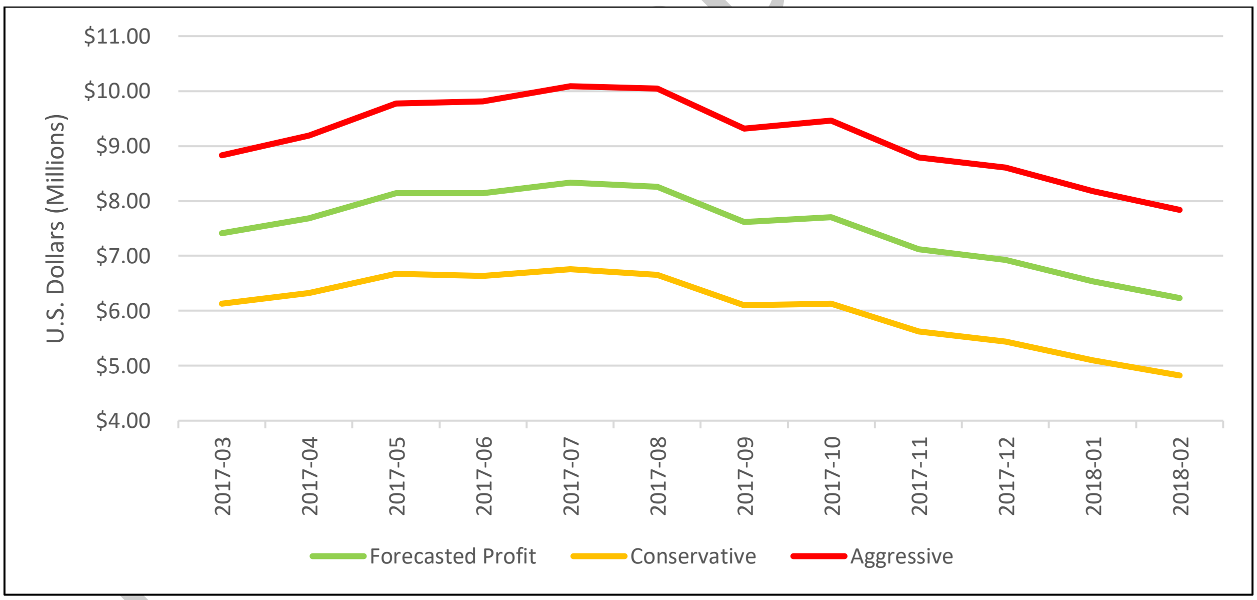

Figure 16: CLK Oil's forecasted profits, March 2017 – Feb. 2018

Post-merger, CLK expects future 12-month net profit from fuel sales to increase to $90.1 ± $17.7M. This is an increase in $5.63M for the next 12 months. Assuming a variable annual

[26] 1.5% for CLK + 0.1% for DCB = 1.6% market share, contingent upon gasoline operations integration
[27] See Appendix A for the calculation of net profit forecasts.

growth rate between 0.35% and 0.47%, profits were extrapolated for four more years (out to 2022).[28]

With a cost of capital at 6.5% (anticipated for CLK post-merger), a 2% margin on gasoline sales, and an estimated integration cost of $19.5M, the **net present value of gasoline operational efficiencies is $4.34M**. The NPV remains positive for this investment if integration cost remains below $23.84M**.** However, if the conservative and aggressive estimates of traffic volume and gas prices are applied, we can assess the volatility of the endeavor.[29] Table 1 shows the net present value of the venture when the conservative and aggressive estimates are applied.[30]

Year	2017	2018	2019	2020	2021	2022	
Period	0	1	2	3	4	5	
	Expected Estimates						Likelihood
Cash Flow	$ (19.50)	$ 5.63	$ 5.63	$ 5.66	$ 5.68	$ 5.71	75%
Present Value	$ (19.50)	$ 5.29	$ 4.97	$ 4.68	$ 4.42	$ 4.17	
Net Present Value	**$ 4.02**						
	Conservative Estimates						
Cash Flow	$ (19.50)	$ 4.53	$ 4.53	$ 4.55	$ 4.57	$ 4.59	10%
Present Value	$ (19.50)	$ 4.25	$ 3.99	$ 3.76	$ 3.55	$ 3.35	
Net Present Value	**$ (0.60)**						
	Aggressive Estimates						
Cash Flow	$ (19.50)	$ 6.87	$ 6.87	$ 6.90	$ 6.93	$ 6.96	15%
Present Value	$ (19.50)	$ 6.45	$ 6.06	$ 5.72	$ 5.39	$ 5.08	
Net Present Value	**$ 9.20**						
Weighted NPV	$ 4.34						

