AVERAGE USE FORECASTING MODEL

1. The purpose of this evidence is to present the forecasting methodology used to forecast average use for Rate 1 revenue class 20 and Rate 6 revenue classes 12, 48, and 73[1]. Rate 1 is the Company's residential rate class while Rate 6 is the Company's small apartment, commercial, and industrial rate class. Revenue class 20 is forecast to comprise 86% of Rate 1 volumes while revenue classes 12, 48, and 73 are forecast to collectively comprise 93% of Rate 6 volumes in 2017. The forecasting methodology for the other revenue classes in Rate 1 and Rate 6 are very similar to the models presented in this exhibit. The evidence validates that the Company's models continue to be accurate predictors of average use.

2. The Company moved to a more objective forecasting methodology starting in the 2001 Budget year in order to address the Board's concern with the systemic bias attributed to the grassroots forecasting process. This forecasting methodology removes systemic or subjective bias by developing regression models to forecast average use for the Company's Rate 1 general service customers and Rate 6 general service customers. This econometric methodology has been in place since 2001, the forecasts of which have been accepted in settlement proposals and Board decisions since. As shown in Tables 1 to 3, 5, and 8, the models exhibit a high R^2 and low Root Mean Squared Percentage Error ("RMSPE") indicating that each of the regression models is a good predictor of average use.

[1] Rate 1 is comprised of: revenue class 10 - residential heating, revenue class 20 - residential space heating and water heating, revenue class 50 - space heating, water heating and pool heating, revenue class 60 – residential general service and revenue class 61 – residential water heating. Rate 6 is comprised of: revenue class 12 – apartment heating and other uses, revenue class 48 commercial heating and other uses, revenue class 73 industrial heating and other uses, revenue class 79 commercial general service, revenue class 83 – industrial general service, revenue class 86 – apartment general service, revenue class 90 – commercial air conditioning and space heating.

3. The year-over-year growth rates in average use for all revenue classes are used as the basis for the average use forecast for Rate 1 and Rate 6 shown at Exhibit C1, Tab 2, Schedule 1 Appendix A. Factors influencing overall average use include new customers (both new construction and replacement customers), the timing of new customer additions to the system, rate migration, gas prices, economic conditions, other external policy changes (e.g., Building Code) , and the Company's DSM programs. While average use changes for Rate 1 are fairly reflective of regression model results because of the homogenous nature of customers within this class, modeled Rate 6 average uses may be adjusted to account for known rate migration or specific changes in usage patterns for customers within this class. Please refer to Exhibit C1, Tab 2, Schedule 1 for a detailed explanation of the derivation of the Company's gas volume budget.

4. Average use is defined as gas volume per unlock customer. The econometric models presented here utilize historical data and relationships to estimate driver variable impacts and derive a top down forecast of average use. The models presented in the exhibit incorporate updated driver variables and historical data obtained from federal and provincial statistical agencies and the Company's database. Maintaining an econometric model is an ongoing process; consequently, the models must be monitored and refined to ensure they are valid and produce accurate forecasts of general service average use.

<u>Error Correction Model</u>

5. The Company uses Error Correction Models ("ECM") to forecast the average use for Rate 1 and Rate 6. The ECM method and two step estimation procedure are

described more fully in Engle and Granger (1987).[2] The ECM uses the concept of cointegration or long-run association between variables.

6. In other words, variables hypothesized to be linked by some theoretical economic relationship should not diverge from each other in the long run. Such variables may drift apart in the short run; however, if they were to diverge without bound, an equilibrium relationship among such variables could not be said to exist. The ECM methodology has been used extensively in the energy field for modeling electricity sales[3] and natural gas prices[4].

7. The major difference between the ECM approach and the standard dynamic single-equation model is the ECM approach explicitly takes into account both long-run equilibrium and short-run dynamic relationships in the determination of average use. It is known that economic theory can provide useful information about the variables relevant in the long-run. However, it is relatively silent on the short-run dynamics between variables. The ECM approach allows the historical data to determine the lag structures and short run dynamics.

8. The estimated models are used to generate a normalized forecast of average use. The main purpose of the normalized forecast is to derive average use such that the weather impact has been taken out. Using the estimated coefficients, weather normalized average use data are obtained by replacing actual degree days in the model with proposed degree days for 2017 for every year so that year-to-year

---

[2] Engle, R.F. and Granger, C.W.J (1987), "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, Vol. 55, No.2.
[3] Engle, R.F., Granger, C.W.J. and Hallman, J.J. (1989), "Merging Short- and Long-Run Forecasts: An Application to Monthly Electricity Sales Forecasting," *Journal of Econometrics*, Vol.40.
[4] Bopp, A.E. (1990), "An Analytical Approach to Forecasting Natural Gas Prices," *AGA Forecasting Review*: American Gas Association.

Witnesses: H. Sayyan
M. Suarez

percentage changes reflect the pure average use trend by eliminating weather variability.

