



**Date:** April 20, 2005  
**To:** HELP Advisory Group  
**From:** M. Sami Khawaja  
**Re:** Division of Public Utilities “Quantec’s Utah HELP Program Evaluation Final Report” Memorandum

---

I offer the following comments regarding the draft DPU review of our report, presented in the order appearing in the memorandum.

1. PacifiCorp did not request that Quantec conduct the study, it was the DPU that requested that the Company “provide a study that establishes a comparison group as a way to properly attribute any benefits that accrue before making a final determination on the success or lack of success of the program. Therefore, the Division believes that the program should continue for another three years and include an evaluation program provided by Quantec.” (Year 3 Annual Report, page 4)
2. The Division finds our findings “less than compelling” because our analysis “fails to satisfactorily isolate HELP impact from other relevant economic variables.” I am not entirely sure why the Division reached such a conclusion. An argument is made later in the memo that we failed to isolate the impact of HELP from HEAT, but nowhere is an explanation given as to why the Division felt that our quasi experimental design failed to isolate the impacts from other relevant economic variables. We believe that our analysis did isolate the impact of HELP from these factors.
3. The Division states that the analysis did not provide any evidence of the changes in any of the performance measures. Again, we are at a loss. We do not feel that the change in arrears (at minimum) is questionable. The evidence was statistically sound that the program did have an impact on arrears accumulation. Later in the memo (page 10), the Division states that where we failed was primarily in the energy consumption, shutoffs, mobility and collections. While numerous, they are small in impacts. An extreme position maybe that the program has had no impact at all on these indicators. If removed from the analysis for lack of evidence (as suggested by the memo), the benefit cost ratios of the various tests will be: 0.80 for the HELP Only, 1.03 for the HEAT&HELP. Under these conditions, there is no distinction between the Ratepayers and the Societal tests. However, lack of

evidence does not mean that the impact is zero. Our studies for other programs have revealed evidence for impacts in these performance indicators. For HELP, we estimated the benefits of reduction in notices and mobility from a societal perspective to be \$2,230,246. What is needed to make the HELP only option cost-effective is only \$950,384 (over three period of the evaluation). In other words, if only 43% of our estimate turned to be accurate, the program is rendered cost effective. Put another way, we could be off by 57% and still have a cost effective program. If anything, we feel that our societal estimates of the cost of mobility are very conservative. Furthermore, our model does not take into account many other societal benefits (e.g., reduced homelessness, improved health, and decreased stress) for lack on monetized values.

4. The Division criticizes our coefficient of determination  $R^2$  estimated at 56%. Assuming that maximizing this value is equivalent to maximizing the quality of the model is dangerous. While overall fit is important, other factors are equally important (e.g., plausibility of the results, concurrency with economic theory, and statistical significance of the important explanatory variables). For example, in our own model, if the dependent variable were changed to “change in arrears” and the pre arrears term was not included as an explanatory variable, the coefficients of all the independent variables would have remained virtually intact and our estimate of the impact of HELP would have been almost identical. The  $R^2$  value would have dropped from 56% to 18%. This does not mean that the second model is less reliable than the first, especially if you are more interested in specific estimated coefficients. Furthermore, the 56% value indicates that there are other components that impact arrears that are not in the model. As long as these components are not correlated with the ones that are in the model, including them would only increase the  $R^2$ , but will not alter the conclusions. We are not accustomed to seeing  $R^2$  values much over 60% using these types of cross sectional/time series data. The only time we see higher values is in pure time series analysis we utilize for our forecasting clients. In those cases, the variables are likely autocorrelated causing the  $R^2$  values to be high.
5. Another criticism is related to the coefficient of variation (CV). The fact is that the average dependent value reported in the model is an aggregation of two “population” estimates. The \$96 average post arrears is composed of an average of \$147 for the comparison and \$47 for the treatment group. The CV then is an estimate of a pooled SD divided by a pooled mean. We are not certain this is a useful estimate. The CVs within each of the groups are likely to be smaller. Additionally, in the model rerun suggested in 4 above (i.e., making the dependent variable the change in arrears and removing the pre arrears from the model), the CV would have increased to 693% with still near identical results. More importantly, we feel, is the individual coefficients that make up the estimated \$78 contribution to HELP. Those are the coefficients of the delta credit,  $\Delta \text{latAgPay}$ , and the interaction terms. Those coefficients have t values of 55.05, -19.43, and

- 31.51. All t-values are statistically significant at the 99% level. We do not view these to be low confidence/precision values. However, the truth is that model building is more of an art than a science. Two econometricians can argue endlessly over many academic as well as practical matters. In the case of our model, our level of confidence increased significantly when we compared its findings to those based on simple averages. Using simple averages, we estimated the changes in arrears for participants that received HELP only to nonparticipants that received neither HELP nor HEAT and got nearly identical results. On average, participants with HELP only had a difference in arrears from the pre period to the post period equal to \$22 while the non participants who received neither HELP nor HEAT had a difference of -\$49, yielding a net difference in arrears of \$71 (compared to \$78 from our regression model).
6. Our biggest concern is with the Division's belief that our treatment and comparison group do not match. As we explained in the proposal and in the report, year 2 participants were used as treatment group for year 2 and comparison for year 1. Year 3 participants act as participants in year 3, but as comparison in years 1 and 2. In other words, the comparison and treatment groups are, more or less, the same customers in different time periods. Changes within households do occur across time, but our main goal was to ensure that the comparison group customers were low income. We matched *each* individual participant to *at least* one nonparticipant from future program years based on *pre arrears*. The Division compared them based on *pre invoice* and declared them dissimilar. If one has ten different measures describing a population, one does not have the luxury of matching on them all, not even two or three most relevant. We chose to match them based on pre arrears and did so very successfully (treatment pre arrears average was \$79 and comparison was \$81). While pre invoice is important, it was not our variable of choice. We opted for pre arrears because we felt it was most important in describing the households' level of ability to pay and to some degree level of burden.
  7. We believe that the DPU should also provide the commission with firm and specific recommendations as to what the program goals need to be and how to best measure them.

Finally, we thank the DPU for the opportunity to provide comments. We will be happy to discuss any or all of them at any time.