AN IMPACT EVALUATION OF A MULTIFAMILY GAS CONSERVATION LOAN PROGRAM IN CHICAGO: METHODS AND RESULTS

Anne Graham, Scott Pigg, and Jeff Schlegel
Wisconsin Energy Conservation Corporation
Madison, Wisconsin

Peter Harrison
The Peoples Gas Light and Coke Company
Chicago, Illinois

Abstract

Multifamily energy conservation programs have not traditionally received as much evaluation attention as single family programs have. Assessing the success of multifamily energy efficiency programs through billing analysis has been problematic because of the variation in building size and characteristics, and particularly in the multitude of possible meter configurations. This paper presents methods and results of a multifamily evaluation in which these variables were effectively managed with respect to an impact evaluation funded by Peoples Gas Light and Coke Company. The evaluation, completed in April 1991, looks at 1987 and 1988 participants in the Chicago Energy Savers Fund's Multifamily Program. The evaluation revolves around a billing analysis performed on a mix of master-metered and individually-metered multifamily buildings with gas accounts in Chicago.

Introduction

Between 1984 and 1989, The Peoples Gas Light and Coke Company (henceforth referred to as Peoples) operated a multifamily loan program through the Residential Energy Conservation Loan Fund, also known as the Chicago Energy Saver's Fund. The program provided reduced-interest loans for energy conservation investments in multifamily buildings with an emphasis on low- and moderate-income multifamily housing. Approved energy conservation measures (ECMs) installed in buildings included heating system replacements and retrofits, building shell insulation, storm and replacement windows, and water heating measures. Loans also covered some building rehabilitation and repair expenses.

The Conservation Loan Fund subcontracted with the Community Investment Corporation (CIC) to administer the multifamily program. CIC contracted with the Center for Neighborhood Technology (CNT) and other community-based organizations (CBOs) in Chicago to deliver the program.

Loans were available for up to $98,000 per building or $3,000 per unit, and could be combined with other sources of funding, including rehabilitation funds. The average loan was about $28,000. During its period of operation, the program provided over $9 million in lending for more than 7,500 dwelling units in 316 multifamily buildings of five or more units.

The program was discontinued in 1989 after Peoples submitted a cost-effectiveness report to the Illinois Commerce Commission (ICC). In the report, the multifamily program was found to be not cost-effective using the tests prescribed by the ICC (Case C' Benefit-Cost Ratio[1] of 0.75). Those findings were based on an internal evaluation of the program examining three years of program operation.

After an intensive program design process, the Peoples Modified Multifamily Program was implemented in April 1990. As part of the process of implementing the new program, a more recent evaluation of the old program was needed to gain further insight into the impact and cost-effectiveness of the program and the performance of individual ECMs.[2] The results of the evaluation are being used to fine-tune the new program to ensure its cost-effectiveness.

Goals

The purpose of the evaluation was to determine the energy savings resulting from the installation of ECMs in multifamily buildings participating in the program. The measured energy savings can be viewed and used in several ways. Five specific goals of the evaluation were:

1. Determine the total energy savings of the program.

2. Evaluate overall program cost-effectiveness.

3. Analyze ECM installation frequencies and expenditures.

4. Evaluate the impact and cost-effectiveness of ECM categories.

5. Analyze cooking gas consumption to see if the program had an effect on the potentially dangerous practice of using ranges for supplemental space heating.

Methodologies

Study Design

The evaluation was designed to use a sample of buildings that were treated under the multifamily program. We compared pre- and post-treatment gas consumption using data obtained from Peoples' main customer file. A comparison group of similar buildings was used to control for other factors that may have influenced energy consumption.

Sample Selection

Treatment buildings. The initial treatment sample consisted of all participating buildings with loan closeout dates between July 1987 and December 1988. As the Program Administrator, CIC was the repository for program data files. We gathered data for each building from CIC's program records, including building address, number of dwelling units, installed ECMs, ECM installation costs, total loan amount, and program participation dates. Buildings in the treatment sample took an average of 410 days to complete the program stages from loan application to loan closeout. A list of 104 participating buildings was compiled.

Comparison buildings. For each treatment building, two comparison buildings were selected to ensure a sufficiently large sample containing high quality data. Provided with the list of addresses and number of dwelling units in each treatment building, Peoples staff selected two buildings with a comparable number of units as close as possible in geographic location to the corresponding treatment building. This selection procedure helped ensure that each comparison building was as close in gas consumption features as possible to its corresponding treatment building. The comparison sample was screened to ensure that these comparison buildings had not participated in the multifamily program. A list of 208 comparison buildings was assembled.

