Filling in the Gaps for A Partially Discontinued Data Series

James R. Knaub, Jr.
Energy Information Administration, EI-53.1
US Dept. of Energy, 1000 Independence Ave. SW, Washington DC, 20585

Key Words:
linear regression; heteroscedasticity; regression weights; model; prediction; combining estimators; standard error of the prediction error; SAS PROC REG; missing 'time series' data

Abstract:
Data on US coal production, imports, producer and distributor stocks, consumption, exports, consumer stocks, and, by default, losses and unaccounted for coal, have been collected, and, considering changes in stock levels, all data have been 'balanced' on a quarterly basis. That is, the relationship between these variates is written as a single equation. These data have been published in the Energy Information Administration (EIA) Quarterly Coal Report for more than sixteen years. Producer and distributor stocks (p/d stocks) will no longer be collected on a quarterly basis, due to budgetary constraints, but will be observed annually. The EIA still wishes to publish these data quarterly, with 'estimates' given for p/d stocks for the first, second and third quarters of each year, and the observed value given for the fourth quarter. (Note that these "estimates" are actually "predictions." However, unlike the usual case with predictions, no values will ever be observed for first, second or third quarter p/d stocks after 1997. Also, we are only interested in a value to approximate the current conditions for each of these publications, not forecasts of the future.)

This paper explores the use of weighted linear regression modeling, prediction and the variance of the prediction error, and the combination of 'predictions' from different models, to help fill in these unobserved p/d stock quarterly values in a reasonable manner, and provide estimated standard errors. Note, however, as indicated by substantial data revisions and the inherently imprecise nature of data collection for coal, nonsampling error is an enormous consideration whether prediction is used or not, and no matter what procedures are employed to predict data.

The procedures found here have substantial potential for use whenever one might consider reducing the frequency of periodic data collection for an established data series.

1. Background:

There are 65 quarters of data (back to 1981) available for modeling p/d stock predictions; data on covariates will continue to be collected, and p/d stocks will continue to be observed annually. (See Table 1 in Quarterly Coal Report (1996) (96/4Q) and other issues of that publication.) We will no longer break this down by State, but only have one national number. A quick examination of past revisions shows that there may have been enough nonsampling error to justify this decision, even if all quarterly observations were continued.

The conventional notion of a "time series" was considered, but the series is to be broken. It is not that there is a need to project ahead to each quarter until the data are later obtained, because data for three out of each four future quarters will never be observed.

Another statistician reminded the author that there could be benefit in separating the data into four groups, depending upon the quarter of the year in which the p/d stocks were collected. However, the author had already considered that, and in the case of these data, the benefit of what might be considered the virtual addition of another regressor (i.e., time-of-year), was not very great, and was offset by the resulting dependence upon smaller data sets, which are made less stable by the presence of possibly substantial nonsampling error.

An attempt was made to use regression against covariates that did not include the previous quarter for p/d stocks. One would then assume that the relationships between these covariates and p/d stocks that were present within the 65 quarters of previous data will not soon change over time. The quality of this assumption could be judged to some extent in the future by compiling annual comparisons of fourth quarter observed data and the predictions that would have resulted were this methodology to be used to predict fourth quarter results.

Without using the previous p/d stocks as a regressor, variance was too high to make results very useful, and the team leader needing the data was not happy with those results. She had emphasized that p/d stocks may be calculated from covariates, except for a losses term. However, this assumes that the previous quarter's p/d stock level was actually observed, which will only be the case in the first quarter of each year. This led back to addressing the problem of using the previous p/d stock level when only a predicted previous value will be available.

2. The Procedure:

A solution lay in using regression weights to account for the fact that the previous p/d stock value was predicted, when that is the case. These regression weights are described below and used in Eq(4.16). They are decomposed into a part that deals with the usual heteroscedasticity, and a part that accounts for the number of time periods since data for the variate of interest were last collected. The inverse square root of a regression weight is the 'nonrandom factor' when residuals are factored into random and nonrandom parts, as in Eq(4.16). This can be implemented easily using STDI under SAS PROC REG, where STDI is an estimate of the standard error of the prediction error. (This is not an endorsement of products for the SAS Institute Inc., but merely reflects the software available to the author.)

Three predictors were then developed, one of which does not use the previous p/d stocks, and the other two do. Both multiple and simple linear regression were used. SAS PROC REG can be used to produce diagnostics and may indicate if there is a problem with multicollinearity. It automatically makes a correction for a matrix singularity. (See SAS Institute Inc. (1985), pages 671 and 672, and Maddala (1977), pages 183-194.) The three predictors, and a combined predictor, were applied to 17 cases using test data from the previous 65 quarters. (See Table 1.) Some of the data were used as the data base, and some were temporarily deleted in order to make 'predictions' and compare results to the values observed. For example, ***the second test case in Table 1 was obtained by using data through the last quarter of 1996 for coal p/d stocks, plus data for the other regressors through the first quarter of 1997, to estimate coal p/d stocks for the first quarter of 1997. That 'prediction' was then compared to the number observed.*** (Note that there have been revisions in the data since Table 1 was compiled, so exact individual results may not be replicated.)

3. Combining Results:

Knaub (1992) contains an example of combining estimates using survey sample data. The method for combining estimates was as described in Granger and Newbold (1986). In their book, Forecasting Economic Time Series, in both the first edition, 1977, and the second edition, 1986, they say (2nd ed., pp. 266-272) that their experience indicates that "ignoring" correlation between errors for different methods when combining estimates works better than trying to account for them. (The formulas proposed ignore bias, but Knaub (1992) suggests that bias may be a factor in the performance of the results which led Granger and Newbold to their conclusion.) The author performed some investigation in the current context, trying different approaches, but came to about the same conclusions for these data. Application of this methodology is shown in section 7.

4. Theory:

As stated earlier, two of the three models employed use the previous p/d stocks value, observed or predicted, and would therefore require an adjustment when those previous p/d stocks are predicted. Using the so-called 'calculation' for 'balancing the books' on coal, and writing it as an equation,

$$y_i \;=\; x_{1_i} + x_{2_i} - x_{3_i} - x_{4_i} - x_{5_i} + x_{5_{i-1}} + y_{i-1} - L_i \;=\; x_{\Sigma_i} - L_i \qquad \text{Eq(4.1)}$$

where y_i and y_{i-1} represent the current and past p/d stocks, and x_{5_i} and $x_{5_{i-1}}$ represent the current and past consumer stocks. Further, $x_{1_i} + x_{2_i} - x_{3_i} - x_{4_i} - L_i$ represents, respectively, coal production, plus imports, minus consumption, minus exports, minus "losses and unaccounted for." (See Table 1 in Quarterly Coal Report (1996) (96/4Q), and other issues.)

