BEFORE THE PUBLIC SERVICE COMMISSION OF UTAH

IN THE MATTER OF THE JOINT APPLICATION OF QUESTAR GAS COMPANY, THE DIVISION OF PUBLIC UTILITIES, AND UTAH CLEAN ENERGY FOR THE APPROVAL OF THE CONSERVATION ENABLING TARIFF ADJUSTMENT OPTION AND ACCOUNTING ORDERS	)))))))	Docket No. 05-057-T01

Surrebuttal Testimony

Artie Powell, PhD

Utah Division of Public Utilities

September 26, 2007

Q: Would you state your name, employer, and for whom you are testifying?

A: My name is Artie Powell; I am employed by and testifying for the Division of Public Utilities.

Q: Have you submitted testimony before in this proceeding?

A: Yes, in this phase of the proceeding I submitted rebuttal testimony on August 8, 2007.

Q: What is the purpose of this testimony?

A: In lieu of striking portions of his testimony, the Commission allowed parties to respond in writing to Dr. Dismukes' surrebuttal testimony within five working days. The purpose of my testimony is to provide limited response to the regression analysis presented in Dr. Dismukes' surrebuttal testimony and summarized in Exhibits attached to his testimony: Exhibit SR CCS 2.2 and Exhibit SR CCS 2.3 (corrected exhibit numbers). Specifically, I offer expert commentary on the regression methods and results presented by Dr. Dismukes in these two models.

Q: Would you briefly summarize your qualifications?

A: I have a doctorate degree in economics from Texas A&M University with a major field in econometrics. Econometrics is a subfield of economics, which applies mathematical and statistical theory, tools, and techniques to the analysis, interpretation and presentation of economic data. As a graduate student, I completed approximately a dozen graduate courses in econometrics, statistics, mathematics and mathematical economics. From 1985 to 2005, I taught economics, econometrics and statistics at the university level. From 1989 to 1995, I taught full-time at the University of Mississippi and helped coordinate the undergraduate and MBA statistical classes for the School of Business. From 1996 to 2005, I taught as an adjunct professor at Weber State University.

Q: Would you briefly explain your concerns with Dr. Dismukes regression analysis?

A: I have several concerns with the regression models and results provided by Dr. Dismukes in surrebuttal testimony purporting to show that there is a significant price effect on usage at the state level or specifically for Questar's GS customers. Specifically, Dr. Dismukes failed to account for the effects of autocorrelation in his regression models, which by itself renders the results of his models suspect. In addition, in the first of his regression models (Exhibit SR CCS 2.2; corrected exhibit numbering), the sample size is relatively small which makes drawing valid conclusions difficult. Thus, the regression analysis, results and conclusions drawn by Dr. Dismukes in his surrebuttal testimony are suspect.

Q: Could you briefly describe Dr. Dismukes' first model?

A: In his first model (CCS Exhibit SR CCS-2.2), Dr. Dismukes regresses the natural log (LN) of usage per customer against four variables, (1) LN of price, (2) LN of price lagged one year, (3) LN of a weather variable, and (4) a time trend. Algebraically, the model can be written as:

$$y = \alpha X_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \qquad (1)$$

where the dependent variable is

y = the natural log of usage per customer;

and the independent or explanatory variables are:

X_0 = an intercept (column of ones);
X_1 = the natural log of the annual price;
X_2 = the natural log of the annual price lagged one year;
X_3 = the natural log of a weather variable; and
X_4 = a time trend represented by the year (1998, 1999, …, 2005).

Q: What is your first concern with this model?

A: My first concern with this model is the sample size: there are only eight (8) observations.[1] (See Table 1; data provided by the Committee of Consumer Services in response to DPU data request 5.1).

[1] Dr. Dismukes lists nine (9) observations, but since the model contains a lag on the price variable only 8 observations are used in the model and analysis.

Table 1: CCS Model 1 Annual Data

Y Var lnUsePerCustomer	**X1 Var** lnPrice	**X2 Var** lnPriceLag	**X3 Var** lnHdd	**X4 Var** Year
4.63178805	1.85248010		8.73520359	1997
4.57468590	1.91285694	1.85248010	8.72176536	1998
4.51436681	1.84989693	1.91285694	8.66836802	1999
4.46606133	1.95815831	1.84989693	8.66888370	2000
4.42675714	2.19260553	1.95815831	8.70334075	2001
4.49891659	1.93963770	2.19260553	8.78063380	2002
4.38807447	2.05339214	1.93963770	8.64611397	2003
4.45794802	2.12777764	2.05339214	8.78109474	2004
4.35715740	2.27282668	2.12777764	8.73004395	2005

With eight observations and five explanatory variables (including the intercept), there are only three (3) degrees of freedom. In statistics, the term ***degrees of freedom*** (DF) is a measure of the number of independent pieces of information on which the precision of a parameter estimate is based. Generally speaking, the greater the DF, the more reliable or precise estimates are. Generally speaking, a larger sample size would increase the DF and improve the reliability of the model and its results.

