But is it True?:
Anticipating Evaluation Results for Planning

Fred Gordon, Energy Trust of Oregon Inc., Portland, OR
Ben Bronfman, Energy Trust of Oregon Inc., Portland, OR
Spencer Moersfelder, Energy Trust of Oregon Inc., Portland, OR
Charlotte Rollier, Energy Trust of Oregon, Inc., Portland, OR

ABSTRACT

Since its establishment in 2002, the Energy trust of Oregon, Inc., has implemented a wide range of comprehensive conservation and renewable energy programs. However, because the organization was required to implement programs quickly, original planning estimates were derived from existing sources and professional judgment. To integrate better assumptions as they became available, the Energy Trust developed a "true-up" process that incorporates new engineering knowledge, anticipated evaluation effects, and finally, evaluations, in an orderly annual cycle. This paper describes the process and presents some results of the "true-up" made to date. Currently, the process is in place for efficiency programs only.

Introduction

The Energy Trust of Oregon, Inc., (the Trust) is a non-profit corporation established in 2002 as part of utility restructuring legislation and charged with acquiring cost-effective conservation and renewable energy resources. The Trust receives funding through a three percent systems benefit charge on electric utility bills of customers of investor-owned utilities in Oregon (Portland General Electric and PacifiCorp) in the range of $40-$50 million per year. In 2003 the Trust was asked by the regional gas company (Northwest Natural) to take on gas efficiency measures in the residential and commercial sectors. Third-party contractors (Program Management Contractors) implement all programs. Currently, the Trust has comprehensive efficiency programs in existing and new buildings, in the residential, commercial and industrial sectors.

Early on, the Trust made a commitment to develop and implement a comprehensive program tracking database. The database (FastTrack) links customer and measure-specific information with utility usage information, a contact database and the Trust's financial system. The data system figures prominently in evaluation, reporting and planning activities. The Energy Trust also committed to using impact and market evaluations to provide the most accurate possible estimates of energy saved through efficiency programs and energy generated through renewable energy programs.

Discussion

The Energy Trust started with an objective of getting programs in the field quickly. Furthermore, the Energy Trust began at the end of a period when regional estimates of energy efficiency resource supply and cost were several years old and no longer very accurate. To get programs implemented quickly, we developed savings and cost estimates that were, in some cases, quite rough. As we staffed up and became more intimate with several technical issues, and as better information came from technical studies conducted by the Northwest Alliance, Northwest Power and Conservation Planning Council and others, it became clear that we were projecting savings on a basis that would not

hold up over time. Both the Energy Trust Board of Directors and the Public Utility Commission are interested in reviewing progress toward savings goals on a quarterly and annual basis and in assessing whether ten-year goals set by the organization (300 Average Megawatts and 19 million therms of energy savings, Oregon using 10% renewable power after ten year) were realistic. The obsolete information clearly did not serve these entities well in tracking the progress of the Energy Trust. Furthermore, while some impact evaluations became available in 2004, for several key programs they would not be available for another year.

This presented a problem: how to build new information (both evaluation and pre-evaluation) into projections in an orderly way. We needed to bridge the gap between "as installed" and "as evaluated" savings estimates, and fold in evaluations as they were completed (sometimes two years after the completion of a program year). A second problem arose, in that "as installed" estimates were developed to serve Energy Trust's need to contract with turnkey contractors for delivery of saving from programs. They generally did not factor in free riders or spillover. Furthermore, in some markets evaluations frequently reveal that engineering estimates of savings are overly optimistic. There was no adjustment for this consideration. Consequently, our estimates of savings were biased towards the high end. A third problem arose in that new information on cost and savings was coming in on an ongoing basis. A series of constant corrections to the basis for savings estimates would create chaos both for those personnel tracking savings and for Board members and PUC personnel trying to review our progress.

By early 2004, a NW Alliance retrospective evaluation provided information that drastically reduced our estimates of savings from compact fluorescent bulbs (fewer hours for more fixtures per home, interactive effects, revised estimate of change in wattage, changes in projected life, etc.). This prompted an ad-hoc correction to forecasts that served as the catalyst for planning a more systematic effort to correct estimated savings. Discussions with the Board revealed that they would rather have us apply preliminary data and judgment to provide our best "unbiased" estimates of savings, rather than withhold judgment but go with estimates that we thought were systematically high.