Note that, $x_{1_i} + x_{2_i} - x_{3_i} - x_{4_i} - x_{5_i} + x_{5_{i-1}} + y_{i-1} = x_{\Sigma_i}$ is y_i if we ignore L, and is used below. Writing $y_i = x_{1_i} + x_{2_i} - x_{3_i} - x_{4_i} - x_{5_i} + x_{5_{i-1}} + y_{i-1} - L_i = x_{\Sigma_i} - L_i$ as

$$y_i = x_{\Sigma_i} + b_0 + (b'_1)(x_{\Sigma_i}) + e_i = b_0 + (b'_1 + 1)(x_{\Sigma_i}) + e_i = b_0 + (b_1)(x_{\Sigma_i}) + e_i , \qquad \text{Eq(4.2)}$$

“losses and unaccounted for” (which will just be referred to as "losses" now) are replaced by a simple linear regression. Graph 1 shows that there is a definite linear relationship between x_Σ and losses, but with substantial variance. A graph of y_i *vs* x_{Σ_i} , Graph 2, shows promise.

As another model, a multiple regression version of the above was used. The components of x_Σ are used separately as regressors. This performed slightly better. (The signs of each estimated coefficient were found to be in the same direction as the signs used in the 'calculation,' Eq(4.1).)

As a third model, already mentioned, multiple regression was considered, but without the use of the observed previous p/d stock levels as a regressor. This guarantees that for one model, a new data point will be added each year. Currently, however, because of the larger associated variance, the results of this model do not contribute heavily to the combined results.

In Table 1 (“Summary of Examples ...”), predictions and estimated standard errors of the prediction errors are given for this last described model first, and its larger estimated standard errors are apparent. Results from the model that performed second best, the simple linear regression, which included the previous p/d stock level, were shown next, and third are the results from the multiple linear regression, also including previous p/d stock level. The combined ‘estimate’ (i.e., ‘prediction’) and the estimated standard error of the prediction error are then shown, next to the observed (or “collected”) value.

Below will be shown what is proposed for cases where y_{i-1} is replaced by a predicted value, which must increase variance to some degree. In the future, this will occur for two quarters (the second and third) out of each year. In the first quarter, p/d stock level will have been collected the previous quarter. For the second and third quarters, this will not be true.

The fourth quarter data will be observed.

Let Q be the number of quarters since the last observed y (i.e., observed p/d stock level). The three models used will first be described for the case of Q = 1:

For the simple linear regression model, resulting in the second predictions in each example in Table 1, using $x_{\Sigma_i} = x_{1_i} + x_{2_i} - x_{3_i} - x_{4_i} - x_{5_i} + x_{5_{i-1}} + y_{i-1}$, one may write

$$y_i = b_0 + b_1 x_{\Sigma_i} + [w_i]^{-1/2} e_{0_i} = \hat{\beta}^{s/\,T} \boldsymbol{X}_i^{/} + b_1 \, y_{i-1} + [w_i]^{-1/2} e_{0_i}, \qquad \text{Eq(4.3)}$$

where $\hat{\beta}^{s/}$ involves only b_0 and b_1. The "s" in the superscript indicates that this is a simple linear regression, but the prime symbol indicates that the previous p/d stock level, y_{i-1}, has been partitioned from the remainder of the regressor. Since the same coefficient, b_1, is used for both parts, we will consider this to be a simple linear regression. Finally, w_i is the regression weight, which is the inverse square root of the nonrandom factor of the residuals.

Now consider the case of the multiple regression model that includes y_{i-1} as a regressor, results being shown as the third, and generally best predictions based on a single model in Table 1. There are seven regressors in this application. The intercept, in this case, was estimated to be essentially indistinguishable from zero, so a zero intercept was specified. The author's experience indicates this may be prudent to avoid model processing anomalies, especially considering the impact of nonsampling error. This may not be necessary here, but it should not be detrimental to any substantial degree either since the intercept was estimated to be near zero. Therefore, here there are seven regressors, including y_{i-1}, but no intercept (i.e., the intercept is zero). The applicable regression equation may be written as

$$y_i = \hat{\beta}^{/\,T} \boldsymbol{X}_i^{/} + b_7 \, y_{i-1} + [w_i]^{-1/2} e_{0_i} \quad . \qquad \text{Eq(4.4)}$$

Finally, a multiple regression model is used, employing six of the regressors above, excluding y_{i-1}. An intercept term is used in this model. The results are shown as the first and generally least accurate of those listed in Table 1. This regression equation may be written as

$$y_i = \hat{\beta}^{T} \boldsymbol{X}_i + [w_i]^{-1/2} e_{0_i} . \qquad \text{Eq(4.5)}$$

In the above, $\hat{\beta}^{T} \boldsymbol{X}_i$ accounts for six regressors, their coefficients and an intercept, but $\hat{\beta}^{T/} \boldsymbol{X}_i^{/}$ and $\hat{\beta}^{T s/} \boldsymbol{X}_i^{/}$ represent cases where y_{i-1} is used, but shown separately. Thus, whenever y_{i-1} is used, it will be written separately, and the primes will be used to indicate that the matrix of other regressors and their coefficients, and an intercept term if included, is incomplete without it.

Let w_i be the usual regression weight due to heteroscedasticity, applicable regardless of the value of Q. In this application, w_i is based on x_{Σ_i}. When Q = 2 or Q = 3, x_{Σ_i} is estimated because y_{i-1} is estimated, but w_i is not very sensitive to this.

Considering the standard form $w_i = x_{\Sigma_i}^{-2\gamma}$, using ideas presented in Knaub (1997), it was decided to use $\gamma = 0.5$ as an improvement over ordinary least squares (OLS). Thus a weighted least squares (WLS) regression is being used.

(Note that in the case of prediction, the value to be predicted may be either smaller or larger than ‘average’ for that variate. However, for inference from cutoff model sampling, the main impetus behind Knaub(1997), the uncollected values of the variate of interest corresponding to members of the population not in the sample, would generally be substantially smaller than ‘average.’ Therefore, since a larger value for γ would put more emphasis on data from the ‘smaller’ members of the population that did respond in the sample, this may be more useful as those entities may be more like the members whose data one is estimating. Because a useful regression weight may be sensitive to this, that is part of what should be considered. See Knaub (1997). With the sample size found in the current application, rather than calculate a value for γ from the data, a default value is used. For cutoff model sampling, we <u>may</u> have wanted a larger value for γ. Here, however, in this application involving prediction, we may feel more confident in using $\gamma = 0.5$.)

Now, when Q = 2, in general, one may write

$$y_i \; = \; \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}} \boldsymbol{X}_i' + b_y \, y_{i-1}^{(Q=2)} + [w_i]^{-1/2} e_{0_i} \quad , \qquad \text{Eq(4.6)}$$

or here

$$y_i \; = \; \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}} \boldsymbol{X}_i' + b_y \, y_{i-1}^{(Q=2)} + x_{\Sigma_i}^{1/2} e_{0_i} \quad , \qquad \text{Eq(4.7)}$$

where

$$y_{i-1}^{(Q=2)} \; = \; \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}} \boldsymbol{X}_{i-1}' + b_y \, y_{i-2} + x_{\Sigma_{i-1}}^{1/2} \, e_{0_{i-1}} \quad , \qquad \text{Eq(4.8)}$$

and y_{i-2} is assumed known well enough from having been ‘observed.’

$$\hat{y}_{i-1}^{(Q=2)} = \hat{\beta}^{T\,/} \boldsymbol{X}_{i-1}^{/} + b_y \, y_{i-2} \qquad \text{Eq(4.9)}$$

is the predicted value.