Average Use Forecasting Methodology

9. The model's specification is based on an objective criterion: to minimize both in-sample and out-of-sample forecast error. The discrepancy between actual average use and the model's forecast can be segregated into three major sources of uncertainty: (1) model specification, (2) forecast error from the driver variables used in the model, and (3) unexpected shocks or structural breaks. Sources (2) and (3) are not within the Company's control and will inevitably occur regardless of which forecasting methodology is adopted. Therefore the objective of the modeling procedure, described below, is to minimize the controllable source of error, the model's specification.

10. The main criteria for assessing the model's predictive ability is the model's forecast accuracy. A comparison of actual un-normalized average use versus the forecasts produced by the model is used to assess predictive ability. Forecast accuracy for 2017 is measured using both in-sample and out-of-sample Mean Percentage Error ("MPE") and RMSPE. In-sample, or ex-post, means that the estimated model incorporates the entire sample, in this case 1985 to 2015. Out-of-sample, or ex-ante, means that the model incorporates only a portion of the sample, in this case 1985 to 2013. Forecasts of average use are produced under both approaches and measured against actual average use from 2014 to 2015 quantitatively via MPE and RMSPE. A two year "hold out" sample is used to compute the out-of-sample forecast accuracy statistics since the forecasting horizon for volumetric budgeting purposes is two years.

11. Table 1 presents the forecast accuracy statistics for Rate 1 and Rate 6. The smaller the MPE and RMSPE, the better the model's forecast performance.

TABLE 1
FORECAST ERRORS - PERCENT VARIANCE & ROOT MEAN SQUARED PERCENTAGE ERROR

Col 1.	Col 2.	Col 3.
Forecast Error Method	Rate 1	Rate 6
In-Sample % Variance (2 Years)	-0.17%	0.21%
In-Sample RMSPE (2 Years)	0.70%	0.40%
Out-of-Sample % Variance (2 Years)	0.06%	-0.35%
Out-of-Sample RMSPE (2 Years)	0.76%	0.44%

$$MPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Forecast_i - Actual_i}{Actual_i}\right)$$

$$RMSPE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Forecast_i - Actual_i}{Actual_i}\right)^2}$$

12. Consistent with the settlement of Issue 1.1 in the RP-2000-0040 Settlement Agreement, Tables 2 and 3 report the results that the models would generate using actual data to allow parties to compare results to the prior year's forecast. Tables 2 and 3 show the results that the models would have produced had all actual driver values been available at the time the forecast was produced. The tables are not updated for 2004 since there are no Board approved average use forecasts for this particular test year. In order to compare the variance between actual and Board Approved average use on the same basis, the actual results for each year have been normalized to the corresponding Board Approved degree days for each

Witnesses: H. Sayyan
M. Suarez

respective test year. The results in Tables 2 and 3 show the regression model is a good predictor of general service average use.

TABLE 2

RATE 1 IN-SAMPLE FORECAST COMPARISON

Col 1.	Col 2.	Col 3.	Col 4.	Col 5.	Col 6.	Col 7.	Col 8.
Fiscal Year	Actual Normalized Average Use Per Customer	Board Approved Normalized Average Use Per Customer[1,3]	Variance Normalized Average Use Per Customer	% Variance Normalized Average Use Per Customer	Model's Normalized Average Use Per Customer[2]	Variance Normalized Average Use Per Customer	% Variance Normalized Average Use Per Customer
	(m3)	m(3)	(2-3)	100*((2-3)/3)	(m3)	(2-6)	100*((2-6)/6)
2001	3,014	3,044	(30)	-1.0%	3,022	(8)	-0.26%
2002	2,980	2,970	10	0.3%	2,963	17	0.57%
2003	2,877	2,892	(15)	-0.5%	2,897	(20)	-0.69%
2004	2,843	n/a	n/a	n/a	2,864	(21)	-0.73%
2005	2,890	2,953	(63)	-2.1%	2,929	(39)	-1.33%
2006	2,796	2,850	(54)	-1.9%	2,816	(20)	-0.71%
2007	2,726	2,687	39	1.5%	2,695	31	1.15%
2008	2,636	2,647	(11)	-0.4%	2,611	25	0.97%
2009	2,616	2,637	(21)	-0.8%	2,623	(6)	-0.24%
2010	2,579	2,622	(43)	-1.6%	2,550	29	1.15%
2011	2,594	2,643	(49)	-1.9%	2,607	(13)	-0.51%
2012	2,529	2,510	18	0.7%	2,528	1	0.02%
2013	2,547	2,568	(22)	-0.8%	2,517	30	1.18%
2014	2,475	2,433	41	1.7%	2,490	(15)	-0.60%
2015	2,427	2,419	9	0.4%	2,404	23	0.97%

[1]Board approved normalized average use from RP-2000-0040, RP-2001-0032, RP-2002-0133, RP-2003-0203, EB-2005-000, EB-2006-0034, EB-2007-0615, EB-2008-0219, EB-2009-0172, EB-2010-0146, EB-2011-0277, EB-2011-0354, EB-2012-0459 and EB-2014-0276 for 2001, 2002, 2003, 2005, 2006, 2007, 2008, 2009, 2010, 2011,2012, 2013, 2014 and 2015 respectively.