Table 1: Net present value calculations for the DCB operations integration

With the volatility applied, the **worth of the project is expected to be $4.03M over five years** with the potential of **either losing $0.59M or gaining $9.20M** over the investment horizon. The **weighted net present value, based on the likelihood of each scenario, is $4.34M**.[31]

[28] A S-Curve model was applied to the traffic volume, then monthly values were aggregated into yearly estimates to determine the growth rate. See appendix D for model application.
[29] In this analysis, the compounding error for an estimate (net profit) using two separate forecasts (traffic volume and gas prices) can directly translate to the volatility of the endeavor (unknown risk).
[30] See appendix A for full calculations.
[31] See appendix A for calculation of weighted net present value.

4.0 Conclusion & Recommendation

CLK's impending acquisition bid prompted CLK's finance division to evaluate the synergy value from gasoline operational efficiencies. Analysis revealed low, but steady growth in annual gasoline consumption, offset by annual declines in prices at the pump. This assumption, combined with the probability of declining gasoline prices, led to the conclusion that integrating CLK and DCB's gasoline operations is a positive NPV investment. Below are recommendations derived from our analysis.

- → With an expected integration cost of $19.5M, CLK can expect a positive NPV investment for operations synergies equaling $4.34M. However, if it is revealed during negotiations that integration cost will not be $19.5M, it will impact the NPV of this investment. If costs exceed $23.84M, it will no longer be a positive NPV investment.
- → Although statistically improbable, conservative estimates project up to $600,000 in net losses over 5 years resulting as a possibility. It is important to note that although the possibility of net losses from gas sales exists, the intrinsic value of building the CLK brand with innovative, sophisticated infrastructure may have a value worth more than $600,000 over 5 years. CLK Finance recommends that the M&A Team consider how other sources of value affect this estimation.
- → In an industry highly exposed to systematic risk like oil and gas, any possibility for mitigating risk should be seized. Using individuals with extensive experience in mergers and acquisitions to negotiate the purchase of the refueling stations would be prudent, given the narrow integration budgetary restraints. CLK's M&A team is highly proficient in negotiations and has a track record of successful purchases and sales in the oil and gas industry. Share this analysis with them if the decision to move forward with this strategic purchase is approved.
- → As time progresses, ensure that both the multiple regression benchmark model for predicting traffic volume and the decomposition model for predicting gas prices are continuously updated with new monthly data. The added data will allow for model refinement and increases in near-term forecast accuracy.

Based on the valuation of post-acquisition gasoline operations, we have no evidence to suggest that the acquisition of DCB Oil is unfavorable. Integrating the $4.34 million NPV with other valuation factors will help CLK determine if this acquisition will benefit the corporation and its stakeholders. Industry competitiveness and a volatile environment reinforce the need to invest in superior sales infrastructure that support our competitive advantage. This uncertainty can be greatly reduced by continuing to improve our forecasting and valuation techniques. Success of this venture and the business analytics supporting it will open the door to identifying other strategic investment opportunities in the future.

Appendices: Contents

The following appendices include all major calculations during our team's analysis and forecasts. There are four appendices:

- → **APPENDIX A:** Financial analysis calculations and supporting data
- → **APPENDIX B:** Includes error measures for model fitting and internal forecasting for both traffic volume and gas prices
- → **APPENDIX C:** Supporting data for all model diagnostics.
- → **APPENDIX D:** Includes calculations made during forecasting

Appendix A: Supporting Financial Analysis

Gallons Sold Constant[32]

Traffic Volume (Billions of Miles) to CLK Gallons Sold (in millions) @ 1.5% market share = 0.56818