For Q = 3, one has

$$y_{i-1}^{(Q=3)} = \hat{\beta}^{T\,/} \boldsymbol{X}_{i-1}^{/} + b_y \, y_{i-2}^{(Q=2)} + x_{\Sigma_{i-1}}^{1/2} \, e_{0_{i-1}} \quad . \qquad \text{Eq(4.10)}$$

and

$$\hat{y}_{i-1}^{(Q=3)} = \hat{\beta}^{T\,/} \boldsymbol{X}_{i-1}^{/} + b_y \, \hat{y}_{i-2}^{(Q=2)} \qquad \text{Eq(4.11)}$$

Considering that the e_{0_i} are iid, and writing the standard error of e_{0_i} as σ_e , and considering the iterative nature of this problem, the solution could be written down. (See the appendix.) However, this may also be seen through the use of conditional and unconditional variances, as suggested by Richard Sigman. (See Acknowledgment.) To do this, consider Hansen, Hurwitz and Madow (1953), page 66. (Note that w is the usual regression weight. We are now describing the full influence on residuals when 'predicted' values are used for previous p/d stocks.) From Hansen, Hurwitz and Madow (1953):

$$\sigma^2_{y_i} = E\sigma^2_{y_i|y_{i-1}} + \sigma^2_{E(y_i|y_{i-1})} \quad . \qquad \text{Eq(4.12)}$$

This leads to

$$\sigma^2_{y_i}(Q=1) \quad = \quad w_i^{-1}\sigma_e^{\,2} \qquad \text{Eq(4.13)}$$

$$\sigma^2_{y_i}(Q=2) \quad = \quad w_i^{-1}\sigma_e^{\,2} \;+\; b_y^{\,2} w_{i-1}^{\;-1}\;\sigma_e^{\,2} \qquad \text{Eq(4.14)}$$

and

$$\sigma^2_{y_i}(Q=3) \quad = \quad w_i^{-1}\sigma_e^{\,2} \;+\; b_y^{\,2} w_{i-1}^{\;-1}\;\sigma_e^{\,2} \;+\; b_y^{\,4} w_{i-2}^{\;-1}\;\sigma_e^{\,2} \qquad \text{Eq(4.15)}$$

where $\sigma_e^{\,2}$ is the variance of the random factor, e_0, of the residual. (See Knaub (1997).)

Therefore, the general regression equation to be used here is of the form

$$y_i \;=\; \hat{\boldsymbol{\beta}}^{\prime\,\boldsymbol{T}}\;\boldsymbol{X}_{\boldsymbol{i}}^{\prime} \;+\; b_y\,\hat{y}_{i-1}^{(Q)} \;+\; e_{0_i}\sqrt{\sum_{q=1}^{Q} w^{-1}{}_{i+1-q}\;b_y^{\,2(q-1)}} \qquad \text{Eq(4.16)}$$

where, with the exception of any term based on a previous y value (this deletion being indicated by prime signs), $\hat{\boldsymbol{\beta}}^{\prime\,\boldsymbol{T}}\;\boldsymbol{X}_{\boldsymbol{i}}^{\prime}$ represents all other coefficients, regressors, and an intercept if used. Further, the term $b_y\,\hat{y}_{i-1}^{(Q)}$ is made up of an assumed constant coefficient (which did appear relatively constant for these data on several examinations during testing), and the previously observed value of y when available, or an estimate when the previous y was not observed. "Q" represents the number of quarters since y was last observed. Thus if Q = 1, the previous y value was observed, and the "hat" in this previous y-value term is not appropriate. That is, $\hat{y}_{i-1}^{(Q=1)} = y_{i-1}$.

Finally, w^{-1}_{i+1-q} is the inverse of the complete regression weight when previous y values are either observed, so Q = 1, or when previous y values are not used in the model. The subscript, $i+1-q$, means that if Q = 1, then only the current w_i value is needed. However, if Q > 1, then at least one previous value of *w* is needed to construct the more complicated form for the residual shown in Eq(4.16), which is used for only one y value, the current one, when predicting that value.

5. Application:

When Q > 1, x_{Σ_i} will be estimated using a predicted value for y_{i-1}. Assuming results are not as sensitive to moderate changes in weights as they are to changes in previous values of coal p/d stocks, let $w^{-1}_{i} = w^{-1}_{i-1} = \ldots\ldots = w^{-1}_{k} = x_{\Sigma_i}$, or $\hat{x}_{\Sigma_i}$ (approximately). Not concerning ourselves with the 'hat' in such cases, neither here nor further below, this yields

$$y_i = \hat{\boldsymbol{\beta}}^{/\,\boldsymbol{T}} \boldsymbol{X}_{\boldsymbol{i}}^{/} + b_y \hat{y}_{i-1}^{(Q)} + e_{0_i} x_{\Sigma_i}^{1/2} \sqrt{\sum_{q=1}^{Q} b_y^{2(q-1)}} \quad , \qquad \text{Eq(5.1)}$$

a special case of Eq(4.16). The test results in Table 1, and the programming code listed make use of Eq(5.1). It would be interesting to make use Eq(4.16) in the future, and see if performance could be improved appreciably.

Note that if there are no previous values of y being used as a regressor, then Eq(5.1) has no y term on the right side of the equation, and then $\hat{\boldsymbol{\beta}}^{/} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{X}_i^{/} = \boldsymbol{X}_i$, so then

$$y_i = \hat{\boldsymbol{\beta}}^{\boldsymbol{T}} \boldsymbol{X}_i + x_{\Sigma_i}^{1/2} e_{0_i} \qquad \text{Eq(5.2)}$$

which is a special case of Eq(4.5). This is the case for one of the three models chosen. This was done so that even if Q were large enough to make the associated variance and reasonableness of this procedure cause for concern in the other models, there would still be one model in use that only employs those other regressors, thus not being dependent upon

Q, and also adding a new data point each year to the historical data used to estimate β in this model.

6. Description of Models Used:

The models used in this application will now be considered in increasing order of test performance (see Table 1), and therefore in order of apparent increasing accuracy, using the data currently available.

In short, the first model used in the table is the one that is multivariate, but does not consider previous y values; the second is the simple linear regression; and the third model (apparently the most accurate at this point in the data collection) is multivariate, as in the first model, but the third model includes a term involving a previous p/d stock variate. Results from these models are given in this order in Table 1, followed by combined predictor results.