[2]Model's normalized average use is generated by running the model using actual data and driver variable information.

[3]There is no Board approved normalized average use for 2004.

Witnesses: H. Sayyan
M. Suarez

TABLE 3
RATE 6 IN-SAMPLE FORECAST COMPARISON

Col 1.	Col 2.	Col 3.	Col 4.	Col 5.	Col 6.	Col 7.	Col 8.
Fiscal Year	Actual Normalized Average Use Per Customer	Board Approved Normalized Average Use Per Customer[1,3]	Variance Normalized Average Use Per Customer	% Variance Normalized Average Use Per Customer	Model's Normalized Average Use Per Customer[2]	Variance Normalized Average Use Per Customer	% Variance Normalized Average Use Per Customer
	(m3)	m(3)	(2-3)	100*((2-3)/3)	(m3)	(2-6)	100*((2-6)/6)
2001	22,510	22,643	(133)	-0.6%	22,706	(196)	-0.86%
2002	22,097	22,125	(28)	-0.1%	21,957	140	0.64%
2003	21,593	21,685	(92)	-0.4%	21,613	(20)	-0.09%
2004	21,472	n/a	n/a	n/a	21,377	95	0.44%
2005	22,241	22,507	(266)	-1.2%	22,334	(93)	-0.42%
2006	22,272	21,999	273	1.2%	22,149	123	0.55%
2007	22,783	21,010	1773	8.4%	22,973	(190)	-0.83%
2008	24,869	24,204	665	2.7%	25,273	(404)	-1.60%
2009	27,654	28,165	(512)	-1.8%	27,875	(222)	-0.79%
2010	29,106	27,949	1157	4.1%	29,691	(585)	-1.97%
2011	29,471	28,029	1442	5.1%	30,240	(769)	-2.54%
2012	28,941	30,122	(1182)	-3.9%	28,634	307	1.07%
2013	29,203	29,878	(675)	-2.3%	28,756	447	1.56%
2014	28,634	28,383	251	0.9%	28,535	99	0.35%
2015	28,600	28,341	259	0.9%	28,375	225	0.79%

[1]Board approved normalized average use from RP-2000-0040, RP-2001-0032, RP-2002-0133, RP-2003-0203, EB-2005-000, EB-2006-0034, EB-2007-0615, EB-2008-0219, EB-2009-0172, EB-2010-0146, EB-2011-0277, EB-2011-0354, EB-2012-0459 and EB-2014-0276 for 2001, 2002, 2003, 2005, 2006, 2007, 2008, 2009, 2010, 2011,2012, 2013, 2014 and 2015 respectively.

[2]Model's normalized average use is generated by running the model using actual data and driver variable information.

[3]There is no Board approved normalized average use for 2004.

13. The primary goal of the average use forecast is to be accurate and objective. Ideally, the forecast error should be small in magnitude and distributed in a random fashion. Although the forecast errors in Tables 1, 2, and 3 are small in magnitude, forecast accuracy is conditional on driver variable forecast accuracy and the absence of any structural break between the historical period and the upcoming forecast period. Consequently, besides testing forecast accuracy, the models were subjected to a battery of diagnostic tests. These tests were run on the model to check for incorrect functional forms, parameter instability, structural breaks, omitted variables and randomness of residuals. Overall the models have been thoroughly tested and are statistically valid.

Witnesses: H. Sayyan
M. Suarez

The following diagnostic tests were run on each model[5] (results are shown in Tables 6 and 9):

Breusch-Godfrey Serial Correlation LM Test

This test is used to test for autocorrelation in the residuals. Autocorrelation occurs when disturbances in a regression equation are serially correlated. The test is set up as follows:

Null Hypothesis: No serial correlation

Alternative Hypothesis: Serial correlation

ARCH Test

This test is used to test for Autoregressive Conditional Heteroskedasticity ("ARCH"). ARCH occurs when the variance of disturbances in a regression equation are not constant and are serially correlated. The test is set up as follows:

Null Hypothesis: No ARCH

Alternative Hypothesis: ARCH

Chow Forecast Test

This test is used to test for stability of a regression model. A regression model is not stable if the estimated coefficients change (and consequently the model's predictions) when estimated over various sample ranges. The test is set up as follows:

Null Hypothesis: No structural change

Alternative Hypothesis: Structural change

[5] The Durbin-Watson test is not used since it is not valid when there are lagged dependent variables in a regression equation. The Durbin Watson test is biased toward the finding of no serial correlation if there are lagged values of the dependent variable in the regression equation.

