

Sweet Substitutes for Sour Research Data

Panel Session

Sponsored by the AESP Market Research & Evaluation Topic Committee

Chair of Panel and 2003 AESP Market Research & Evaluation Topic Committee

Chair: Lori Megdal, Ph.D. Megdal & Associates

Evaluators and researchers can be faced with smaller elements of their work where the data desired is unobtainable or far too expensive to collect. Developing and using proxies or substitute data/methods is one of the most creative areas within evaluation and market research. Yet, as a creative part of this work, it can not readily be taught. However, there is significant value in hearing about how others may have solved similar problems that can allow practitioners to recognize potential solutions when they run into something similar or provide ways to think about finding proxies that can help lead them to developing their own creative solutions. This unusual panel and audience discussion will provide a unique experience to help foster improvement and further development in this area.

The panelists will provide practical examples of where substitutes were made for data that was unavailable or too expensive to obtain that worked well for the evaluation or market research effort. Each panelist will be providing one specific practical application of proxy or substitute data that they think were particularly good (good fit and much more cost-effective/practical than the initially desired data). Then an open discussion with the audience will be held to maximize the benefit to attendees of finding alternative ways to handle the problems they may be experiencing in this area, based upon the experience of the panelists and other members of the audience.

Panelists: Angela Jones, Vanward Consulting and Mary O'Drain, Pacific Gas & Electric Company

Using Census Data to Estimate A Current-Year Joint Distribution of Household Income and Household Size

The problem: A current-year estimate of the joint distribution of household income and household size was needed to estimate current-year eligibility for California's Statewide CARE program, a rate discount program for low income customers. PUMS data from the Census provides the right data but is not a current-year estimate. Current-year estimates can be obtained from data vendors but only of the separate distributions.

The solution: The most recent PUMS data is used to create a joint distribution of household income and household size from the current-year estimates of the separate

distributions of household income and household size obtained from data vendors, using iterative proportional fitting.

The basic method used for creating a current-year estimate of the joint distribution of household income and household size will be discussed including mention of an additional data source that is suitable to serve in PUMS' place for this method.

Angela Jones, Vanward Consulting, will be presenting the statistical methods discussed above.

Mary O'Drain, Pacific Gas & Electric Company, will make a presentation in the use of these types of analyses.

Panelist: Shel Feldman, Ph.D., Shel Feldman Management Consulting
Careful Creation of Panel Design and Scenario Development to Create Defensible Data from a Delphi Panel

The use and value of Delphi panels to develop market penetration projections under conditions of changing technologies, standards, program designs, and other uncertainties will be discussed. The presentation will focus on aspects of panel design and scenario development crucial to obtaining a defensible set of projections, and on the importance of forcing the explication of assumptions held by program sponsors and by panelists. In describing both strengths and drawbacks of the method, Dr. Feldman will contrast it with other methods for projecting the information required by program planners and policy makers.

Panelist: Brenda Weeks, Tennessee Valley Authority
Using Proxy Data to Estimate Saturations

An example of the use of proxy data in the process of estimating saturations as conducted by the Tennessee Valley Authority will be presented by Brenda Weeks.

Panelist: Lori Megdal, Ph.D., Megdal & Associates
Useful Examples Using Ratios and Comparisons within Billing Data to Proxy Unmeasured Effects within Billing Analysis

There can be many hypotheses concerning why people or firms differ in their energy usage patterns. Sometimes these crop up as billing analysis is being conducted when the research design is already finalized whereby gathering primary data to include in the analysis is not feasible. This can occur when customer surveys are not possible in the budget, already concluded as part of earlier phasing of the data collection tasks, or where the survey effort is based on a smaller sampling method due to budget constraints will the billing analysis is based upon a previously adopted large inclusive evaluation philosophy. These varied circumstances can often lead to times when additional data would be greatly desired.

The presentation will provide several examples where ratios and comparisons from within the billing data were used to test hypotheses concerning different usage patterns (and to help obtain models that better fit the full sample of billing data) and provide more defensible billing analysis results. These examples include proxies for effects of economic trends within larger businesses, differing sizes of water heaters, snowbirds, and others.