Traffic Volume (Billions of Miles) to CLK Gallons Sold (in millions) Sold @ 1.6% market share = 0.60606[33]

CLK Monthly Profit Forecast

Date	Traffic Volume Forecast	CLK Gallons Sold (Millions) Pre-Acquisition	CLK Gallons Sold (Millions) Post-Acquisition	Gas Price Forecast	Cost	Pre-Acquisition Profit (Millions USD)	Post-Acquisition Profit (Millions USD)
Mar-17	266.03	151.15	161.23	$ 2.30	$ 2.25	$ 6.95	$ 7.42
Apr-17	265.79	151.02	161.09	$ 2.39	$ 2.34	$ 7.21	$ 7.69
May-17	277.20	157.50	168.00	$ 2.42	$ 2.38	$ 7.64	$ 8.14
Jun-17	276.65	157.19	167.66	$ 2.43	$ 2.38	$ 7.63	$ 8.14
Jul-17	283.81	161.26	172.01	$ 2.42	$ 2.37	$ 7.81	$ 8.33
Aug-17	283.74	161.22	171.96	$ 2.40	$ 2.35	$ 7.74	$ 8.26
Sep-17	261.41	148.53	158.43	$ 2.40	$ 2.36	$ 7.14	$ 7.62
Oct-17	269.29	153.01	163.21	$ 2.36	$ 2.31	$ 7.22	$ 7.70
Nov-17	252.38	143.40	152.96	$ 2.33	$ 2.28	$ 6.67	$ 7.12
Dec-17	252.75	143.61	153.18	$ 2.26	$ 2.22	$ 6.49	$ 6.93
Jan-18	237.80	135.12	144.12	$ 2.27	$ 2.22	$ 6.13	$ 6.54
Feb-18	227.36	129.18	137.79	$ 2.26	$ 2.22	$ 5.84	$ 6.23
						$ 84.49	$ 90.12

Table A.1: CLK's Monthly Profit for pre- and post-acquisition of DCB Oil from March, 2017 – February, 2018

[32] Gallons Sold Constant used to calculate the number of gallons sold based on the forecast of Traffic Volume is based on the market share parameters as well as the current average fuel economy in the U.S.)
[33] Both Constants are based off the U.S. Average Fuel Economy of 26.4 Miles per Gallon

Monthly Profit Calculation

*CLK Gallons Sold (in Millions) = Traffic Volume (in Billions of Miles) * (Gallons Sold Constant)*

*Monthly Profit (Millions of USD) = [Gallons Sold (in Millions) * Gas Price] – [Gallons Sold * Cost]*[34]

Example for March, 2017 (pre-acquisition profit):

*CLK Gallons Sold (in Millions) = 266.03 Billions of Miles * 0.56818 = 151.15 Million Gallons*

*Monthly Profit = [151.15M gallons * $2.30] – [151.15M gallons * $2.30 * 0.98] = $6.95 Million*

Fuel Demand Growth Rate Estimation for Investment Horizon

To estimate growth in demand for fuel, an S-curve exponential model was fit to the monthly traffic time series data. Once the parameters were estimated, a forecast could be extrapolated. Because we were only concerned with annual growth, the monthly data was aggregated into annual forecasts.

The S-curve model takes the following model structure:[35]

$$y_t = \frac{1}{\beta_0 + \beta_1 * \beta_2^t} + e_t$$

When fit to the traffic volume data, the model's parameters were estimated with the form:

$$\widehat{y_t} = \frac{1}{0.0036 + 0.0057 * 0.99^t}$$

Figure A.2 shows the S-Curve's fits and future forecast for the investment horizon.

