

Prepared by:
Proctor Engineering Group, Ltd.
San Rafael, CA 94901
(415) 451-2480

ROLE OF BILLING ANALYSIS IN DSM EVALUATION

Proceedings of Energy Program Evaluation:
Uses, Methods and Results

August 25 - 27, 1993
Chicago, Illinois

Contributors:
Michael Blasnik, GRASP
Margaret Fels, Princeton University
Roger Wright, RLW Analytics, Inc.
Miriam Goldberg, XENERGY, Inc.

Reprinted by:
Proctor Engineering Group

Creators of CheckMe!®



PANEL 5G

ROLE OF BILLING ANALYSIS IN DSM EVALUATION

Moderator: Kenneth Keating, Bonneville Power Administration

PANELISTS:

- Margaret Fels, Center for Energy and Environmental Studies, Princeton University
- Michael Blasnik, GRASP
- Roger Wright, RLW Analytics, Inc.
- Miriam Goldberg, XENERGY, Inc.
- Philip Hummel, Electric Power Research Institute

Summary

The extent to which billing analysis is used in DSM program evaluations varies widely. Billing analysis—the mainstay of residential energy conservation program evaluations in the 1980s—offers the advantage of requiring only readily available data, and provides meter-based estimates of whole-building energy savings at the individual-house or individual-building level. Its feasibility has been well demonstrated by widespread application of a particular billing analysis tool, PRISM (PRinceton Scorekeeping Method). Nevertheless, DSM evaluation presents new challenges, in terms of an increasing demand for savings estimates at the end-use level and a growing concern about self-selection bias and free-ridership issues. Given these challenges, the desirability of measuring whole-building energy savings in addition to end-use savings becomes a key issue that warrants consideration by utility planners and other parties interested in DSM program design.

In this panel, potential roles of billing analysis in DSM evaluation are explored by panelists with a wide range of experiences and opinions. Examples of questions addressed by the panel are:

- How should whole-building billing analysis be used in conjunction with end-use specific approaches? Can whole-building savings from metered billing data provide a useful calibration for engineering estimates?
- How are decisions made concerning the use of billing analysis in an evaluation?
- Are there some types of programs for which billing analysis is not relevant or appropriate?
- On the other extreme, when should estimation of whole-building savings be required independent of other methods used?
- To what extent does billing analysis offer the desired “transparency” feature in an evaluation?
- When estimates of whole-building savings and end-use savings are inconsistent, how should this inconsistency be reconciled (by evaluators, by utility planners)?
- How can utilities make billing information more easily accessible to evaluators? How can data preparation for billing analysis be simplified?

Margaret F. Fels, Princeton University

Based on readily available data, billing analysis is appealing because of its widespread applicability: Using a quasi-experimental design approach, it can be applied to nearly all customers participating in a DSM program, as well as to a sample of buildings selected as a comparison group, to compare distributions of weather-adjusted whole-building energy savings in the treatment and comparison groups and thus to determine net savings. Furthermore, the desired feature of transparency is ensured in simple billing analysis, wherein the savings estimates remain closely related to the original metered data and thus to the true consumption picture.

In the face of pressures from newer DSM programs to estimate end-use specific savings and to answer difficult questions about attribution, evaluators are trying varied and elaborate approaches. One possible outcome is the loss of transparency as well as the loss of comparability across programs. Although in theory many of these methods could provide the desired levels of detail, in practice they often rely on opportunity samples of metered data (both in terms of small subsamples of building and short periods of time), filling in the gaps with engineering estimates that experience has shown may not reflect reality.

Even the staunchest billing analysis advocates would not argue that whole-building billing analysis can meet all the needs of DSM program evaluation. Nevertheless, as evaluators invent new end-use specific approaches, and even as the reliability of the estimation methods increases, the enormous resource offered by billing data should not be ignored. A comparison with whole-building savings derived from actual consumption data can bolster the end-use specific estimates. Or, if they disagree, the comparison can help to shed light on whether the end-use savings do not translate to whole-building (and ultimately to system-wide) savings or whether flawed assumptions in the methodologies need to be addressed. Billing analysis need not be used at the exclusion of other approaches, but, rather, in creative combination with them to add strength and confidence to the evaluation.

Michael Blasnik, GRASP

The use of billing data to estimate energy savings from conservation programs has been transformed by the recent growth in DSM programs and utility cost-recovery mechanisms which depend on "measured" savings. It appears that, in the eyes of many leaders in the DSM field, the relatively simple and obvious approach of comparing pre-treatment and post-treatment usage (with appropriate normalization as needed) for each participating building and a sample of comparable nonparticipants is no longer an acceptable approach, nor even a desirable component, for program evaluation.

Instead, billing data must be analyzed with grand multiple regression models that supposedly control for all extraneous factors. A particular coefficient is proclaimed as the true program savings, baffling most readers into silent acceptance. While such models may sometimes prove useful, they are often inappropriate, unstable, and fail to satisfy the assumptions which they rely upon. In the hands of less skilled (or less scrupulous) evaluators, these models can easily give misleading or biased results due to common problems such as model misspecification, outliers, undue influence, multicollinearity, and heteroscedasticity. I believe that evaluations based on coefficients of multiple regression models need to include critiques of the models selected and should provide sufficient information for an informed reader to assess the model for themselves. Additionally, the results of simpler analyses should be provided for comparison. Unfortunately, these principles are rarely followed in current "state-of-the-art" DSM evaluations, leading to results that are inscrutable and therefore questionable.

Roger Wright, RLW Analytics, Inc.

The greatest challenges in impact evaluation, especially for C&I programs, are to measure: (1) savings by measure category and market, (2) net vs. gross savings, and (3) persistence. The objective is unbiased, reliable, and cost-effective estimates.

Progress will be made by addressing many different technical problems, including the following:

- Developing a combined heating/cooling version of PRISM.
- Combining survey data with billing analysis and tracking data.
- Combining short-term end-use metering with billing analysis.
- Dealing with model specification and heteroscedasticity.

The key questions are (1) which evaluation methods are most appropriate in particular circumstances, and (2) can programs be better designed to yield measurable savings?

Miriam Goldberg, XENERGY, Inc.

The purpose of DSM programs is to reduce the demand for energy, both instantaneously (kW) and cumulatively (kWh). If a program is effective, the reduction should be reflected in customer billing records. However, the billing records alone cannot tell the whole story of a program, but an impact estimate that is inconsistent with the evidence from the bills will generally not be considered credible.

There are several challenges to interpreting the evidence from the bills. Except in occasional statistically designed pilot studies, there is no perfect comparison group. Some kind of model must be used to control for systematic differences between participants and comparison customers. The use of these models, however, often raises issues of credibility, transparency, and the sensitivity of results to the particular model specification used. In addition, demand (kW) savings are not directly measured in the bills, but must be inferred from the energy savings or estimated by separate means. Billing analysis by itself typically does not provide explicit estimates of snapback, measure degradation, or free-ridership—only their combined or net effects. In addition, some customer segments are inaccessible to billing analysis, due to incomplete billing records.

Because of these and other limitations, billing analysis increasingly involves integration with other data sources, such as customer surveys and audits, and other estimates such as *a priori* engineering estimates. Results from special metering studies can be linked to billing analysis through model calibration or ratio estimation.

One of the great strengths of billing analysis is that it can readily be conducted for large fractions of customers, both participating and nonparticipating. When billing information is used in conjunction with more specialized or rare data, the value of both is enhanced.