Developing this system necessarily involved the planning and evaluation functions at the Trust as well as program staff, communications staff and information technology staff. The system had to minimize confusion in reporting, synchronize with data systems, provide the appropriate reporting basis for program contracts, and also serve planning, forecasting, and evaluation needs.. The first attempt took most of a year. This is because we were inventing the system to some extent by implementing it, and we needed to educate both users and programmers as we went. The corporate database was being developed in parallel, and some of the early data corrections were both complex and dramatic. Going forward, the process will be repeated annually, with adjustments anticipated to be less onerous.

Planning and Reporting Structure

The Energy Trust's tracking systems count measures installed and assign cost and savings on that basis. The reporting function tracks past savings by program, fuel, and year. Information that is more detailed is available and analyzed for special purposes. The planning function uses historic data on savings to create trends, and it produces forecasts that reflect various strategic choices regarding program focus and direction, given the Energy Trust's limited financial resources. These form the basis for strategic decisions by management and the Board. Key factors include both the volume and unit cost of savings.

For residential programs, planning usually begins with cost and savings assumptions for individual measures, along with estimates of penetration of the measures in participating housing units and in the number of participating housing units. Savings are typically engineering-based and are weighted averages of typical measure installations. For business programs, since the measures and facilities are quite diverse in their character and size, forecasts are more often based on average cost per saved kWh and therm and on volumes of kWh and therms saved.

Structure of the True-up

For some residential programs, new information about potential unit savings and cost came in for specific equipment. In these cases, the "true-up" needed to work at the measure level. In other cases, it was unnecessary to work in any more detail than the fuel type and individual program for each year.

In sorting through the new information that came to the Energy Trust, we identified several different species, and developed specific adjustment factors for each. Rarely were more than one or two factors pertinent for a specific program or measure.

1. Corrections to "as installed" estimates of savings. In parallel to the true-up, the Fast Track data system was being completed and data were being imported from the individual data systems of our Program Management Contractors into this central system. This process was used to audit data and some corrections were implemented
2. New Data Factors. This is new information on unit savings, cost, measure life, or load shape that does not come from evaluation of Energy Trust programs. Some examples of new data factors follow:
    a. The need for corrections to CFL savings estimates was already mentioned. The Energy Trust pays for the installation of CFLs through several programs and promotional channels. Each has a different mix of heating fuels (influencing space heating take-back), likely install rate, early uninstall rate, and mix of technologies. Using a consistent spreadsheet format, a "trued up" estimate of savings was developed for CFLs delivered through each market channel. For example, a lower "install rate" was used for promotional giveaways than for discount coupons. The install rate was highest for programs where the Energy Trust pays for direct installation.
    b. Contractors initially assumed that whenever resistance heat was replaced with a high efficiency heat pump with an Energy Trust incentive, the Energy Trust program induced the change in space heating type, and the contractor took full credit for the efficiency gain in moving from resistance to heat pump, and then to high efficiency heat pump. Given that the incentive was $200 and heat pumps cost several thousand dollars , we suspected that in many cases the customer had decided to buy a heat pump and the Energy Trust incentive helped persuade the customer to buy a more efficient one. This assumption has a dramatic impact on projected savings. In the absence of data on what motivated the contractors, we came to an agreement with the contractor that perhaps half of the conversions to heat pumps were inspired by the program, which was incorporated into a "new data factor" to adjust the average unit savings. This assumption will be examined in the program process evaluation, which in turn may lead to further adjustments.
    c. At the time of program inception, the Energy Trust had very little data on the average gas use per home, and staff had limited experience with gas programs. Based on a supply curve study, we developed measures savings estimates based on typical household use that later information proved were much too high to be realistic. For gas weatherization

measures, revised estimates of savings were developed based on converting estimates for electric weatherization measures from the Regional Technical Forum (a fairly well-vetted data source) to gas savings, using a kWh/therm conversion factor and an assumed average gas furnace efficiency.