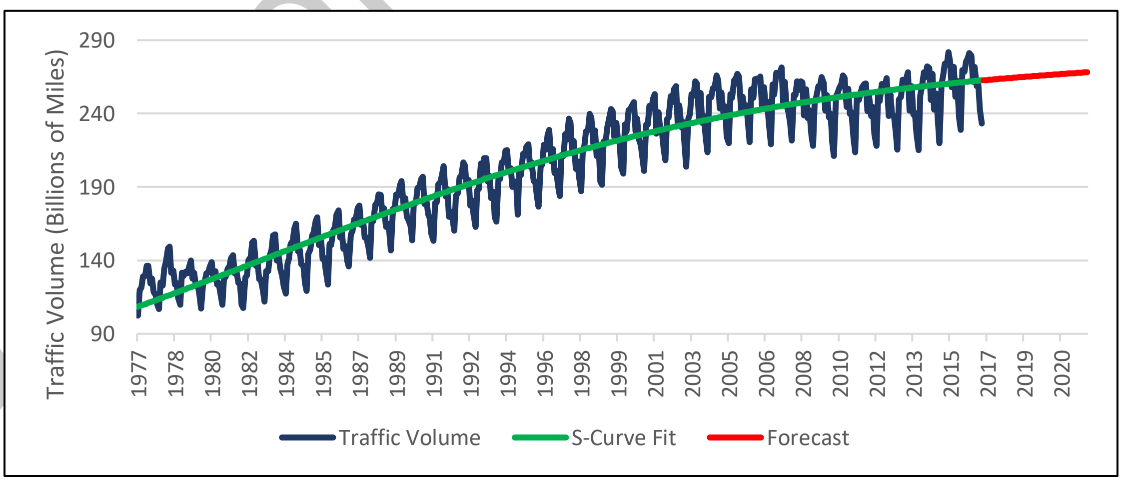

Figure A.2: S-curve fit to traffic volume data with forecast covering the investment horizon (5 years)

---

[34] Cost per gallon is determined by applying a two percent profit margin to each gallon sold.
[35] Random error (e_t) is added to the model structure to account for unknown or unpredictable variation.

Table A.3 shows the aggregated forecast for traffic volume. The annual growth is estimated. These growth rates serve as a proxy for demand growth during net present value estimation.

Year	Annual Traffic Volume Forecast	Annual Growth
2017	3,158.98	0.47%
2018	3,173.72	0.47%
2019	3,187.57	0.44%
2020	3,200.48	0.41%
2021	3,212.50	0.38%

Figure A.3: Aggregated annual traffic volume forecast with associated growth rate (Traffic Volume in Billions of Miles)

Weighted Net Present Value Calculation

It was estimated that the likelihood of the expected Net Present Value (NPV) for this project was 75%. The likelihood of observing conservative values for both traffic volume and gas prices is 10%. The likelihood of observing aggressive values for traffic volume and gas prices is 15%.

Expected NPV = $4.03M

Conservative NPV estimate = – $0.59M

Aggressive NPV estimate = $9.20M

*Weighted NPV = ($4.03 * 0.75) – ($0.59M * 0.10) + ($9.20 * 0.15) = $4.34M*

Appendix B: Error Measurements

Traffic Volume

	Model Fit		Internal Forecasting	
	Benchmark	Decomposition	Benchmark	Decomposition
ME	2.28E-15	(0.005)	(0.006)	(0.007)
MSE	0.014	0.018	0.009	0.019
MAPE	0.66%	0.74%	0.50%	0.74%

Table B.1: Error Measurements for each model used to forecast Traffic Volume[36]

[36] Mean error (ME) was included in appendices B and C to highlight each models' under/over-forecasting characteristics. (Positive ME indicates under-forecasting and vis versa)

Gas Price

	Model Fit		Internal Forecasting	
	Double Exp. Smoothing	Decomposition	Double Exp. Smoothing	Decomposition
ME	(0.025)	(0.026)	0.142	0.082
MSE	0.407	0.362	0.149	0.162
MAPE	3.29%	3.15%	2.26%	2.30%

Table B.2: Error measurements for each model used to forecast Gas Prices

Appendix C: Model Diagnostics

Traffic Volume

The diagnostics results for the traffic volume forecasting models. Plotting both models' fits against their residuals revealed no observable pattern to indicate non-constant variance or non-linearity. Homoscedasticity and linearity were further confirmed by performing a Score Test[37] on each model, both of which passed. Runs Tests[38] revealed that both models' data are not independent, but Sign Tests[39] confirmed the model data to be stationary. This is consistent with the reality that today's traffic volume is dependent on the previous period's. Finally, normality of the residuals of both models were checked using the Shapiro-Wilk Test[40]. The decomposition model failed the test for normality, while the multiple regression model passed, but required the removal of a single outlier for the residuals.