Ramsey RESET Test

This is a general test which tests for omitted variables, incorrect functional form and correlation between the independent variables and disturbances. The test is set up as follows:

Null Hypothesis: Normally distributed disturbances (zero mean, constant variance)

Alternative Hypothesis: Non- normally distributed disturbances (non-zero mean, constant variance)

14. The following tables present the mnemonics used in the models (Tables 4 and 7), the regression equations for each model (Tables 5 and 8), and the diagnostic tests results run on the models (Tables 6 and 9). For the t tests in the regression equations shown at Tables 5 and 8, the p-values indicate the probability of obtaining a forecast at least as extreme as one that was actually observed, assuming that the null hypothesis (coefficient is not significant) is true. The p-value is compared to a significance level which is often 0.05 or 0.10, so that if its value is smaller, the null hypothesis is rejected at the 95% or 90% confidence level, respectively. The smaller the p-value, the more strongly the test rejects the null hypothesis, thereby supporting the statistical significance of the coefficient. In any instance where insignificant variables were retained within the models, it was for the purposes of (1) improving the significance of other coefficients or (2) optimizing forecast accuracy. In contrast, for the diagnostic test results shown in Tables 6 and 9, the null hypotheses tested are the desired outcomes. In each case, to support the null hypothesis, p-values in excess of 0.10 are preferred. Overall, diagnostic test results in Table 6 and 9 show that the models in Table 5 and 8 are statistically valid and no assumptions appear to be violated at the 95% confidence level.

Witnesses: H. Sayyan
M. Suarez

TABLE 4 - RATE 1 MODEL MNEMONICS

Mnemonic	Definition
C	Constant Term
LOG(X)	Logarithm of Variable X
DLOG(X)	$LOG(X_t) - LOG(X_{t-1})$, First Difference of Logarithm of Variable X
CDD, EDD, NDD	Balance Point Heating Degree Days for Central, Eastern and Niagara Weather Zones
MET20VINT	Vintage Variable for the Metro Region, Central Weather Zone
WES20VINT	Vintage Variable for the Western Region, Central Weather Zone
CEN20VINT	Vintage Variable for the Central Region, Central Weather Zone
NOR20VINT	Vintage Variable for the Northern Region, Central Weather Zone
ERC20VINT	Vintage Variable for the Eastern Weather Zone
NRC20VINT	Vintage Variable for the Niagara Weather Zone
REALCRCRPG	Real Residential Natural Gas Price for the Central Weather Zone
REALERCRPG	Real Residential Natural Gas Price for the Eastern Weather Zone
DUM2008-DUM2009-DUM2010	Dummy Variables for Recession Impact
CENTEMP	Central Weather Zone Employment
ECM_Region	Error Correction Term for Each Region

Witnesses: H. Sayyan
M. Suarez

TABLE 5 - RATE 1 REVENUE CLASS 20 REGRESSION EQUATIONS

Metro Region - Central Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	2.64	10.11	0.00
LOG(CDD)	0.70	21.35	0.00
LOG(REALCRCRPG)	-0.03	-1.57	0.13
LOG(MET20VINT)	0.83	11.25	0.00
DUM2008	-0.04	-3.72	0.00
DUM2010	-0.04	-2.84	0.01
R-squared	0.99		
Adjusted R-squared	0.99		
S.E. of regression	0.01		
F-statistic	540.86		0.00

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.00	0.04	0.97
DLOG(CDD)	0.75	30.56	0.00
DLOG(MET20VINT)	0.99	1.86	0.07
DUM2008	-0.01	-1.45	0.16
ECM_MET20(-1)	-0.34	-1.62	0.12
R-squared	0.98		
Adjusted R-squared	0.98		
S.E. of regression	0.01		
F-statistic	284.16		0.00

Western Region - Central Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	1.16	1.57	0.13
LOG(CDD)	0.69	26.49	0.00
LOG(REALCRCRPG)	-0.09	-5.74	0.00
LOG(WES20VINT)	0.42	5.66	0.00
LOG(CENTEMP)	0.17	2.00	0.06
DUM2008	-0.04	-3.67	0.00
DUM2010	-0.05	-4.28	0.00
R-squared	0.99		
Adjusted R-squared	0.99		
S.E. of regression	0.01		
F-statistic	629.90		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	-0.005	-2.17	0.04
DLOG(CDD)	0.72	37.68	0.00
DLOG(REALCRCRPG)	-0.08	-4.18	0.00
DUM2008	-0.01	-2.45	0.02
ECM_WES20(-1)	-0.64	-2.95	0.01
R-squared	0.98		
Adjusted R-squared	0.98		
S.E. of regression	0.01		
F-statistic	395.36		0.000

Central Region - Central Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.78	1.01	0.32
LOG(CDD)	0.70	20.92	0.00
LOG(REALCRCRPG)	-0.03	-1.85	0.08
LOG(CEN20VINT)	0.54	8.29	0.00
LOG(CENTEMP)	0.21	2.40	0.02
DUM2008	-0.05	-4.29	0.00
R-squared	0.99		
Adjusted R-squared	0.99		
S.E. of regression	0.01		
F-statistic	553.60		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.00	0.11	0.91
DLOG(CDD)	0.70	28.35	0.00
DLOG(REALCRCRPG)	-0.04	-1.63	0.12
DUM2008	-0.01	-1.43	0.17
DLOG(CEN20VINT)	0.37	1.81	0.08
ECM_CEN20(-1)	-0.92	-4.57	0.00
R-squared	0.98		
Adjusted R-squared	0.97		
S.E. of regression	0.01		
F-statistic	197.01		0.000