3. Anticipated Evaluation Factor. This factor is used to reflect the possibility that as-installed engineering estimates are too optimistic, and to factor in anticipated levels of free riders or free drivers. Generally, planning and evaluation staff developed these factors by reviewing results of prior evaluations of similar programs elsewhere, with consideration to the maturity and design of the Energy Trust programs. Generally, we assumed that evaluations would find that savings were 10 to 30 percent lower than "as installed' estimates of savings, depending on the type of program. In our experience, industrial programs generally have the highest realization rates (sometimes about 100% and sometimes in the range of 90%, so we used 90%) and realization rates for new commercial construction and weatherization measures for homes are often far lower. We chose 70% for both of these cases. In the case of the Home Energy Savings weatherization program, this was *after* applying new data adjustments.

   The most difficult case for the anticipated evaluation factor was programs that are anticipated to contribute to market transformation for commodities. These include new homes, new buildings, furnace, and appliance programs. In these cases, we applied a simplifying assumption that spillover would equal free riders, and made adjustments for neither. We are in the process of building models to forecast savings for these programs based on an explicit market transformation model that accounts for each of these factors over a period of several years. Once these models are complete, both evaluation and new data factors will be applied, although in a more complex matter, to reflect on unit savings and cost assumptions, market share assumptions, and net savings assumption.

4. Evaluation Factor. Factors #2 and 3 as applied to loads corrected for Factor #1, above, provide the "best" estimate of savings until an impact evaluation is completed. At that point, the impact evaluation provides a "realization ratio" which can be used to correct the "as installed" estimate of savings. Impact evaluations generally provide reliable savings estimates only for programs as a whole (by fuel) or for major clusters of measures, not for each individual measure.. Assuming the evaluation is successful and defensible this replaces items 2 and 3 above in producing the "best" estimate of savings. However, if New Data modified the basic engineering estimates of savings for specific measures, it is important to adjust the program savings to reflect the correct overall savings from the program, but in a way that reflects the *relative* measures savings incorporating the new data. Thus, the new data factor is incorporated for individual measures, then the overall savings estimate is scaled up or down to match the evaluation results. This provides the best picture of overall savings and measure savings.

So, in summary:

Trued Up Savings = "as installed" corrected for audited results

either

x New Data Factor x Anticipated Evaluation Factor

or

x Evaluation Factor as applied to incorporate new data in base estimates.

Arriving at Agreement on True-up Factors

Program contractors were very interested in new data adjustments because they formed the basis whereby the contractors could achieve program goals. The Energy Trust used independent expert consulting engineers to balance the interests of the Program Contractors and help arrive at agreement on appropriate new data factors.

Program contractors were not generally held accountable for "anticipated evaluation factors". This is because contractors generally could not discern a free rider when they participate in a program and could not practically withhold services from those customers, even if they could be identified. Furthermore, the hedging for optimistic saving estimates is speculative, and while it is useful to produce reasonable overall estimates, we did not hold contractors responsible for these factors either.

Thus, where evaluations were not yet available we faced the prospect of producing separate estimates of trued-up savings for program contract compliance (adjusted for new data factor but not for anticipated evaluation factor), and for reporting projections for the Board (adjusted for both). For some programs, we avoided this by getting program managers and contractors to agree to scale down goals in contracts to reflect anticipated evaluation factors. This allowed us to use one set of reporting estimates, avoiding significant confusion.

Establishing the Base

As previously mentioned, the "as installed" estimates of savings were being updated and corrected as the true-up was being completed, causing some delay and temporary confusion. In addition, we discovered that for the Home Energy Savings Program, methods of estimating "as installed" savings had evolved over the program history. This reflected both better communication of intent between the contractor and the Energy Trust, and some attempts to streamline and simplify costly house-by-house estimation procedures. So the base that the True-up was adjusting *from* needed to be reflected in adjustments.

Data System Integration

The true-up was first implemented as a series of spreadsheets. Later the summary conclusions were integrated into Fast Track. The purposes of the spreadsheet system were two: complete a true-up while FastTrack was being finalized, and provide a transparent-as-possible and thoroughly error-checked true-up to use as a check while the process was being automated in FastTrack. In future years, the process will occur largely in FastTrack.