The first model below will operate with a data base that increases over time, but this is not true for the other two models. In this model, there are six regressors, none involving previous coal p/d stock (i.e., "y") values. An intercept term (i.e., not set equal to zero) is included. Here one has

$$y_{M1_i} = \hat{\beta}^T \boldsymbol{X}_i + e_{0_i} x_{\Sigma_i}^{1/2} \quad , \qquad \text{Eq(6.1)}$$

which corresponds to Eq(5.2), where "M1" simply indicates that it is associated with the first predictor (i.e., "model 1") results in Table 1.

For the second model in the order taken for Table 1 (the order of apparent test accuracy), there is one regressor and an intercept term. (Call this "model 2.") The previous coal p/d stock level, y_{i-1} , or an estimate of it, is a part of the regressor. Here one has

$$y_{M2_i} = \hat{\beta}^{T\,s/} \boldsymbol{X}_i^{/} + b_1\, \hat{y}_{i-1}^{(Q)} + e_{0_i} x_{\Sigma_i}^{1/2} \sqrt{\sum_{q=1}^{Q} b_1^{2(q-1)}} \quad , \qquad \text{Eq(6.2)}$$

which is a form of Eq(5.1), where "M2" indicates the second model associated with Table 1, $\hat{\boldsymbol{\beta}}^{s/}$ involves only b_0 and b_1, and this is the same "b_1" that is the coefficient for the y_{i-1} related portion of the single regressor. (Refer to Eq(4.3).)

In the third model (in order of use in Table 1), seven regressors were used, including a previous y value (or an estimate of it), but, as stated earlier, the intercept was found to be very nearly zero, and was set equal to zero. Call this "model 3." Here one has

$$y_{M3_i} = \hat{\boldsymbol{\beta}}^{T/} \boldsymbol{X}_i^{/} + b_7 \, \hat{y}_{i-1}^{(Q)} + e_{0_i} x_{\Sigma_i}^{1/2} \sqrt{\sum_{q=1}^{Q} b_7^{2(q-1)}} \quad , \qquad \text{Eq(6.3)}$$

another form of Eq(5.1). Here b_7 is used to indicate that one of the regressors (y_{i-1}) is being considered separately from the rest of $\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i$, so we use prime marks on the first term, as before.

7. Combining Results in Practice:

The "STDI," ("standard error of the individual predicted value," page 663 in SAS Institute Inc. (1985)), an option in SAS PROC REG, was used for each of the three models above. (*Note that this is not an official endorsement of that product. Other vendors may supply this feature. The author is simply stating what was used with regard to the software that was available to the author.*)

From STDI, and the above models, one has $(\hat{y}_{M1}, \hat{\sigma}_{M1})$, $(\hat{y}_{M2}, \hat{\sigma}_{M2})$, and $(\hat{y}_{M3}, \hat{\sigma}_{M3})$, or this could be written as $(\hat{y}_{Mj}, \hat{\sigma}_{Mj})$, representing the predicted new quarter producers'/distributors' stocks and corresponding estimated standard error of the prediction error, using model j.

Using STDI and section 9.2.4 in Granger and Newbold (1986), $\hat{\rho}(M2, M3)$ was about 0.48. However, using techniques displayed in Granger and Newbold (1986), weighting factors used for combining predictions which employed this value of rho were, in effect, not so very different from those found using $\rho(M2, M3) = 0$. Considering this, and considering the volatility of results, the overall empirical results found in Granger and Newbold (1986), pages 266 through 272, those in Knaub (1992), and extending these results from two to three models, we have

$$\hat{y}_c = \frac{\hat{\sigma}^2_{M2}\hat{\sigma}^2_{M3}\ \hat{y}_{M1} + \hat{\sigma}^2_{M1}\hat{\sigma}^2_{M3}\ \hat{y}_{M2} + \hat{\sigma}^2_{M1}\hat{\sigma}^2_{M2}\ \hat{y}_{M3}}{\hat{\sigma}^2_{M1}\hat{\sigma}^2_{M2} + \hat{\sigma}^2_{M1}\hat{\sigma}^2_{M3} + \hat{\sigma}^2_{M2}\hat{\sigma}^2_{M3}} \qquad \text{Eq(7.1)}$$

and

$$\hat{\sigma}_c^{\,2} = \frac{\hat{\sigma}^2_{M1}\hat{\sigma}^2_{M2}\hat{\sigma}^2_{M3}}{\hat{\sigma}^2_{M1}\hat{\sigma}^2_{M2} + \hat{\sigma}^2_{M1}\hat{\sigma}^2_{M3} + \hat{\sigma}^2_{M2}\hat{\sigma}^2_{M3}} \qquad \text{Eq(7.2)}$$

Thus, the "bottom line" information we are seeking is $(\hat{y}_c,\ \hat{\sigma}_c)$. The first value in this pair is a prediction of the current (missing) quarterly coal p/d stock level, and the latter value is an estimate of the standard error of that prediction error. That estimated standard error is impacted by every kind of error, with the exception of a consistent bias, and so it is an excellent overall measure of the quality of the estimate of p/d stocks data. Further, because it is the result of combining predictors, biases for these models are likely to counteract each other to some extent, thus preventing a relatively large bias in any one of them from being likely to be extremely influential.

8. Testing:

As for application, 17 examples using test data from within the 65 sets of quarterly data already observed, and employing these equations, performed well. (This conclusion is based upon the relationship of predicted values to estimated standard errors of the prediction errors and the corresponding number actually observed, and considering the magnitude of revisions made.) Using constant coefficients within each model appears to be reasonable. Under several test data set conditions examined, corresponding b values seemed fairly stable.

Examining the test results, it is clear that the accuracy of the three models seems generally subject to ranking as indicated earlier. Also, as must necessarily be the case, larger "Q-values" mean larger variances. However, Q did not appear to be as influential here as the model chosen. Since a combined prediction was then developed, and, as just indicated, Q = 3 was not a lot worse than Q = 1, these were positive developments. However, it would be interesting to see what differences would occur in such test data, should experimentation with Eq(4.16) be used, instead of using Eq(5.1).

Note on the next page that when considering these results (rounded to millions of short tons), in 12 out of 17 tests, the absolute difference between the predicted and observed values is approximately less than or equal to the estimated standard error. For this number of test cases, this is a good result. However, the extreme cases (numbers 4 and 14) are cause for some concern.

Note also that it appears that this method may tend to underestimate. However, if one were to plot predicted values on the x-axis of a graph, and the corresponding observed values on the y-axis, it would appear that this method overestimates. By eliminating test case number 14, however, a regression through the remaining data points nearly extends through the origin.