	Traffic Volume	
	Model Fit	
	Benchmark	Decomposition
ME	2.3E-15	-0.005
MSE	0.014	0.018
MAPE	0.007	0.007
Score Test	Passed	Passed
Shapiro-Wilk Test	Passed	Failed
Runs Test	Passed	Passed
Sign Test	Passed	Passed

Table C.1: Diagnostics for models used to forecast Traffic Volume

[37] Checks a model for constant variance by assessing the "goodness of fit" of the model's fitted values regressed against its variance. The test statistic is derived from the regression's MSE/2 and confirms or denies a hypothesis by comparing to a chi-squared critical value.
[38] A non-parametric statistical test that checks a randomness hypothesis for two datasets. It's used to test the hypothesis that the two datasets are mutually exclusive or independent.
[39] (Keller 2009, page 776)
[40] A test to determine if a dataset is normally distributed. The null hypothesis assumes normality, the resulting p-value must be larger than the desired alpha to confirm normality.

Gas Price

The diagnostics results for the gas price index forecasting models. The models' residuals were plotted against their fits to reveal that the variance was non-constant and non-linear. This was verified by checking each model with a Score Test, both of which failed. Histograms of the residuals revealed both models' residuals looked to be normally distributed. Runs and Sign Tests revealed that the data was neither independent nor stationary. This is in line with the assumption that the value of this period's GPI is dependent on the value of the previous period's GPI. Finally, normality of each model's residuals was checked. Shaprio-Wilk Tests revealed that both models' fits resulted in non-normal residuals.

	Gas Price Index	
	Model Fit	
	DES	Decomposition
ME	(0.025)	(0.026)
MSE	0.407	0.362
MAPE	0.033	0.032
Score Test	Failed	Failed
Shapiro-Wilk Test	Failed	Failed
Runs Test	Failed	Failed
Sign Test	Failed	Failed

Table C.2: Diagnostics for models used to forecast Gas Prices

Appendix D: Forecasting[41]

Traffic Volume Internal Forecast using Benchmarking with Multiple Regression

Model Structure:

$$\widehat{TrafficVolume}_i = -1.36 + 0.86 * Lag1TrafficVolume - 0.0015 * GasPriceIndex + 1.46E-8 * U.S.WorkingPopulation + 1.80E-7 * TotalAutoSales - 0.012 * LTGovernmentBondYield + 0.0088 * CPI - 0.0033 * TimeIndex + 1.30 * Mar + 0.42 * Feb \ldots - 0.069 * Jan$$

Applying Benchmarks (example for March, 2016):

$$\widehat{TrafficVolume}_{March,2016} = -1.36 + 0.86 * (15.13) - 0.0015 * (273.62) + 1.46E-8 * (205{,}208{,}000) + 1.80E-7 * (710{,}817) - 0.012 * (2.27) + 0.0088 * (238.65) - 0.0033 * (470) + 1.30 * (1)$$

$$\widehat{TrafficVolume}_{March,2016} = 16.25 \rightarrow 16.25^2 = 263.91\ Billions\ of\ Miles$$[42]

[41] All forecast calculations are explained in this section except the use of Double Exponential Smoothing for the internal forecast of Gas Price Index. This model was completely estimated using Minitab statistical software.
[42] The internal forecast of 16.25 is the transformed value. Squaring the value provides the forecast in Billions of Miles.

Internal Forecast Compared to observed value:

$$TrafficVolume_{March,2016} = 269.71\ Billions\ of\ Miles$$

Traffic Volume Internal Forecast using Decomposition Model

Model Structure:

$$\widehat{TrafficVolume}_{t+k} = \hat{T}_{t+k} * \hat{S}_{t+k} * \hat{C}_{t+k}$$
$$= (11.19 + 0.012 * TimeIndex) * SeasonalIndex * CyclicalComponentForecast$$

Example for March, 2016 *(time index = 470)*:

$$\widehat{TrafficVolume}_{March,2016} = (11.19 + 0.012 * 470) * (1.00098) * (0.96)$$ [43]

$$\widehat{TrafficVolume}_{March,2016} = 16.26 \rightarrow 16.26^2 = 264.26\ Billions\ of\ Miles$$