TABLE 5 CONTINUED - RATE 1 REVENUE CLASS 20 REGRESSION EQUATIONS

Northern Region - Central Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.91	1.10	0.28
LOG(CDD)	0.68	23.36	0.00
LOG(REALCRCRPG)	-0.08	-4.73	0.00
LOG(NOR20VINT)	0.47	7.55	0.00
LOG(CENTEMP)	0.22	2.25	0.03
DUM2009	-0.05	-4.46	0.00
R-squared	0.99		
Adjusted R-squared	0.99		
S.E. of regression	0.01		
F-statistic	838.60		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.00	-0.55	0.58
DLOG(CDD)	0.68	30.54	0.00
DLOG(REALCRCRPG)	-0.06	-2.79	0.01
DLOG(NOR20VINT)	0.26	1.51	0.14
ECM_NOR20(-1)	-1.03	-4.69	0.00
R-squared	0.98		
Adjusted R-squared	0.97		
S.E. of regression	0.01		
F-statistic	264.24		0.000

Eastern Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	1.69	5.56	0.00
LOG(EDD)	0.77	20.72	0.00
LOG(REALERCRPG)	-0.03	-1.89	0.07
LOG(ERC20VINT)	0.43	18.51	0.00
DUM2009	-0.04	-3.77	0.00
R-squared	0.99		
Adjusted R-squared	0.99		
S.E. of regression	0.01		
F-statistic	845.32		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	-0.01	-2.96	0.01
DLOG(EDD)	0.78	25.34	0.00
DLOG(REALERCRPG)	-0.02	-0.64	0.53
DUM2008	-0.01	-1.78	0.09
ECM_ERC20(-1)	-1.07	-4.67	0.00
R-squared	0.97		
Adjusted R-squared	0.97		
S.E. of regression	0.01		
F-statistic	209.91		0.000

Niagara Weather Zone

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	2.19	6.38	0.00
LOG(NDD)	0.74	17.06	0.00
LOG(NRC20VINT)	1.23	19.09	0.00
DUM2008	-0.05	-3.79	0.00
R-squared	0.99		
Adjusted R-squared	0.98		
S.E. of regression	0.02		
F-statistic	645.25		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	-0.01	-4.43	0.00
DLOG(NDD)	0.74	25.10	0.00
ECM_NRC20(-1)	-0.63	-3.60	0.00
R-squared	0.96		
Adjusted R-squared	0.96		
S.E. of regression	0.02		
F-statistic	350.81		0.000

Witnesses: H. Sayyan
M. Suarez

TABLE 6 - RATE 1

Model Diagnostic Tests

Col 1.	Col 2.	Col 3.	Col 4.	Col 5.	Col 6.	Col 7.	Col 8.
Test		Metro Region	Western Region	Central Region	Northern Region	Eastern Weather Zone	Niagara Weather Zone
Breusch-Godfrey Serial Correlation LM Test	Test Statistic	1.33	0.24	0.45	0.40	0.07	0.96
	P Value	0.25	0.62	0.50	0.53	0.79	0.33
ARCH Test	Test Statistic	0.16	1.34	0.64	0.15	0.05	0.87
	P Value	0.69	0.25	0.42	0.70	0.83	0.35
Chow Forecast Test: Forecast from 2015 to 2015	Test Statistic	0.70	0.67	0.87	2.07	0.22	0.09
	P Value	0.41	0.42	0.36	0.15	0.64	0.77
Ramsey RESET Test	Test Statistic	1.14	1.67	0.61	0.05	1.91	0.04
	P Value	0.30	0.21	0.44	0.83	0.18	0.84

TABLE 6 - RATE 1

Model Diagnostic Tests

Col 1.	Col 2.	Col 3.	Col 4.	Col 5.	Col 6.	Col 7.	Col 8.
Test		Metro Region	Western Region	Central Region	Northern Region	Eastern Weather Zone	Niagara Weather Zone
Breusch-Godfrey Serial Correlation LM Test	Test Statistic	0.83	0.29	0.49	0.27	0.07	0.75
	P Value	0.36	0.59	0.48	0.60	0.79	0.39
ARCH Test	Test Statistic	0.24	3.15	0.62	0.02	0.00	0.80
	P Value	0.62	0.08	0.43	0.89	0.95	0.37
Chow Forecast Test: Forecast from 2014 to 2014	Test Statistic	0.44	0.80	0.03	0.31	0.74	1.05
	P Value	0.51	0.38	0.86	0.58	0.40	0.31
Ramsey RESET Test	Test Statistic	0.77	1.12	0.25	0.10	1.77	0.01
	P Value	0.39	0.30	0.62	0.75	0.20	0.91