Data Retrieval

For the 312 addresses on the list, Peoples downloaded consumption data from their customer main files dated between January 1985 and May 1990. The data sent to WECC contained gas consumption information for every account in every building on the list; account types included all combinations of individually metered heating, cooking, and master-metered accounts. Each gas account contained the following data: account and service pipe numbers, building address, meter read type, billing dates, therms used, and appliance codes.

We were fortunate that the consumption records kept by Peoples included a field indicating which appliances each gas account serviced. This allowed us to separate accounts for the different analyses on heating and cooking gas consumption. The field was eight characters wide, with each character representing a gas end-use, such as water heater, central heating system, and cooking range. Each character of the field was either "1" or "0," representing that the appliance was either serviced by gas or not. All combinations of appliances were present in the accounts analyzed.

Initial and Final Sample Sizes and Sources of Attrition (Heating Analysis)

The original data request made to Peoples contained many more accounts than eventually were used in the evaluation because of (expected) attrition due to bad or incomplete data. Gas consumption data for all accounts in 104 treatment buildings and 208 comparison buildings were received from Peoples. Five treatment buildings were dropped because of insufficient background information about a building. On examining the consumption data, 34 treatment and 59 comparison buildings were dropped because of insufficient pre- or post-treatment consumption data. A few treatment buildings were unoccupied in the pre-treatment period, causing them to be dropped from the analysis.

In addition, 89 comparison buildings were dropped to create a one-to-one match between treatment and comparison buildings. For each treatment building, one comparison building was randomly selected from the two available. In cases where one comparison building was already lost (due to the above reasons), the remaining comparison building was used.

Of the remaining 65 treatment buildings comprising the sample, five buildings did not have matching comparison buildings. These five treatment buildings were very large, with between 77 and 163 dwelling units. The comparison sample was made up of 48-unit or smaller buildings (in the entire sample of 208 comparison buildings, only one building had more than 48 dwelling units). Because the five large buildings lacked comparison buildings, as well as other differences between the five very large buildings and the 60 smaller buildings, we decided do some separate analyses of the two groups. Sections of the evaluation compared the two building sub-groups in some detail.

Preparation and Weather Normalization

Preparation for the energy savings analysis began with a careful building-by-building screening of the end-use appliance codes to verify that accounts were correctly coded. Looking at each building's annualized consumption, we ensured that all high consumption accounts were flagged as heating accounts. This preliminary analysis uncovered a few accounts whose appliance codes were suspicious when compared to their consumption levels. Generally, there was a wide difference in annualized consumption between accounts coded for space heating and those coded for other gas end-uses. A few accounts (four) with very high consumption were re-coded as space heating.

All accounts with an appliance code indicating space heating were chosen for use in the energy savings and cost-effectiveness parts of the evaluation. Pre- and post-treatment periods were assigned to each account. All consumption data before the loan application date were used for analyzing pre-treatment consumption, and all consumption data after the loan closeout date were used for analyzing the post-treatment period. Comparison buildings were assigned the same pre- and post-dates as their matched treatment building. Thus every building had its own treatment period, but the distribution of treatment periods between the treatment and comparison groups were the same.

We weather-normalized the data using the same variable degree-day model of gas consumption employed in the Princeton Scorekeeping Method (PRISM).[3] The model predicts gas consumption as a function of heating degree days, where the reference temperature for calculating heating degree days is individually estimated for each account. Degree days were calculated using average daily temperature data for O'Hare Field in Chicago. Normals were based on a 32-year period from 1958 through 1989. Consumption data were consolidated into periods between actual meter reads. The above model requires at least four actual consumption periods.

For buildings with multiple heating accounts, we were concerned that account attrition might lead to under-representation of a building's total consumption. We therefore compared the number of accounts in the final sample to the original number for each building. Buildings whose accounts in the final sample totalled less than 90% of the original (un-normalized) annual consumption were dropped from the sample. Thirty-four buildings in the treatment group and 59 buildings in the comparison group were dropped for this reason. In addition, 89 comparison buildings were dropped to create a one-to-one match between treatment and comparison buildings. For each treatment building, one comparison building was randomly selected from the two available. In cases where one comparison building was already lost (due to the above reasons), the remaining comparison building was used. In the final sample, the majority of the buildings had a single heating account (59 of 65 for the treatment group, and 57 of 60 for the comparison group).