Panelist: Ed Smyth, RLW Analytics, Inc.

Use of U.S. Census Data from Selected Metropolitan Statistical Areas (“MSAs”) in the 2002 Oncor Baseline Assessment for Energy Star® Windows

In order to have statistical relevance and a cost effective effort, this project constructed a study design maximizing the use of secondary data. In conducting this market research, a basic geographic perimeter was established that matched the kinds of data parameters normally established in secondary data, such as cities, counties, states, or regions. In this particular project, the area of concern (the Oncor service area) covers various, non-contiguous portions of the state of Texas. Within those portions the distribution areas cover over partial areas of various counties and towns. A perfect statistical representation of data that only relates to homes strictly within the Oncor service area would create primary research tasks that would be onerous, time consuming, and ineffective in cost.

In order to maintain a workable template of data that can be cross-analyzed with other secondary data, RLW confined all sampling and analysis tasks work within the Metropolitan Statistical Areas (MSA) that captures the bulk of the major population centers in the Oncor service territory. These are the Odessa-Midland MSA, the Forth Worth-Arlington MSA, the Dallas MSA (the two latter grouped together as the greater Dallas-Fort Worth Consolidated MSA), and the Killeen-Temple-Waco MSA (Figure 1, below). Use of the MSA based methodology enabled the RLW team to use various data sources to cleanly and accurately extrapolate the sampling data to the larger population.

The retailer sample design required the interviews of 74 retailers in all three target MSAs. The sample was required to include representations from the smaller MSAs of Killeen-Temple-Waco and Odessa-Midland. It was known from previous studies, and Oncor new connect data that 85% of the new construction was within the Dallas-Fort Worth MSA¹. Therefore, 85% of the sample was devoted to that MSA and the remaining 15% were assigned to the smaller MSAs. The final contact list contained 60 sites in the Dallas-Fort Worth MSA and the remaining 14 sites were allocated evenly between Killeen-Temple-Waco and Odessa-Midland. The list of identified retailers in these MSAs was randomized to allow the creation of a contact list devoid of bias. The percentage of big box retailers was a product of this random selection process.

¹ TXU Electric and Gas Residential New Construction Baseline Assessment, RLW Analytics, Inc, January 2001.

The first step in determining the installed volumes of different window types was to determine the total area of installed windows for the subject year, in the Oncor service territory. This was evaluated using the Oncor new connection data for 2001 along with other information known about the residential home market in the region. The connection data provided the number of single family and multi-family homes receiving new Oncor service. Table 1 presents the findings from the new connect data for the target Metropolitan Service Areas (MSA) of Dallas-Fort Worth, Killeen-Temple-Waco, and Odessa-Midland. These MSAs are the proxy for the overall Oncor service territory. This table indicates that 76% of the new connections are new single family homes, and 24% are multi-family units.

| New Residents in Target Metropolitan Service Areas | Oncor New Connection Data | | | |
|---|---------------------------|------------|--------------|------------|
| | Single Family | | Multi-Family | |
| | Number | % of total | Number | % of total |
| Dallas | 20,311 | 74% | 7,146 | 26% |
| Fort Worth | 13,161 | 81% | 3,162 | 19% |
| Killeen-Temple-Waco | 3,258 | 79% | 843 | 21% |
| Odessa-Midland | 621 | 64% | 355 | 36% |
| Total | 37,351 | 76% | 11,506 | 24% |

Table 1: Oncor New Connections

Table 2 below compares the new connection data of the target MSAs with the total new connection data for the entire Oncor service territory. The percentage of new single family homes (76.4%) in the target MSAs is very close to percentage for the entire Oncor territory. These target MSAs include 80% of all new residential construction within the Oncor territory, ensuring that the data collected from these areas provide an accurate representation of the activity throughout the Oncor territory.