An acceptable sample size will depend on a number of factors including the number of regressors in the model, the desired level of accuracy of each parameter being estimated and the desired level of model power (R^2). One "rule of thumb" suggests that for every parameter to be estimated you should have 30 data points or observations.[2] A more precise formula for computing the minimum sample size is given by:

$$n = \left(\frac{Z}{E}\right)^2 \left(\frac{1 - R^2}{1 - R^2_{xx_j}}\right) + k \qquad (2)$$

Where Z is the critical value corresponding to the standard normal distribution for a given test size (α); E is the desired margin of error or half of the width of the desired confidence interval for β_j; R^2 is the desired explanatory power of the model or coefficient of determination; $R^2_{xx_j}$ is the desired coefficient of determination for a model regressing X_j on the other regressors or explanatory variables of the model; and k is the total number of regressors including the intercept.[3]

[2] See for example, William Mendenhall, James E. Reinmuth, and Robert J. Beaver, "Statistics for Management and Economics," 7th ed., [Belmont, California: Duxbury Press, 1993], pp.251-261.

[3] Ken Kelly and Scott E. Maxwell, "Sample Size for Multiple Regression: Obtaining Regression Coefficients That Are Accurate, Not Simply Significant," Psychological Methods, (Vol. 8, No. 3), 2003, pp. 305-321. As the authors explain, the formula given here will result in a confidence interval no larger than the desired width (E) about 50% of the time. Thus, the formula can be considered a conservative estimate of the sample size. That is, the actual sample size necessary to ensure a confidence interval no greater than E may be considerably larger than that calculated from Equation (2).

As an example, suppose we use a test size of five percent (α = 0.05, Z = 1.96), R^2 = 0.80, $R^2_{xx_j}$ = 0.70, and k = 5, then to achieve a sample size of only 30, the margin of error would be approximately E = 0.3. If we assume that the "true" price elasticity from Dr. Dismukes' model (β_1) is between -0.5 to -0.01, then a margin of error of E = 0.3 seems relatively large.[4] If we use a margin of error E = 0.15, then the minimum sample size would be n = 119. Of course, the necessary sample size will vary depending on the values chosen for Z, E, R^2, and $R^2_{xx_j}$ but, I think this illustrates that we would have expected a sample size of more than 30 and possibly more than 100 in order to ensure the accuracy or reliability of the models estimates. As one expert states,

> It should be clear that the sample size is important. When the sample size is too small, the analyst cannot compute adequate measures of error in the regression results, and there can be no basis for checking model assumptions.[5]

Given the small sample, drawing valid conclusions from Dr. Dismukes' models is problematic. Compounding the problem is the likelihood that autocorrelation is present in this type of data.

---

[4] Given a margin of error E = 0.3, any estimate of β_1 greater than -0.3 would be insignificant. For example, given and estimate of -0.1, the 95 percent confidence interval would be -0.1 ± 0.3, or -0.4 to 0.2.

[5] Raymond H. Myers, "Classical and Modern Regression with Applications," 2nd Ed., [Boston, Massachusetts: PWS-Kent Publishing Company, 1990], p. 6.

Q: You indicated that autocorrelation might be problem. Could you explain your concern about autocorrelation?

A: Autocorrelation is a violation of one of the basic assumptions in regression models and refers to the dependent relationship among the regression errors (ε). When using economic data, it is not unusual for the regression errors to follow a first order autoregressive process:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \qquad t = 2, 3, ..., T \qquad (3)$$

where ε_t is the error term for observation "t"; ε_{t-1} is the error term for observation "t-1"; ρ is the correlation coefficient between ε_t and ε_{t-1}; and u_t is an error term that satisfies the fundamental regression assumptions[6]. As one author explains,

> The presence of the autocorrelation causes difficulty in the estimation of error variance and, as a result, in tests of hypotheses and confidence interval estimation.[7]

The presence of autocorrelation, in other words, would make it difficult to draw valid conclusions from Dr. Dismukes' regression results. I would note, that the presence of positive autocorrelation ($\rho > 0$), which is typical of economic data, makes it doubly difficult:

---

[6] The error terms u_t (t = 2, 3, …, T) are identically, independently, normally distributed random variables: $\mathbf{u} \sim IIN(\mathbf{0}, \sigma^2\mathbf{I})$, where **u** is the T-1x1 vector of errors and **I** is an identity matrix of dimension T-1.

[7] Myers, p. 288.