- Analysis spreadsheets were used as a basis for negotiating engineering adjustments to estimates of measure savings for the Home Energy Savings and Efficient New Homes programs.[1] We expect that this step will be necessary to do in spreadsheets every year.
- For the Home Energy Savings program, for some measures, an adjustment needed to be applied to measures on a house-by-house basis to reflect the differing "as-installed' estimates of savings. These data were then summarized to create new data and anticipated evaluation factors by major housing type and fuel and year.

[1] New data for the Efficient New Homes program became available based on evolving regional modeling of savings as the contract for this program was being finalized. These were incorporated in the first set of savings estimates used for tracking and did not need to be "trued up" at the measure level.

- A summary true-up sheet applied each appropriate adjustment factor at the program level, using the summary of the Home Energy Savings measure sheet, and applying program-wide factors for all other programs.

Integration into our FastTrack data system proved more complicated than anticipated. Adjustments needed to be made to address several dichotomies:

- Is the adjustment for forward looking savings, backward looking savings, or both?
- Is the adjustment applicable for program goal tracking, reporting to the Board, or both?

The capability was developed to retain multiple versions of the savings data to reflect pre-adjustment and post-adjustment values. This is critical so evaluations, which in some cases occurred in parallel to the true-up, could provide realization rates and have the data set they started with available to adjust from.

Results

An Example: Single-family Electric Water Heaters 2004

To adjust 2004 savings for the Electric Water Heater measure[2], we queried an average reported savings number for each measure in the tracking system. Results are shown in Table 1.

A. In some instances, our initial estimates included now-obsolete estimates of savings associated with reduced losses from Transmission and Distribution. We developed a multiplier to remove these savings from this level of record keeping.
B. We adjusted the average savings number by removing savings associated with losses from T&D.
C. There are reduced savings for electric water heaters due to the increased Federal efficiency standard for water heaters. For our earlier programs, this needed to be factored in as a new estimate.
D. We calculated a multiplier to adjust from Average savings without T&D losses to our number derived from engineering analysis.
E. In this particular instance there is no Future Evaluation Factor for engineering because we had just completed an engineering analysis for the measure.
F. However, we did include a Future Evaluation Factor for free riders pending the upcoming evaluation.
G. We developed a multiplier to reflect the total adjustment resulting from the engineering reduction and the anticipated reductions from the pending evaluation.
H. We used this multiplier to recalculate the average savings numbers after removing T&D. For this particular measure we used this as a deemed savings number and multiplied this by the total number of units that have been installed.

For many of our programs we have not yet conducted an updated engineering analysis for implemented measures. Therefore, the adjustment process is limited to a future evaluation factor for engineering and a future evaluation factor for free riders.

[2] The letters below coincide with the column headings in Table 1.

Table 1: 2004 Savings Reductions: Electric Water Heater

	A	B	C	D	E	F	G	H	I	J
	Average Savings Number From Measures In The Tracking System (kWh)	Multiplier to Subtract T&D Adder (1-0.070)	Average Savings Number (kwh) After Removing T&D (A*B=C)	Prescriptive Savings Number (kWh) From Updated Engineering Analysis	New Data Factor (1-D/C=E)	Future Evaluation Factor - Engineering	Future Evaluation Factor - Free-Riders	Measure Adjustment After Adjusting to Subtract T&D Adder (1-(1-E)*(1-F-G)=I)	Total Average Adjusted Savings Per Measure (kWh) (C*(1-I)=J)	Total Measure Adjustment Multiplier (J/A=K)
Single-family Electric Water	243	0.93	226.05	128	0.43	0.00	0.22	0.56	100	0.41

Program Summary

The first true-up was successful completed in February 2005, and incorporated into our 2004 annual report. Table 2 below provides comparative information by sector for electric savings for 2004. The "*Corrected Tracking*" column incorporates a comprehensive analysis of savings data performed in early 2005, based on a review of invoices and corrections to data from Program Management Contractors that were made as data was entered into FastTrack. The column labeled "*Trued Up"* reflects changes to the Corrected Tracking figures.