Analysis and Results

Program Energy Savings

The goal for this part of the analysis was to find energy savings resulting from participation in the program. Weather-normalized savings for each account were summed to get total per-building savings figures. Confidence levels were also calculated, based on the model's ability to weather-normalize consumption. By subtracting the median change in energy consumption in the comparison buildings from the median change in energy consumption in treated buildings, we were able to determine net (control-adjusted) savings in therms. The mean and median of per-dwelling-unit energy savings were also calculated and presented.

Participating buildings achieved median control-adjusted energy savings of 374 therms per dwelling unit (with a 95% confidence interval of 238 to 511 therms) after treatment. This represents a savings of 28.4% (19.9 to 36.9) of the pre-treatment energy use. Without adjusting for the comparison group's increased energy use between pre- and post-treatment periods, the median savings were 277 therms (188 to 420) per dwelling unit, or 21.9% (16.9 to 39.1) of pre-treatment energy use. Figure 1 illustrates these findings.

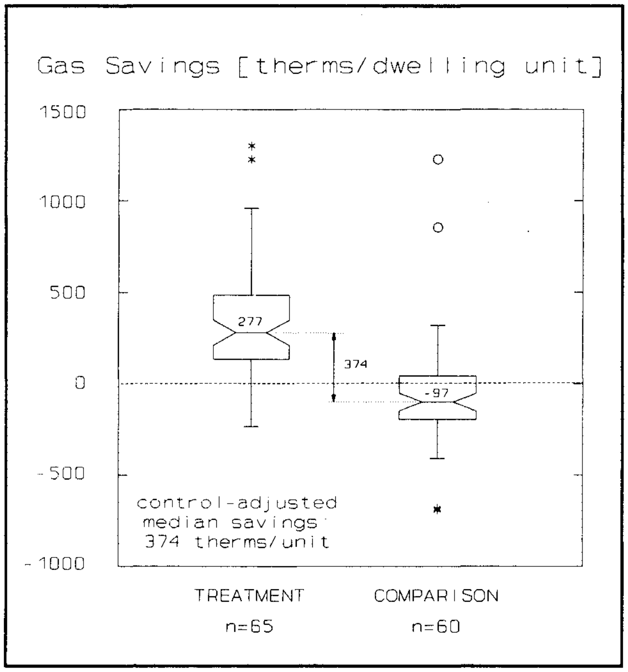

Figure 1. Per-unit Gas Savings for Treatment and Control Groups

Program Cost-effectiveness

Another goal in this evaluation was to find out how cost effective the program and the ECMs were based on measured energy savings rather than using engineering estimates. Three benefit-cost ratios were derived using the following tests:

- BCR using the ICC's total resource cost test (Case C'). Benefits: amount of energy saved by all buildings used in the sample times the marginal cost per therm of gas. Costs: energy loan amounts for all buildings in the sample plus administrative costs to run the program for every dwelling unit in the sample.
- BCR for the package of ECMs installed in a building. Benefits: each building's total energy savings times the retail cost of gas. Costs: amount spent on that building's ECM installations.
- BCR from the participant's perspective. Benefits: each building's total energy savings times the retail cost of gas, plus the interest rate subsidy provided by the program. Costs: energy loan amount plus other owner costs.

Using three cost-effectiveness tests, benefit-cost ratios were calculated. The program had a BCR of 0.77 using the Case C' test, the test used and prescribed by the ICC. The program was not cost effective using the Case C' test, but did pass cost-effectiveness tests from other perspectives. The Participant's Perspective BCR was 1.80, and the ECM Package BCR was 1.77. Even with 28% savings, the program was not found to be cost effective because of high expenditures.

ECM Installation Frequencies and Expenditures

A total of $1,751,959 was spent on ECMs installed in the 65 building sample, with a median per-unit expenditure of $1,467 (an extremely wide range of per-unit expenditures existed, with a low of $222 and a high of $3,334). Total loan amounts were $1,826,988. Of the total ECM expenditures, 47% was spent on storm and replacement windows, 23% was spent to replace heating systems, and the remaining 30% was almost evenly split between heating system retrofits, domestic hot water measures, insulation, infiltration reduction, and repair/rehab/lighting. The most commonly installed measures were storm windows, followed by indoor thermostats, boiler replacements, radiator work, and ceiling cavity insulation, in that order. About 9% of all expenditures were on several ECMs not identifiable as gas conservation measures, such as repairs, rehabilitation, and (electric) lighting measures.

Savings Attributed to ECMs

In order to better understand the sources of the energy savings in the treatment group, we undertook a regression analysis of energy savings as a function of the ECMs that were installed in the building. The goal of the analysis was to establish which ECMs contributed to energy savings and were the most cost effective.