Witnesses: H. Sayyan
M. Suarez

TABLE 7 - RATE 6 MODEL MNEMONICS

Mnemonic	Definition
C	Constant Term
LOG(X)	Logarithm of Variable X
DLOG(X)	LOG(X_t) - LOG(X_{t-1}), First Difference of Logarithm of Variable X
CDD, EDD, NDD	Balance Point Heating Degree Days for Central, Eastern and Niagara Weather Zones
CENTEMP	Central Weather Zone Employment
EASTEMP	Eastern Weather Zone Employment
NIAGEMP	Niagara Weather Zone Employment
REALERCCPG	Real Commercial Gas Price for the Eastern Weather Zone
REALNRCCPG	Real Natural Gas Price for the Niagara Weather Zone
ONTGDP	Ontario Real Gross Domestic Product
CRCCOMVAC	GTA Commercial Vacancy Rate
TIME	Time Trend
DUMRegion	Dummy Variable for Migration Impact
DUMXXXX	Dummy Variable for the Break in the Year XXXX
AR(p)	pth-order Autoregressive Process Term
ECM_Region	Error Correction Term for Each Region

Witnesses: H. Sayyan
M. Suarez

TABLE 8 - RATE 6 REVENUE CLASS 12 REGRESSION EQUATIONS

Central Revenue Class 12 (Apartment)

Single Equation Model

Variable	Coefficient	t-Statistic	p-Value
C	2.27	1.91	0.07
LOG(CDD)	0.60	7.86	0.00
LOG(CENTEMP)	0.56	4.98	0.00
DUM1996	-0.09	-4.10	0.00
DUMCRC12	0.21	5.45	0.00
AR(1)	0.49	2.43	0.02
R-squared	0.97		
Adjusted R-squared	0.97		
S.E. of regression	0.03		
F-statistic	181.731		0.000

Eastern Revenue Class 12 (Apartment)

Single Equation Model

Variable	Coefficient	t-Statistic	p-Value
C	3.26	2.30	0.03
LOG(EDD)	0.57	8.72	0.00
LOG(TIME)	-0.05	-4.00	0.00
DUMERC12	0.26	10.23	0.00
DUM2011	-0.13	-5.35	0.00
LOG(REALERCCPG)	-0.14	-3.24	0.00
LOG(EASTEMP)	0.47	2.55	0.02
DUM2014	0.05	2.32	0.03
R-squared	0.97		
Adjusted R-squared	0.97		
S.E. of regression	0.02		
F-statistic	120.79		0.000

Niagara Revenue Class 12 (Apartment)

Single Equation Model

Variable	Coefficient	t-Statistic	p-Value
C	4.64	4.02	0.00
LOG(NDD)	0.54	8.29	0.00
LOG(TIME)	-0.03	-2.86	0.01
LOG(NIAGEMP)	0.34	2.15	0.04
LOG(REALNRCCPG)	-0.06	-1.74	0.10
DUMNRC12	-0.07	-4.23	0.00
DUM2011	-0.10	-4.50	0.00
AR(1)	-0.34	-1.57	0.13
R-squared	0.90		
Adjusted R-squared	0.87		
S.E. of regression	0.03		
F-statistic	29.49		0.000

Witnesses: H. Sayyan
M. Suarez

TABLE 8 CONTINUED - RATE 6 REVENUE CLASS 48 REGRESSION EQUATIONS

Central Revenue Class 48 (Commercial)

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.23	0.28	0.78
LOG(CDD)	0.84	17.52	0.00
LOG(TIME)	-0.12	-9.14	0.00
LOG(CRCCOMVAC)	-0.07	-4.36	0.00
LOG(ONTGDP)	0.26	4.39	0.00
DUMCRC48	0.11	10.78	0.00
R-squared	0.97		
Adjusted R-squared	0.96		
S.E. of regression	0.02		
F-statistic	153.58		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.00	0.14	0.89
DLOG(CDD)	0.82	31.63	0.00
DLOG(TIME)	-0.06	-3.07	0.01
DLOG(CRCCOMVAC)	-0.05	-3.88	0.00
DUMCRC48	0.02	2.41	0.02
ECM_CRC48(-1)	-0.81	-5.36	0.00
R-squared	0.98		
Adjusted R-squared	0.97		
S.E. of regression	0.01		
F-statistic	203.68		0.000

Eastern Revenue Class 48 (Commercial)

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	1.73	1.75	0.09
LOG(EDD)	0.74	10.28	0.00
LOG(TIME)	-0.16	-12.00	0.00
LOG(ONTGDP)	0.19	3.44	0.00
DUMERC48	0.10	4.74	0.00
DUM2010	0.11	5.40	0.00
R-squared	0.96		
Adjusted R-squared	0.95		
S.E. of regression	0.03		
F-statistic	111.59		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.01	1.19	0.24
DLOG(EDD)	0.69	8.83	0.00
DLOG(TIME)	-0.13	-2.79	0.01
ECM_ERC48(-1)	-0.73	-2.52	0.02
R-squared	0.80		
Adjusted R-squared	0.77		
S.E. of regression	0.03		
F-statistic	34.03		0.000

Niagara Revenue Class 48 (Commercial)