Internal Forecast Compared to observed value:

$$TrafficVolume_{March,2016} = 269.71\ Billions\ of\ Miles$$

Traffic Volume Future Forecast using Benchmarking with Multiple Regression

Model Structure:

$$\widehat{TrafficVolume}_i = -1.65 + 0.86 * Lag1TrafficVolume - 0.0014 * GasPriceIndex + 1.59E-8 * U.S.WorkingPopulation + 1.92E-7 * TotalAutoSales - 0.013 * LTGovernmentBondYield + 0.0098 * CPI - 0.0041 * TimeIndex + 1.31 * Mar + 0.41 * Feb \ldots - 0.074 * Jan$$

Applying Benchmarks. Example for March, 2017 (*time index = 482)*:

$$\widehat{TrafficVolume}_{March,2017} = -1.65 + 0.86 * (15.28) - 0.0014 * (257.78) + 1.59E-8 * (205{,}903{,}000) + 1.92E-7 * (707{,}086) - 0.013 * (2.03) + 0.0098 * (243.60) - 0.0041 * (482) + 1.31 * (1)$$

$$\widehat{TrafficVolume}_{March,2017} = 16.31 \rightarrow 16.31^2 = 266.03\ Billions\ of\ Miles$$

Gas Price Internal Forecast using Decomposition Model

Model Structure:

$$\widehat{GasPriceIndex}_{t+k} = \hat{T}_{t+k} * \hat{S}_{t+k} * \hat{C}_{t+k}$$
$$= (7.11 + 0.019 * TimeIndex) * SeasonalIndex * CyclicalComponentForecast$$

[43] Decomposition Model's Trend Component = *11.19 + 0.012*TimeIndex*, Seasonal Index for March = 1.00098, Cyclical Component was estimated using Double Exponential Smoothing (0.96).

Example for March, 2016 *(time index = 470)*:

$$\widehat{GasPriceIndex}_{March,2016} = (7.11 + 0.019 * 470) * (0.986) * (0.868)$$

$$\widehat{GasPriceIndex}_{March,2016} = 13.72 \rightarrow 13.72^2 = 188.30$$

$$\widehat{GasPrice}_{March,2016} = \frac{188.30}{100} * \$1.19 = \$2.24$$

Internal Forecast Compared to observed value:

$$GasPriceIndex_{March,2016} = 170.36$$

$$GasPrice_{March,2016} = \$2.03$$

Gas Price Future Forecast using Decomposition Model

Model Structure:

$$\begin{aligned}\widehat{GasPriceIndex}_{t+k} &= \hat{T}_{t+k} * \hat{S}_{t+k} * \hat{C}_{t+k} \\ &= (7.21 + 0.018 * TimeIndex) * SeasonalIndex * CyclicalComponentForecast\end{aligned}$$

Example for March, 2017 *(time index = 482)*:

$$\widehat{GasPriceIndex}_{March,2016} = (7.21 + 0.018 * 482) * (0.985) * (0.879)$$

$$\widehat{GasPriceIndex}_{March,2016} = 13.90 \rightarrow 13.90^2 = 193.29$$

$$\widehat{GasPrice}_{March,2016} = \frac{193.29}{100} * \$1.19 = \$2.30$$

References

Bureau of Economic Analysis. https://www.bea.gov/. Accessed 01 May, 2017.

DeLurgio, S. (1998). Forecasting Principles and Applications. Irwin/McGraw-Hill.

Digital Raves. Traffic Sources, Offers They Work Best With. https://digitalraves.com/. Accessed 01 June, 2017.

Federal Reserve Bank of St. Louis. FRED Economic Data. https://fred.stlouisfed.org/series/HSN1F. Accessed 01 May, 2017.

Keller (2009). Statistics for Management and Economics, 10th edition. Cengage Learning.

"The U.S. Petroleum Industry: Statistics, Definitions | NACS Online ...." http://www.nacsonline.com/YourBusiness/FuelsReports/GasPrices_2013/Pages/Statistics-and-Definitions.aspx. Accessed 6 Jun. 2017.