Table 1 - *Summary of some testing done as of January 5, 1998:*

(Predictions, estimated standard errors, and collected values are in millions of short tons.)

obs. #	Q Value	M1 Pred 1, s.e. 1		M2 Pred 2, s.e. 2		M3 Pred 3, s.e. 3		Pred.	s.e.	Collected Value
1	3	31	4	31	3	29	2	30	2	31
2	1	36	5	40	3	36	2	37	2	38
3	1	38	5	42	3	40	2	40	2	42
4	2	33	4	38	3	38	3	37	2	42
5	3	32	4	35	3	34	3	34	2	36
6	1	34	4	38	2	35	2	36	1	37
7	2	32	4	36	3	35	2	35	2	37
8	3	32	4	31	3	31	2	31	2	34
9	1	33	5	36	2	31	2	33	2	34
10	2	30	5	35	3	33	2	33	2	36
11	3	31	5	35	3	33	3	33	2	33
12	1	40	5	39	3	38	2	38	2	38
13	2	36	4	35	3	36	2	36	2	35
14	3	40	6	31	3	32	3	32	2	27
15	1	37	5	40	3	38	2	38	2	40
16	2	38	5	38	3	39	3	38	2	41
17	3	34	4	35	3	36	3	35	2	35

All estimates are for quarters in 1992, or more recent years, using only data from previous periods (except that the b_y estimates are made using all data). So ... the worst case is N=44, n=43.

9. File maintenance for future prediction of producer/distributor coal stocks:

Note that nonsampling error is a large influence here. ("Nonsampling error" usually refers to processing and other errors that are part of a survey whether it is a sample survey or a census survey (enumeration). Here it is being used to cover the same types of errors as "nonsampling error" would usually cover in a survey.) Revisions of more than a million short tons are typical for production. However, one change noted of more than 2 million short tons for producer/distributor stocks was far more substantial. That is, a change of 2 million out of hundreds of millions for a regressor is relatively minor, except that a great deal of precision is typically used. However, a change of 2 million out of tens of millions, for the variate of interest, is of much greater impact.

....

The purpose of this effort was to be able to estimate producer/distributor stocks (p/d stocks) for the first, second and third quarter of each calendar year. The fourth quarter data will be observed, so that a prediction for the fourth quarter will not be necessary. (Future prediction of the fourth quarters may be compared to collected (observed) values for purposes of evaluating the continued efficacy of the models, but the observed p/d value will continue to be the value reported for each of the fourth quarter periods in the future.) As described below, the size of the data file being used to predict the p/d stocks value each of the first three quarters in the future will not be increased over the 67 data points collected through the end of 1997, except that the first of the three models will be able to use an additional data point observed each year.

SAS and FORTRAN source code found in the file QCR.PDSTOCKS.PROGRAM (a hardcopy is seen below) may be used as a module to be inserted into the Quarterly Coal Report (**QCR**) production software. Most of the work is done by SAS PROC REG, and as long as results from that SAS PROC may be used, there will not be a tremendous amount of programming to be done.

QCR.PDSTOCKS.PROGRAM outputs the prediction for p/d stocks, and the estimated standard error of the prediction error. This standard error is typically 1.5 to 2 million short tons for the first quarter when in the previous quarter we actually observed data for p/d stocks. It is typically closer to 2.5 million short tons for the third quarter, when the previous two quarters have estimated p/d stocks. The second quarter results generally fall between these extremes. Therefore, the typical standard error of the prediction error will normally be about 2 million short tons. It is suggested that a footnote in the QCR should state this, and provide the standard error estimate for the current quarter as well.

The input file for QCR.PDSTOCKS.PROGRAM is COAL.PDSTOCKS.DATAQCR. (See a hardcopy of this file below.) This input file will need to be updated for each new quarter.

The records in COAL.PDSTOCKS.DATAQCR are fixed with nine fields each. The first field contains the "Q value." If that value is 1, 2 or 3, then that indicates the number of quarters since the last collection/observation of the p/d stocks volume datum. For the historical data, that value is a "1," indicating that the previous p/d stock value was a collected value. If we are estimating for the first quarter of a new year, then all Q values will be "1," except as noted here. Also, in this case, all p/d values will be actually collected/observed values, except as noted.

An exception (to the value "1") for the Q values is that once a predicted value is inserted for the p/d value, that record is to be flagged for exclusion in any future quarterly predictions for models M2 and M3. To that end, Q will be shown as a "0," or a "4," as will be explained.

Observed data will be available quarterly through 1997. In the first quarter of 1998, all Q values will be "1." In the second quarter of 1998, all Q values will be "1," except the first quarter of 1998, which will be changed to a "0," and the second quarter of 1998, which will be a "2." In the third quarter of 1998, Q values will all be "1," except the first two quarters of 1998 which will both be "0," and the third will have a "3" in that field. For the fourth quarter of 1998, a new p/d value will be collected. The Q value for the fourth quarter of this and subsequent years will always be "4," (another 'flag' value, like "0"), because the previous quarter p/d stock value will never be reported. In the first quarter of 1999, all Q values will be "1," except for "0" in each of the first three quarters of 1998, and a "4" in the fourth quarter 1998 record. Notice that from that point forward, the Q value for the current quarter will be a "1," "2," "3," or "4," depending on the quarter of the year; the Q value of every record up through the fourth quarter of 1997 will be a "1;" and all other Q values will be "0," except other fourth quarter values that will each be a "4." The reason for using the flag "4" is that unlike the records marked "0," which are ignored in all of the predictions, a record with a Q value of "4" will be used in one of the three models (M1) being combined with the other two to provide the final prediction. (See the SAS and FORTRAN code on pages 21 through 24.)

The second of the nine fields in COAL.PDSTOCKS.DATAQCR is the production value observed for the given quarter. The third field is for the volume of imports. (Note that all values in the record, other than the "Q value," are in thousands of short tons.) The fourth field is for the current producer/distributor stocks. The fifth and sixth fields are for consumption and exports. The seventh and eighth fields are for current and previous quarter consumer stocks, respectively. In the last field, we find the previous quarter value for producer/distributor stocks. It may be either observed or predicted, as indicated by the value of Q.

After the 67th record in COAL.PDSTOCKS.DATAQCR is completed, using data for the fourth quarter of 1997, the 68th record will need to be filled with production and import volume values, etc. There will be one blank field, the fourth one, for the first quarter 1998 p/d stocks value. At that point, we may begin using QCR.PDSTOCKS.PROGRAM or its equivalent, embedded in the **QCR** production software, to 'predict' the 'missing' p/d stock values.

Acknowledgments:

Many thanks to Richard S. Sigman whose comments on corrections to an early draft were essential. Also, others have brought the author's attention to a technical correction and to points of interest that otherwise would not have been addressed. Thanks to all.

References:

Energy Information Administration, Office of Coal, Nuclear, Electric and Alternate Fuels (1996), Quarterly Coal Report, DOE/EIA-0121(96/4Q) (Washington DC).

Granger, C.W.J., and Newbold, P. (1986), Forecasting Economic Time Series, 2nd edition, Academic Press.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953), Sample Survey Methods and Theory, Volume II (Theory), John Wiley & Sons.

Knaub, J.R., Jr. (1992), "More Model Sampling and Analyses Applied to Electric Power Data," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 876-881.