| Target MSAs Compared to Oncor Totals | Oncor New Connection Data | | | | |
|--|---------------------------|------------|--------------|------------|-------------------|
| | Single Family | | Multi-Family | | Total Connects |
| | Number | % of total | Number | % of total | |
| Target MSAs | 37,351 | 76.4% | 11,506 | 23.6% | 48,857 |
| Entire Oncor Service Area | 46,554 | 75.9% | 14,768 | 24.1% | 61,322 |
| Target MSAs as % of Oncor Total Area | 80.2% | | 77.9% | | 79.7% |

Table 2: Target MSAs Compared to Oncor Totals

This approach of using these MSAs as proxy footprints to the actual utility service area allows future use of this baseline study for further cross-tabulation or other comparative analysis of the results against any residential or new home construction data reported in the US Census. This makes the final delivered study of continued future value to Oncor beyond the immediate report.

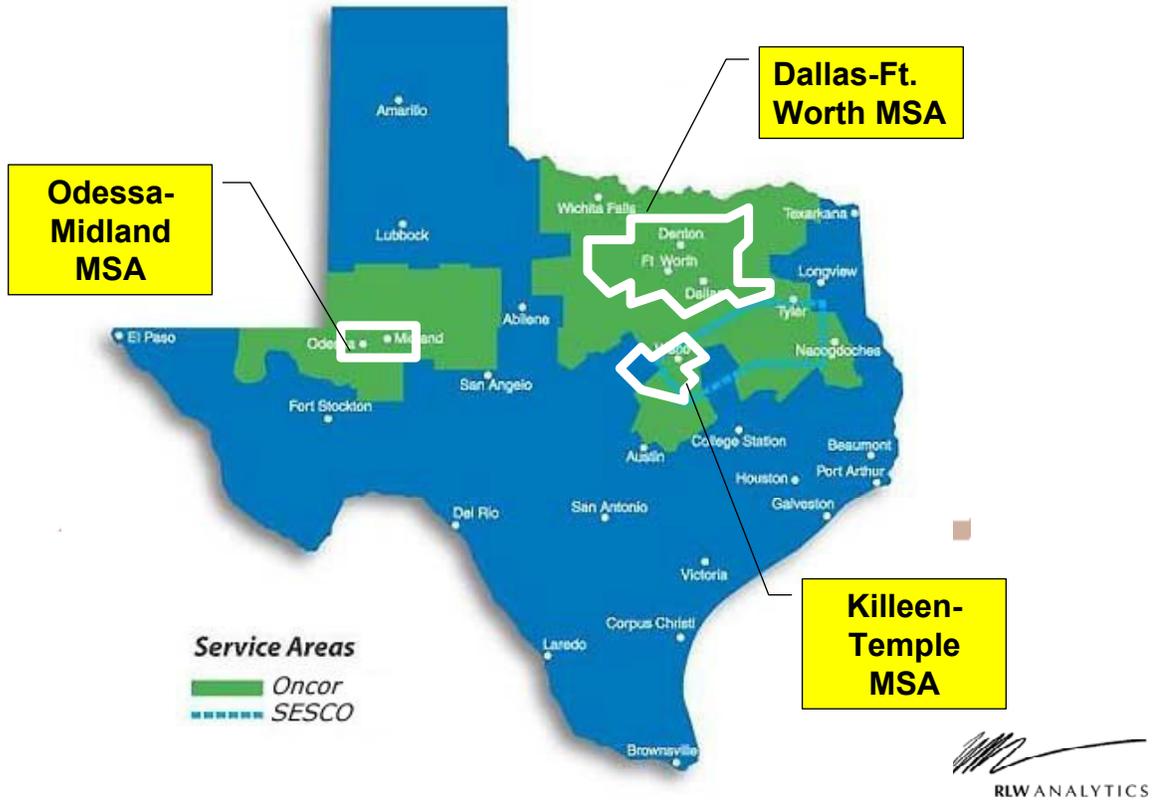


Figure 1: Oncor Service Area with Corresponding MSAs