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	-1.61	-1.19	0.25
LOG(NDD)	0.73	14.86	0.00
LOG(TIME)	-0.10	-5.33	0.00
LOG(REALNRCCPG)	-0.18	-5.04	0.00
LOG(ONTGDP)	0.43	4.39	0.00
DUMNRC48	0.11	4.75	0.00
DUM2010	-0.11	-3.80	0.00
R-squared	0.94		
Adjusted R-squared	0.92		
S.E. of regression	0.02		
F-statistic	62.12		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	0.00	-0.37	0.72
DLOG(NDD)	0.75	13.62	0.00
DUMNRC48	0.11	3.42	0.00
DUM2010	-0.12	-3.44	0.00
DLOG(REALNRCCPG)	-0.07	-1.54	0.14
ECM_NRC48(-1)	-0.87	-2.72	0.01
R-squared	0.91		
Adjusted R-squared	0.89		
S.E. of regression	0.03		
F-statistic	50.39		0.000

Witnesses: H. Sayyan
M. Suarez

TABLE 8 CONTINUED - RATE 6 REVENUE CLASS 73 REGRESSION EQUATIONS

Central Revenue Class 73 (Industrial)

Long Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	1.63	0.58	0.56
LOG(CDD)	0.47	2.68	0.01
LOG(TIME)	-0.15	-3.78	0.00
LOG(ONTGDP)	0.45	2.73	0.01
DUMCRC73	0.52	12.64	0.00
R-squared	0.92		
Adjusted R-squared	0.90		
S.E. of regression	0.07		
F-statistic	71.84		0.000

Short Run Equation

Variable	Coefficient	t-Statistic	p-Value
C	-0.02	-2.21	0.04
DLOG(CDD)	0.55	8.84	0.00
DLOG(ONTGDP)	0.71	2.37	0.03
DUMCRC73	0.24	6.51	0.00
DUM2009	-0.18	-4.84	0.00
ECM_CRC73(-1)	-0.62	-6.35	0.00
R-squared	0.87		
Adjusted R-squared	0.84		
S.E. of regression	0.03		
F-statistic	32.34		0.000

Eastern Revenue Class 73 (Industrial)

Single Equation Model

Variable	Coefficient	t-Statistic	p-Value
C	-319,256.00	-3.67	0.00
EDD	34.36	2.59	0.02
DUM2003	57,733.93	3.37	0.00
DUM2004	-160,946.50	-7.23	0.00
DUMERC73	122,447.20	11.50	0.00
EASTEMP	671.43	4.91	0.00
TIME	-5,567.67	-5.07	0.00
R-squared	0.95		
Adjusted R-squared	0.94		
S.E. of regression	15,508.71		
F-statistic	73.99		0.000

Niagara Revenue Class 73 (Industrial)

Single Equation Model

Variable	Coefficient	t-Statistic	p-Value
C	-1.59	-0.44	0.66
LOG(NDD)	0.80	3.63	0.00
DUM2002	-0.36	-4.17	0.00
DUMNRC73	0.47	4.44	0.00
DUM2010	0.39	3.66	0.00
LOG(NIAGEMP)	1.26	2.44	0.02
AR(1)	0.66	3.75	0.00
R-squared	0.97		
Adjusted R-squared	0.96		
S.E. of regression	0.10		
F-statistic	106.18		0.000

TABLE 9-RATE 6
Model Diagnostic Tests

Col 1.	Col 2.	Col 3.	Col 4.	Col 5.	Col 6.	Col 7.	Col 8.	Col 9.	Col 10.	Col 11.
		Revenue Class 12 (Apartment) Model Diagnostic Tests			**Revenue Class 48 (Commercial) Model Diagnostic Tests**			**Revenue Class 73 (Industrial) Model Diagnostic Tests**		
Test		Central Weather Zone	Eastern Weather Zone	Niagara Weather Zone	Central Weather Zone	Eastern Weather Zone	Niagara Weather Zone	Central Weather Zone	Eastern Weather Zone	Niagara Weather Zone
Breusch-Godfrey Serial Correlation LM Test	Test Statistic	2.24	1.18	0.27	1.45	1.96	1.19	1.64	0.06	2.34
	P Value	0.13	0.28	0.60	0.23	0.16	0.27	0.20	0.80	0.13
ARCH Test	Test Statistic	0.10	1.42	1.22	0.66	0.04	2.33	0.74	0.00	1.96
	P Value	0.75	0.23	0.27	0.42	0.85	0.13	0.39	0.98	0.16
Chow Forecast Test: Forecast from 2015 to 2015	Test Statistic	0.01	0.09	1.94	0.16	1.18	0.18	0.24	1.69	0.00
	P Value	0.93	0.76	0.18	0.69	0.29	0.68	0.63	0.20	0.96
Ramsey RESET Test	Test Statistic	0.95	2.92	0.39	1.07	0.29	1.00	1.44	1.34	2.50
	P Value	0.34	0.10	0.54	0.31	0.59	0.33	0.24	0.26	0.13

15. Major driver variables in the models are balance point heating degree days adjusted for billing cycles, vintage, a time trend, real natural gas prices and economic variables. Driver variable assumptions are shown in the Economic Outlook at Exhibit C2, Tab 1, Schedule 1.