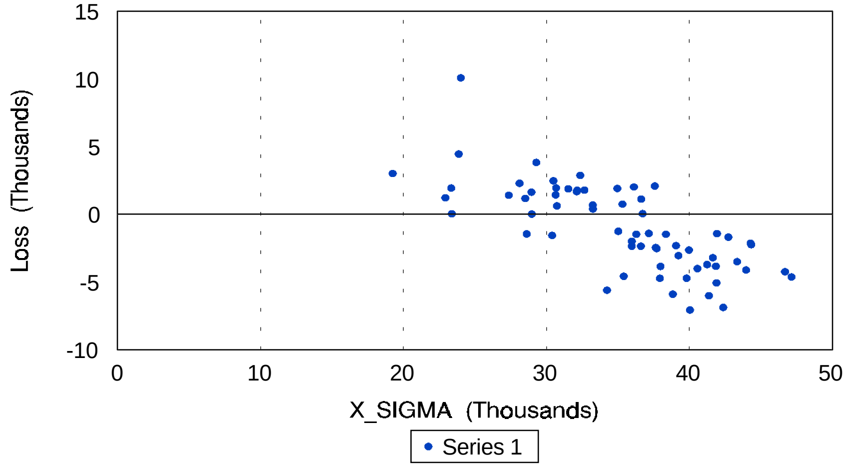
Knaub, J.R., Jr. (1997), "Weighting in Regression for Use in Survey Methodology," InterStat, April 1997, http://interstat.stat.vt.edu/InterStat. (Note shorter, more recent version in ASA Proceedings of the Section on Survey Research Methods, 1997.)

Maddala, G.S. (1977), Econometrics, McGraw-Hill, Inc.

SAS Institute Inc. (1985), SAS User's Guide: Statistics, Version 5 Edition, Cary, NC: SAS Institute Inc.

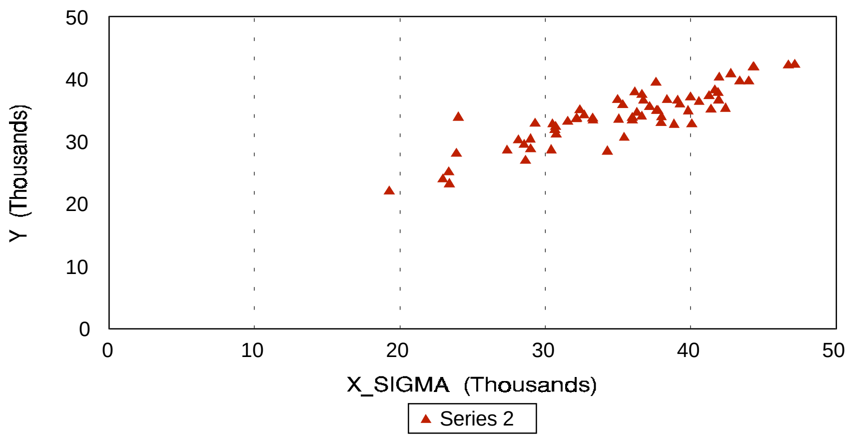
Graph 1

Loss as a function of **X**_SIGMA



Graph 2

P/D Stocks (Y) as a function of **X**_SIGMA



Note:
Heteroscedasticity is not always apparent from a graph. When the number of observations is small, it often appears to change substantially with minor changes to the data, as it does in this application. Also, calculated regression weights can be very volatile. (Further, when there are many observations, some points on a graph may be obscured by others.)

EXAMPLE COMPUTER CODE:

Following is some simple SAS code to implement equations 6.1, 6.2 and 6.3. That is followed by very simple FORTRAN code for implementing equations 7.1 and 7.2. This program may be improved upon, but did generate Table 1, when parts of the data file shown below this program file were removed for testing purposes. (When testing was done, the Q-value of the data record for the quarter treated as "current" was set according to the information supplied above, in section 9, ***"File maintenance for future prediction of producer/distributor coal stocks."***)