The process of building such a model is not a straightforward task. We wanted to explain as much of the variation in energy savings as possible, but the risk of putting too many predictors in the equation was that we would reduce the ability to make sense of the model and may come to the wrong conclusions about some aspect of it. Since it would be difficult to make any predictions of energy savings from the installation of the 35 individual ECMs used in the program, we grouped them into eight categories:

- Heating system replacement.
- Heating system retrofits.
- Domestic hot water measures.

- Insulation.
- Storm windows.
- Replacement windows.
- Infiltration reduction.
- Repair/rehab/lighting.

We hypothesized that the spending levels for ECMs might be a better predictor of energy savings than simply using dummy variables for the presence or absence of an ECM. Such a model would yield estimates of the energy savings achieved per dollar spent on the ECM. We also theorized that a building that used more energy per dwelling unit had a higher potential for energy savings and would achieve more savings. We therefore included a variable for the pre-retrofit NAC per dwelling unit.

Our initial formulation of an energy savings model predicted per-unit energy savings as a function of per-unit, pre-NAC, and per-unit installation costs in each of the eight ECM categories. We ran the model (and variations of it) on the data for the group of 60 treatment buildings and 60 comparison buildings. These trial runs uncovered three things:

- Spending in two ECM categories, infiltration reduction and repair/rehab/lighting had no correlation to gas savings. We therefore dropped these categories from the model.
- There was significant negative interaction between heating system replacement and insulation. In other words, energy savings from a building that received both measures was less than the sum of their individual energy savings. This makes sense, since insulation reduces the heat load on the building, and thus reduces the savings from a new heating system. To account for this effect, we added an interaction term to the model. (We found no significant interaction effects among the other variables.)
- One comparison building with extremely high energy savings (48%, or more than 1200 therms per unit) significantly reduced the fit of the model. We dropped this building from the equation as an unexplained outlier, but retained other outliers in both groups.

After incorporating these changes, the model was run again, and the results examined for evidence of confounding variables, collinearity, and interaction (aside from the one exception noted below, we found none).

The adjusted coefficient of determination (r^2) was 0.59, indicating that about 60% of the variation in energy savings is accounted for by the model and about 40% remains unexplained, typical for this type of analysis. All but two of the variables were significant at a probability level of less than 0.01. The lack of confidence in the coefficients for heating system retrofits and domestic hot water retrofits may be due to collinearity between these variables. It is important to remember that the coefficients were not particularly well determined, but are useful to reveal broad differences between ECM categories.

The regression analysis of energy savings as a function of spending in six ECM categories showed that

Table 1. Energy Savings and Cost-effectiveness for ECM Categories, using the Results of the Regression Analysis (with 95% Confidence Limits)

ECM	Number of Installations	Estimated Energy Savings		Estimated Average ECM life (years)	Estimated Package BCR	
		Unit	Therms]		Unit	Therm
1 Heating System Replacement	29	373	169	19.5	1.58	0.50
2 Heating System Retrofit	45	81	106	7.6	1.30	1.44
3 Domestic Hot Water	28	64	111	7.1	1.25	1.83
4 Insulation	33	207	150	22.5	4.43	2.45
5 Storm Windows	40	81	85	10.0	0.37	0.28
6 Replacement Windows	15	174	171	23.8	0.72	0.43
1-4 Combined	18	458	134	21.0	1.55	0.61

insulation and heating system replacement provided statistically significant energy savings and were cost effective from the ECM package perspective. Storm windows, on the other hand, were found not to be cost-effective (see Table 1). Savings generally increased as pre-treatment normalized annual consumption of gas (pre-NAC) and per-unit expenditures increased, indicating greater savings where greater opportunity existed and more money was spent.

Cooking Gas Consumption

The cooking analysis was based on accounts with appliance codes indicating cooking as the only gas end-use. We used cooking accounts from all available treatment and comparison buildings—not just the samples used for the heating analysis. We did not screen out account changes during the analysis period; while this increased the variability of the sample, it also allowed us to retain a large sample size, since account changes during the three- to four-year period of interest were common. The analysis of cooking consumption centered on two parameters: (1) annual gas consumption and (2) a comparison of daily consumption during the heating and non-heating seasons.