16. Natural gas prices have an important impact on average use. Sharp increases typically have two effects. First, they influence customer's fuel use habits, for example, the lowering of thermostat settings. Second, price increases likely factor in customers' decision-making around the purchase of more efficient furnaces and other appliances. In addition, homeowners may also respond by retrofitting older residences in order to reduce energy consumption. In the models, real natural gas prices are used. The Consumer Price Index ("CPI") is used to convert nominal gas prices to real gas prices. Nominal energy price forecast for 2017 is based on the consensus Henry Hub price forecast produced in April 2016.

17. A linear time trend is used as a proxy measure for energy conservation. However, a linear time trend only reflects constant annual changes in appliance efficiency; it will not be able to reflect the time-varying impact of new residential construction on appliance efficiency. Consequently, a vintage variable serves as either a supplementary or complementary variable to the time trend in the model.

18. The vintage variable (for revenue class 20 only) is employed as a proxy measure of gas space heating and gas water heating efficiency gains and residential thermal efficiency. Newer homes with improved thermal envelope characteristics and older homes adding insulation and storm windows/doors reduce the typical amount of gas needed for space heating. Residential thermal efficiency will continue to improve as newer, better-insulated residences account for a larger portion of the

housing stock. The vintage variable captures the impact of both furnace efficiency and new home thermal efficiency on average use.

19. Vintage is defined as the calendar year in which the customer became a customer (new gas service main date) and is not based on the age of the building. This data includes both new construction and conversion customer additions. As space heating efficiency gains have a greater impact on average use than thermal improvements to homes, customers by vintage is a better variable than age of the building in terms of explaining the percentage decline in residential average use.

20. An illustration of the vintage ratio for 1992 follows:

$$V_{1992} = \frac{\sum_{y=1987}^{1991} V_y}{\sum_{yy=1987}^{1992} V_{yy}}$$ where V denotes vintage.

21. Calendar 1992 is used as the reference year for the vintage ratio since the Energy Efficiency Act prohibited selling of the conventional low-efficiency furnace in January 1992.[6] Consequently, this ratio will capture the increasing market share of both mid-efficiency and high-efficiency furnaces at the expense of declining market share of conventional furnaces over time. Generally, regions with stronger new construction additions experience a sharper decline in the ratio than established regions like Metro. As more new customers are added to the revenue class the declining ratio leads to lower average use over time.

[6] During the 1970s natural gas furnaces averaged about 65% Annual Fuel Utilization Efficiency ("AFUE"). The Energy Efficiency Act imposed 78% AFUE as a minimum for gas furnaces manufactured after January 1, 1992.

Witnesses: H. Sayyan
M. Suarez

22. Thus the sign of this variable's coefficient is positive.

23. Economic variables such as employment, vacancy rates, and gross domestic product can impact demand for new gas appliances as well as impact demand for natural gas for space heating and manufacturing processes. Stronger employment and demand for products both domestically and abroad will generally increase natural gas demand.

Risks to the Forecast

24. The impact of customer mix on average use is not static and changes over time. New customers may have different gas use characteristics than existing customers and may be influenced by builder specifications for inclusion/exclusion of new gas appliances. Thus, aggregate average use will be affected even if customers take no actions that could affect their average use. Advances in the future penetration of gas appliances above historical penetration levels implicit in the model could result in increased average use. Conversely, builder specification of non-gas water and/or space heating equipment represents a risk to the forecast as it could result in lower gas consumption than forecast.

25. The impact of Cap and Trade is not explicitly modeled within the average use equations as it is unclear how including Greenhouse Gas emissions compliance costs as part of distribution charges will be perceived by customers such that this would drive behavioral changes in consumption. Given that Cap and Trade will increase the overall price of natural gas service for customers, the price difference may drive lower consumption than forecast.

26. New Building Code requirements come into effect in January 2017 that could potentially result in lower average uses than forecast. The potential reductions in average use are largely dependent on the installation options or compliance packages implemented by designers and builders, as well as when permits were applied for. While savings are difficult to model, it is estimated that the impacts will be minimal as forecast average uses are relatively close to the target reduction.

27. The use of more efficient water heaters across the franchise area and/or the loss of natural gas water heating to other fuels could result in a permanent decrease in baseload usage and natural gas consumption relative to the forecast.

28. Gas consumption for space heating is very sensitive to thermostat settings. Customers may set their thermostats lower under extremely warm weather like that experienced in 1998, 2001, 2006, and most recently in 2012.

29. Economic activity can impact both demand for appliances and natural gas. If the economy slows more significantly and natural gas prices are higher than indicated in the Economic Outlook (Exhibit C2, Tab 1, Schedule 1), average use will decline further.

30. A structural break in the historical estimated relationship between average use and the driver variables will increase forecast risk as will forecast uncertainty in the driver variables.

Conclusion

31. The model employed by the Company passes a battery of statistical tests and is valid given current and historical information. Continual evaluation and testing is required, as new information becomes available. The model has been estimated

over volatile periods in history – recent years of unexpected warm and cold weather, historically high energy prices, and increased energy price volatility. In light of these volatile economic and weather conditions, continuous model evaluation ensures that ongoing impacts in the relationship of average use and its driver variables is captured to produce the most accurate and objective forecast as possible.