QCR.PDSTOCKS.PROGRAM -

```
//JK7UPRED JOB (6944,LEN,,30),'PREDICT',TIME=(1,55),REGION=6000K
/*ROUTE PRINT HST
/*PHOLD
//SASINY EXEC SAS,REGION=5000K,WORK='80,40',SORT=100
//IN  DD  DSN=JK76944.COAL.PDSTOCKS.DATAQCR,DISP=SHR
//OUT DD  DSN=JK76944.ASA.PRELIM.RESULTS1,DISP=SHR
//SASLOG   DD  SYSOUT=A
//SASLIST  DD  SYSOUT=A
//SYSIN   DD *
 OPTIONS  NOSOURCE  NONOTES  REPLACE LINESIZE=132;
 DATA SASREG;
  INFILE IN;
 INPUT  Q 1-1  X1 5-10  X2 15-20  Y 25-30
        X3 35-40  X4 45-50  X5 55-60  X6 65-70  YLAST 75-80;
 *LOSS = Y - (YLAST + X1 + X2 - X3 - X4 - X5 + X6);
  XT = X1 + X2 - X3 - X4 + YLAST - X5 + X6;
  W = 1/XT;
 IF Q EQ 0  THEN  W = 0;
 *PROC PLOT; *PLOT LOSS*XT / HZERO VZERO;
 *PROC PLOT; *PLOT Y*XT / HZERO VZERO;
```

```
PROC REG NOPRINT; MODEL Y=X1 X2 X3 X4 X5 X6 / P CLI;
    WEIGHT W;
       OUTPUT OUT=ODW P=YHPW STDI=YHSTPW;
DATA _NULL_; SET ODW;
 FILE OUT;
 PUT   @11 YHPW 10.2    @31 YHSTPW 10.2;

/*
//SASIXT  EXEC SAS,REGION=5000K,WORK='80,40',SORT=100
//IN DD   DSN=JK76944.COAL.PDSTOCKS.DATAQCR,DISP=SHR
//OUT DD  DSN=JK76944.ASA.PRELIM.RESULTS2,DISP=SHR
//SASLOG   DD  SYSOUT=A
//SASLIST  DD  SYSOUT=A
//SYSIN   DD *

OPTIONS  NOSOURCE  NONOTES  REPLACE LINESIZE=132;

DATA SASREG;
 INFILE IN;

INPUT  Q 1-1  X1 5-10  X2 15-20  Y 25-30
       X3 35-40  X4 45-50  X5 55-60  X6 65-70  YLAST 75-80;
LOSS = Y - (YLAST + X1 + X2 - X3 - X4 - X5 + X6);

XT = X1 + X2 - X3 - X4 + YLAST - X5 + X6;
IF Q EQ 1  THEN  W = 1/XT;
IF Q EQ 2  THEN
 W = (1/XT)*(1/(1+(0.63**2)));
IF Q EQ 3  THEN
 W = (1/XT)*(1/(1+(0.63**2)+(0.63**4)));
IF Q EQ 0  THEN  W = 0;
IF Q EQ 4  THEN  W = 0;


*PROC PLOT; *PLOT LOSS*XT / HZERO VZERO;
*PROC PLOT; *PLOT Y*XT / HZERO VZERO;

PROC REG; MODEL Y=XT / P CLI;
    WEIGHT W;
   *RESTRICT X1 = 1;
   OUTPUT OUT=ODW P=YHPW STDI=YHSTPW;
    *PROC PRINT;
```

```
*DATA ODW;*VAR YHSTPW; *PROC PRINT;
DATA _NULL_; SET ODW;
 FILE OUT;
 PUT   @11 YHPW 10.2    @31 YHSTPW 10.2;
/*
//SAS7NI    EXEC SAS,REGION=5000K,WORK='80,40',SORT=100
//IN DD   DSN=JK76944.COAL.PDSTOCKS.DATAQCR,DISP=SHR
//OUT DD  DSN=JK76944.ASA.PRELIM.RESULTS3,DISP=SHR
//SASLOG   DD  SYSOUT=A
//SASLIST  DD  SYSOUT=A
//SYSIN   DD *
OPTIONS  NOSOURCE  NONOTES  REPLACE LINESIZE=132;
DATA SASREG;
 INFILE IN;
INPUT  Q 1-1  X1 5-10  X2 15-20  Y 25-30
       X3 35-40  X4 45-50  X5 55-60  X6 65-70  YLAST 75-80;
LOSS = Y - (YLAST + X1 + X2 - X3 - X4 - X5 + X6);
XT = X1 + X2 - X3 - X4 + YLAST - X5 + X6;
IF Q EQ 1  THEN  W = 1/XT;
IF Q EQ 2  THEN
 W = (1/XT)*(1/(1+(0.76**2)));
IF Q EQ 3  THEN
 W = (1/XT)*(1/(1+(0.76**2)+(0.76**4)));
IF Q EQ 0  THEN  W = 0;
IF Q EQ 4  THEN  W = 0;
*PROC PLOT; *PLOT LOSS*XT / HZERO VZERO;
*PROC PLOT; *PLOT Y*XT / HZERO VZERO;
PROC REG; MODEL Y=X1 X2 X3 X4 X5 X6 YLAST / P NOINT CLI;
    WEIGHT W;
       OUTPUT OUT=ODW P=YHPW STDI=YHSTPW;
DATA _NULL_; SET ODW;
 FILE OUT;
 PUT   @11 YHPW 10.2    @31 YHSTPW 10.2;
```

```
/*
//FORSTR  EXEC FORT1CLG
//FORT.SYSIN DD *
      REAL*16 SPM1,SPM2,SPM3,YPM1,YPM2,YPM3,SC,YC,D
      INTEGER Y1,S1,Y2,S2,Y3,S3,YL,SL,Y,S
C
   11 READ(21,*,END=12)YPM1,SPM1
      GO TO 11
   12 READ(22,*,END=13)YPM2,SPM2
      GO TO 12
   13 READ(23,*,END=14)YPM3,SPM3
      GO TO 13
   14 D=(SPM1**2)*(SPM2**2)+(SPM1**2)*(SPM3**2)+(SPM2**2)*(SPM3**2)
      SC = SQRT(((SPM1**2)*(SPM2**2)*(SPM3**2))/D)
      YC =      (SPM2**2)*(SPM3**2)*YPM1
      YC = YC +(SPM1**2)*(SPM3**2)*YPM2
      YC = YC + (SPM1**2)*(SPM2**2)*YPM3
      YC = YC/D
C
C     WRITE(6,*)YC,SC
      Y1 = YPM1 + 0.5
      Y2 = YPM2 + 0.5
      Y3 = YPM3 + 0.5
      S1 = SPM1 + 0.5
      S2 = SPM2 + 0.5
      S3 = SPM3 + 0.5
      YL = YC + 0.5
      SL = SC + 0.5
      Y = (YC/1000.) + 0.5
      S = (SC/1000.) + 0.5
      WRITE(6,123)Y1,S1,Y2,S2,Y3,S3
  123 FORMAT(3(9X,I5,5X,I4),//)
      WRITE(6,234)YL,SL
  234 FORMAT(13X,I5,5X,I4,////)
      WRITE(6,345)Y,S
  345 FORMAT(13X,I5,2X,I4)
      STOP
      END
//GO.FT21F001  DD  DSN=JK76944.ASA.PRELIM.RESULTS1,
//    DISP=SHR
//GO.FT22F001  DD  DSN=JK76944.ASA.PRELIM.RESULTS2,
```

```
//    DISP=SHR
//GO.FT23F001  DD  DSN=JK76944.ASA.PRELIM.RESULTS3,
//    DISP=SHR
//*    DISP=SHR,LABEL=(,,,IN)
/*
```

DATA:

Here are the data, with revisions, as available in January 1998, as described in section 9 above. Data were being revised at the same time testing was being done, and more revisions are likely to be found for the data below in future publications of the Quarterly Coal Report.

COAL.PDSTOCKS.DATAQCR:

1.	1	135908	298	22265	169358	20514	158274	207340	23859
2.	1	234578	232	23417	193011	33978	164970	158274	22265
3.	1	238178	297	24149	182874	35773	185274	164970	23417
4.	1	222250	114	32995	186675	25564	179484	185274	24149
5.	1	216434	128	36876	164067	31621	198377	179484	32995
6.	1	201384	271	39678	184091	25249	189967	198377	36876
7.	1	198043	230	36784	172078	23842	195254	189967	39678
8.	1	192282	269	39867	173144	15144	192315	195254	36784
9.	1	186844	379	37761	165381	20345	197033	192315	39867
10.	1	198816	299	35249	205741	22045	173743	197033	37761
11.	1	204150	323	33931	192406	20238	168654	173743	35249
12.	1	223115	276	34265	199942	15125	174283	168654	33931
13.	1	230226	269	30841	185817	23737	194065	174283	34265
14.	1	244592	435	29701	208113	25263	208019	194065	30841
15.	1	197988	306	34090	197424	17357	197211	208019	29701
16.	1	213238	330	35371	205484	18544	179454	197211	34090
17.	1	228160	500	35197	193529	24208	188013	179454	35371
18.	1	222827	623	32632	213937	25831	176195	188013	35197
19.	1	219414	500	33133	205099	24097	170234	176195	32632
20.	1	227974	485	38024	206314	17245	166398	170234	33133
21.	1	220001	576	38148	188673	24170	176018	166398	38024
22.	1	218681	537	33804	212671	23687	164885	176018	38148
23.	1	223659	614	32093	196654	20416	175226	164885	33804
24.	1	222199	331	36560	199523	16576	173173	175226	32093
25.	1	218823	483	33939	199627	20113	176037	173173	36560
26.	1	232958	475	28775	229397	21033	165598	176037	33939
27.	1	244782	459	28321	208394	21885	185459	165598	28775
28.	1	236889	542	36764	220787	16061	175279	185459	28321
29.	1	226645	587	36079	205735	24900	173308	175279	36764
30.	1	241622	437	31360	238672	27691	154331	173308	36079
31.	1	245109	567	30418	218448	26371	158413	154331	31360
32.	1	247179	531	35508	223486	21429	149238	158413	30418
33.	1	239022	687	30598	208025	28445	159013	149238	35508
34.	1	243060	925	28848	232026	23991	147165	159013	30598
35.	1	251468	708	29000	226163	26949	146087	147165	28848