Table 2: 2004 Electric Savings Compared to 2004 Action Plan: (Average MW)[3]

	Action Plan	Corrected Tracking	Trued Up	Percent of Plan
Residential	4	6.8	5.7	155%
Commercial/Industrial Facilities	6	5.5	5.6	96%
Industrial Process	17	11.2	10.2	60%
T&D Savings	3	2.6	2.3	N/A
Total	30	27	24	78%

Within the residential sector, the 2004 electric savings target was exceeded, primarily due to NW Energy Efficiency Alliance activity in the Trust service territory exceeding the forecast.[4] This is also a result of improved performance and strong delivery efforts by program staff and contractors and dividends accrued through the initial development of the trade ally program marketing network. Conversely, the rollout for the Efficient New Homes program occurred later than anticipated due to the need to revisit cost-effectiveness assumptions, effectively delaying most savings initially projected in 2004 until 2005. Early program implementation can be looked at as being like a flywheel, with much effort and resources required at the outset to build the infrastructure and foundation for sustained delivery. Once the mass is at full speed, less effort is needed and more results can be readily supported and achieved.

---

[3] The measure "Average MW" is a convention used in the Pacific Northwest reflecting the reliance on a hydro system for baseload generation. It is calculated as ((kWh/1000)/8760).

[4] The Northwest Energy Efficiency Alliance is a regional market transformation organization. The Energy Trust supports the Alliance financially, and takes credit for market transformation effects in its service territory.

Electric savings from efficiency improvements to commercial and industrial facilities are slightly less than forecasted in 2004 primarily because projects within the New Building Efficiency program and for NW Alliance programs were completed slower than forecast. However, Building Efficiency (existing buildings) and LED traffic signal programs exceeded goals in 2004. Already in early 2005, commitments to future projects in commercial and industrial facilities are significantly accelerating.

Industrial process energy savings are significantly less than projected because project schedules for larger efforts, in particular, required longer completion cycles than expected and because there were fewer "very large projects" than anticipated. The overall volume of 2004 program activity in the Production Efficiency program, when measured in projected savings, was less than forecast, raising questions about whether initial projections are realistic over the long term. The targets appear plausible only if more very large and very low-cost projects are identified.

Next Steps

The true-up process is in place and is being integrated into the program data tracking system (FastTrack). The system will undergo an annual "true-up", with the next one scheduled for the first quarter of 2006. For that true-up, major impact evaluations for residential, commercial and industrial retrofit programs will be completed, and results will be integrated for 2003 and 2004 program years. Also, decisions will be made as to whether those results will be applied to 2005 programs, or whether the 2005 program estimates will hold until those specific evaluations are completed.

Evaluations are designed so that – as far as possible – results can be "seamlessly" integrated into the tracking system. Where possible, evaluations will provide specific savings and free ridership estimates by major class of efficiency measure. Because of this priority, evaluation contractors are discouraged from developing evaluation models that estimate net impacts directly.

Where measure-specific estimates are not possible, end-use results will be applied. In those cases where even end-use results are problematic, program-level adjustments will be made. For the residential retrofit program (Home Energy Savings), a preliminary assessment of the program tracking database reveals sample sizes in the hundreds for single major measures such as furnaces, heat pumps, ceiling and wall insulation, windows and duct sealing. Commercial retrofit results are much more likely to be at the end-use level – lighting and mechanical (HVAC). Industrial sector evaluation is site-specific, and it more likely will result in a program-level adjustment, until a "library" of measure specific results is developed. The true-up will also be able to expand to include new evaluation components. This round of evaluation will include estimates of program spillover, both project (customer) specific and "current practice" (from surveys of trade allies). The next true-up will include these estimates, if appropriate.

In summary, the Energy Trust has developed a systematic approach to "true-up" program performance results on an annual basis. True-up analyses are derived from a variety of information sources, including:

1. *New data on measure performance and more detailed measure simulations*. This approach particularly impacted the initial savings projections for the Home Energy Savings Program.

2. *Anticipating Evaluation Results***.** Experience shows that evaluated estimates are often lower than tracking estimates based on an "as installed" engineering analysis. However, to have enough

comparative post-installation energy use data available, impact evaluations are undertaken eight months to two years after a year of program activity has been completed.

3. *Evaluations***.** When finalized, evaluations provide the most reliable representation of realized savings and replace the refined projections identified in items 1 and 2 above.