> The existence of positively correlated errors can result in an estimate of σ^2 [error variance] that is a **substantial underestimate**. This, of course, tends to inflate t-statistics on coefficients and deflate the width of confidence intervals on coefficients.[8]

A simple test for first-order autoregression based on the fitted residuals (e_t) is known as the Durbin-Watson test.[9] The test statistic "d" is given by the formula:

$$d = \frac{\sum_{t=2}^{T}\left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{T} e_t^2} \qquad (4)$$

Lower (d_L) and upper (d_U) bounds (or critical values) for the Durbin-Watson statistic are specified with respect to the sample size (T) and the desired testing or confidence level (α). The traditional Durbin-Watson test[10] for positive autocorrelation, where ρ is the correlation coefficient for the models error terms, is:

Reject H_0: $\rho = 0$, if $d < d_L$;

Fail to reject H_0: $\rho = 0$, if $d > d_U$; and

Declare the test inconclusive if $d_L < d < d_U$.

---

[8] Myers, p. 288, (emphasis added).

[9] Myers, pp. 289-290.

[10] The null and alternative hypotheses for positive autocorrelation are respectively H_0: $\rho < 0$ and H_a: $\rho \geq 0$.

However, there are two important qualifications or conditions to the use of the Durbin-Watson test. First, the regression must contain an intercept term. Second, the independent variables cannot contain a lagged dependent variable.[11] Dr. Dismukes' first regression meets both conditions. However, the inconclusive range presents an especially particularly "awkward problem" in small samples.[12]

This is illustrated in tables of critical values for the Durbin-Watson test by the absence of computed values for small samples and relatively large numbers of regression variables. For example, Johnston reports critical values starting with samples as small as six observations but lacks values when there is more than one independent variable. Indeed, for a given set of regressors, as the sample size declines it appears that the inconclusive range widens to the point where the test is incapable of detecting autocorrelation even if it exist. Given five regressors (including the intercept) and 100 observations, the inconclusive range is from 1.592 to 1.758 (a difference of 0.166); with 50 observations the range is 1.378 to 1.721 (a difference of 0.343); with 10 observations the range is 0.376 to 2.414 (a difference of 2.038); and with 8 observations no critical values are reported.[13]

---

[11] J. Johnston, "Econometric Methods," 3rd ed., [New York, New York: McGraw-Hill Book Company, 1984], p.316.

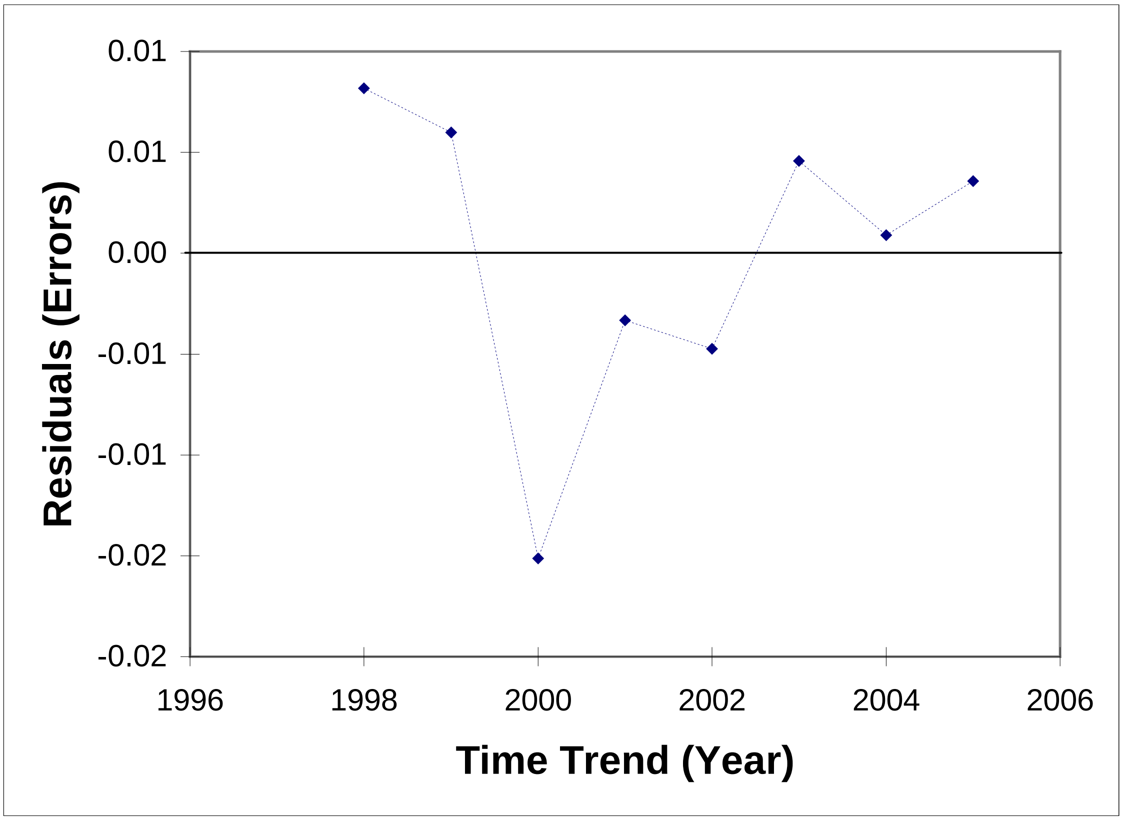
[12] Johnston, p. 316.