36.	1	264184	735	35099	217014	22383	160782	146087	29000
37.	1	254279	674	36895	211666	27733	173061	160782	35099
38.	1	254760	514	33659	240821	29497	161639	173061	36895
39.	1	255853	776	33418	225978	26191	168210	161639	33659
40.	1	254746	938	42162	219208	22318	171485	168210	33418
41.	1	237006	730	41054	208757	26214	173663	171485	42162
42.	1	251438	984	33628	236093	31197	163860	173663	41054
43.	1	252794	738	32971	223562	29239	167711	163860	33628
44.	1	255956	679	39853	220594	24731	168632	167711	32971
45.	1	242735	1043	40513	210037	27010	173270	168632	39853
46.	1	249055	882	35198	237698	26481	161878	173270	40513
47.	1	249799	1199	33993	224093	24294	163692	161878	35198
48.	1	243417	1213	38453	229165	18870	152619	163692	33993
49.	1	233750	1093	34827	214820	19946	154842	152619	38453
50.	1	227131	2142	27183	249872	18522	121909	154842	34827
51.	1	241127	2861	25284	232087	17181	120458	121909	27183
52.	1	255153	1850	34139	237596	14877	112278	120458	25284
53.	1	256964	1577	35758	223145	17940	126694	112278	34139
54.	1	260853	2304	32955	245820	19704	121225	126694	35758
55.	1	260535	1853	33219	223640	18838	136139	121225	32955
56.	1	266244	1795	42460	227695	18988	144004	136139	33219
57.	1	248613	1609	42104	217496	23184	151657	144004	42460
58.	1	257097	1725	36193	259415	22175	131739	151657	42104
59.	1	257782	2071	34444	236274	24201	134639	131739	36193
60.	1	259756	1713	36851	243360	20516	124760	134639	34444
61.	1	263397	1552	37344	229264	23039	134267	124760	36851
62.	1	272118	2071	33780	259657	23504	127595	134267	37344
63.	1	268585	1790	28648	251053	23414	123024	127595	33780
64.	1	273927	1331	37544	245813	20011	119847	123024	28648
65.	1	269701	1708	42529	232945	20603	128087	119847	37544
66.	1							128087	42529
67.	1								
68.	1			.					

(The last part of this file is to be completed as data become available.)

Appendix: Alternative Development of Equation (4.16)

Consider the general equation in the form

$$y_i \; = \; \hat{\boldsymbol{\beta}}^{\prime T} \boldsymbol{X}_i^{\prime} + b_y \, y_{i-1}^{(Q)} + w_i^{-1/2} e_{0_i} \quad , \qquad \text{Eq(A.1)}$$

a generalized version of Eq.(4.6), where Q is the number of periods since y was last observed, and therefore $y_{i-1}^{(Q)}$ is given the previous observed value of y_{i-1} if Q = 1, or is described by one or more regression equations when Q > 1.

If Q = 1, then

$$y_{i-1}^{(Q)} = y_{i-1} \quad , \qquad \text{Eq(A.2)}$$

and

$$\sigma^2_{y_i}(Q=1) \; = \; w_i^{-1} \sigma_e^{\,2} \qquad \text{Eq(A.3) and Eq(4.13)}$$

If Q = 2, then

$$y_{i-1}^{(Q)} \; = \; \hat{\boldsymbol{\beta}}^{\prime T} \boldsymbol{X}_{i-1}^{\prime} + b_y \, y_{i-2} + w_{i-1}^{-1/2} \, e_{0_{i-1}} \qquad \text{Eq(A.4)}$$

Substituting Eq(A.4) into Eq(A.1), for Q = 2 one has

$$y_i \;=\; \hat{\beta}^{T\,\prime} \boldsymbol{X}_i^{\prime} \;+\; b_y \; (\hat{\beta}^{T\,\prime} \boldsymbol{X}_{i-1}^{\prime} \;+\; b_y \; y_{i-2} \;+\; w_{i-1}^{\;-1/2} \; e_{0_{i-1}}) \;+\; w_i^{\;-1/2} e_{0_i}$$

Eq(A.5)

This is not a very useful format. However, since the component variances may be added, we have

$$\sigma^2_{\;y_i}(Q=2) \;=\; w_i^{\;-1} \sigma_e^{\;2} \;+\; b_y^{\;2} w_{i-1}^{\;-1} \; \sigma_e^{\;2}$$ Eq(A.6) and Eq(4.14)

and therefore (for Q = 2)

$$y_i \;=\; \hat{\beta}^{T\,\prime} \boldsymbol{X}_i^{\prime} \;+\; b_y \; \hat{y}_{i-1}^{\;(Q=2)} \;+\; e_{0_i} \sqrt{b_y^{\;2} \; w_{i-1}^{\;-1} \;+\; w_i^{\;-1}}$$ Eq(A.7)

Further, for Q = 3, one has

$$\sigma^2_{\;y_i}(Q=3) \;=\; w_i^{\;-1} \sigma_e^{\;2} \;+\; b_y^{\;2} w_{i-1}^{\;-1} \; \sigma_e^{\;2} \;+\; b_y^{\;4} w_{i-2}^{\;-1} \; \sigma_e^{\;2}$$ Eq(A.8)

and Eq(4.15)

So, in general,

$$\sigma^2_{\;y_i}(Q) \;=\; \sigma_e^{\;2} \sum_{q=1}^{Q} w^{\;-1}_{\;i+1-q} \; b_y^{\;2(q-1)}$$ Eq(A.9)

Therefore, Eq(A.1) can be rewritten as

$$y_i \;=\; \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}}\, \boldsymbol{X}_{\boldsymbol{i}}' \;+\; b_y\, \hat{y}_{i-1}^{(Q)} \;+\; e_{0\,i} \sqrt{\sum_{q=1}^{Q} w^{-1}{}_{i+1-q}\; b_y^{\,2(q-1)}} \qquad \text{Eq(A.10) and Eq(4.16)}$$

where $\hat{y}_{i-1}^{(Q=1)} = y_{i-1}$, $\quad \hat{y}_{i-1}^{(Q=2)} = \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}}\, \boldsymbol{X}_{i-1}' + b_y\, y_{i-2}$,

$$\hat{y}_{i-1}^{(Q=3)} = \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}}\, \boldsymbol{X}_{i-1}' + b_y\, \hat{y}_{i-2}^{(Q=2)} \quad , \ldots$$

$$\hat{y}_{i-1}^{(Q=\nu)} = \hat{\boldsymbol{\beta}}'^{\boldsymbol{T}}\, \boldsymbol{X}_{i-1}' + b_y\, \hat{y}_{i-\nu+1}^{(Q=\nu-1)}$$