[13] For large samples, the Durbin-Watson statistic will approximately equal $d = 2(1 - \rho)$. Since ρ, the correlation coefficient, is a number between -1 and 1, d will range between 0 and 4. Values close to 2 ($\rho = 0$) indicate the absence of autocorrelation. Since the statistical distribution of d is unknown, critical values,

Given that Dr. Dismukes' model contains only 8 observations, the application of the Durbin-Watson test is impractical. However, a visual inspection of the error terms from his model indicates the presence of positive autocorrelation. The typical pattern for positive autocorrelation is for some sequential errors (or residuals) to be positive change to negative for a group of sequential errors and then switch to negative again. This pattern is repeated for the entire sample similar to a sine wave. A plot of the errors from Dr. Dismukes' model is provide in Figure 1.

d_L and d_U, are estimated using Monte Carlo simulations. In essence, when the sample size is too small, the critical values become 0 and 4, and the test is unable to detect the presence of autocorrelation when it exists.

Figure 1: Error Plot from Dr. Dismukes' Model 1



In the first of Dr. Dismukes' models, the sample size is relatively small at T = 8. With five regressors, that leaves only three (3) degrees of freedom. Additionally, the Durbin-Watson statistic is 1.75. However, it appears that the sample size is too small to carry out the hypothesis test for the presence of autocorrelation: it appears that upper and lower bounds are not tabulated for sample sizes this small with five regressors.[14] A graphical plot of the error terms against the time trend variable does suggest that autocorrelation is present. Given the small sample size and the apparent presence of autocorrlation in the data, drawing valid conclusions from this model is difficult.

Q: Could you explain how these concerns relate to Dr. Dismukes' second regression model?

A: The Division submitted a data request to the Committee asking for information on the diagnostic test performed by Dr. Dismukes for his second regression model (Exhibit SR CCS 2.3; corrected exhibit numbering). This is the data request that the Commission instructed the Committee to clarify. For convenience, the request and response are reproduced in Table 2.

From Dr. Dismukes' response to the data request, it appears that positive autocorrelation is present in the data for model 2 (Exhibit SR CCS 2.3; corrected exhibit number). That is, the P-value for the Durbin-Watson statistic is less than 0.0001, which indicates that we would reject the null hypothesis that $\rho = 0$ (no autocorrelation).

[14] See for example Johnston, Table B-5, pp. 554-557; or Meyers, Table C-7, p. 485. Johnston provides an upper bound of 2.588 for a sample size of nine (9) with five (5) regressors. Using the conservative approach as described herein, $d = 1.75 < 2.588$ and we would reject the null hypothesis: it appears that autocorrelation is present in Dr. Dismukes' first regression model.

Table 2: DPU Data Request 5.1

DPU Data Request:

5.1 For purposes of this request, please refer to Exhibits SR CCS-1.2 and SR CCS-1.3 of Dr. Dismukes's surrebuttal testimony.

…

d. Please provide all statistical diagnostic tests used to examine the statistical results.

Dr. Dismukes' Response:

Questar Monthly - with moving average
Durbin D-test
Positive Autocorrelation: Pr<DW : < 0.0001
Negative Autocorrelation: Pr>DW : 1.000
White's Test
Pr>ChiSq: <.0001

There are several simple corrections or transformations that can be performed on the data to remove the effect of autocorrelation on the error variance estimates. However, it appears from the testimony and data response that Dr. Dismukes did not perform any of these corrections. Thus, I would conclude that drawing a valid conclusion from this model is not possible. Again, the presence of positive autocorrelation could substantially over-inflate the t-statistics of the model making the finding of statistical significance in the model suspect.

Q: Would you summarize your testimony?

A: Dr. Dismukes offered two models in his surrebuttal testimony that purportedly show that there is a statistically significant relationship between usage per customer and the price of natural gas. While this relationship may be consistent with economic theory, conclusions to this effect cannot be validly drawn from the models and results presented by Dr. Dismukes. First, both models appear to exhibit the presence of autocorrelation, which can cause over-estimation of the t-statistics. In the absence of correcting for this problem, the model can lead to false conclusions that its coefficients are statistically significant. In addition, in Dr. Dismukes' first model, the sample size is too small to allow valid conclusions to be drawn. Thus, I would recommend that the Commission place little or no weight on this portion of Dr. Dismukes' surrebuttal testimony.

Q: Does that conclude your testimony?

A: Yes it does.