We first annualized gas consumption for each account (without weather correction). We then divided the pre- and post-treatment analysis periods into a heating season (October through April) and a non-heating season (May through September), and looked at daily gas consumption in each period. We were particularly interested in the ratio of heating season to non-heating season consumption. A high ratio would indicate much higher gas consumption during the heating season and would suggest the use of a range for space heating. Biederman and Katrakis found that heating season gas consumption was 50% to 500% higher than non-heating season gas use (ratio = 1.5 to 5.0) in buildings where tenants said they frequently used their gas ranges for space heating (Ref. 2).

Given that it is difficult to unambiguously identify space heating in cooking accounts, we were unable to ascertain a clear answer for the extent to which this practice occurs. But we found no significant differences in consumption patterns between the treatment and comparison groups. The median cooking consumption was about 75 therms in both groups and showed more than a 4-therm reduction between the two analysis periods. Except for the slight overall decline in cooking consumption in both groups, we found no systematic changes in cooking gas use between the pre- and post-weatherization periods. It is important to recognize, though, that the analysis was only possible with individually metered accounts. The use of ranges for space heating is arguably less likely among these customers—who directly pay the cost of range gas use—than among residents in master-metered buildings where the landlord pays the direct cost of range gas consumption.

Discussion

Meter Configurations and Aggregation by Building

Most of the treatment and comparison buildings had master-metered heating accounts. Some of these master accounts also included the entire building's hot water and cooking gas. Other master-metered buildings had individual accounts for cooking gas in addition to master-metered heating and hot water. All buildings in the comparison group had individually-metered cooking accounts, since the number of cooking accounts present at an address was the only way to know how many dwelling units were present in those buildings. The number of dwelling units for the treatment group was already known from program records, so not all buildings had meter configurations like the comparison group. Few buildings in either group had individually metered heat as part of their meter configurations.

For the heating analysis, all accounts marked as central or individual-unit heating accounts were assembled. Some of those accounts included other gas end uses. After weather normalization was completed, all accounts were assembled and grouped by service pipe number (building). The buildings reaching this stage constituted our final sample size of 65 treatment and 60 comparison buildings.

For the cooking analysis, all accounts marked as cooking accounts were assembled. No accounts with other gas end-uses were included. The accounts reaching this stage comprised our final sample size of 574 treatment and 1,737 comparison accounts.

Issues in the Presentation and Reporting of Results

Normalization by number of dwelling units. The first step in understanding the results of the weather normalization and the analysis of ECM costs was to "normalize" with respect to the number of dwelling units in each building. There was a wide range of building size in the evaluation—from 5 to 163 units—and energy savings results had to take into account the amount of space they applied to. Square footage data would have been a more precise measure of area for which energy savings results referred to, but building square footage data was not available from the data sources we used for this evaluation. Nearly every result (except for the cooking analysis) of the evaluation was first divided by the number of units it represented before being reported.

Measures of centrality. Whenever practical, both means and medians were used to report central tendencies. This approach was intended to aid users of the evaluation to gain a more in-depth understanding of the evaluation's results.

While both means and medians are used to represent central tendencies, each has its unique value. The mean is a reliable measure of centrality where there is a symmetrical distribution. It is also used to estimate total program benefits. On the other hand, the median helps explain what a typical building has experienced and can include extreme values without throwing them out. If the mean and median values are extremely different, the user of the evaluation is alerted to the fact that the distribution of the data is skewed.

Quantification of uncertainty. To quantify the uncertainty in total consumption and savings for the program, we calculated the root-mean-square of the 95% confidence intervals for NAC resulting from the weather normalization model for each building. This procedure incorporated the uncertainty in the ability of the model to predict building gas consumption under standard conditions—an important step, since we allowed buildings with minimal consumption data to remain in the analysis. Note, however, that this approach is fundamentally different from one that treats the evaluated buildings as a sample of the larger program. The latter approach would typically estimate total program benefits as the mean (or median) of the sample savings times the number of participants, and the uncertainty would be based on the sampling error of the mean (which implicitly includes the uncertainty from the weather normalization procedure).

Our uncertainty estimates for total program savings, on the other hand, do not contain this sampling error, and therefore measure only the uncertainty in the measured savings for our set of 65 buildings. The single exception to this is that we incorporated the sampling error for the median percentage savings in the comparison group which we used to control-adjust the total impacts. The rationale behind this is that the comparison groups savings were, in fact, a sample of the general rental population.

Along with confidence intervals, another tool we used to help evaluation users understand the value of the reported results was to present confidence intervals and coefficients of determination where applicable. When there is a wide confidence interval or a low r^2 value, the user is alerted that the reported result is far from being certain.

Control-adjustment of total sample savings. Total savings were needed in order to calculate the aggregate cost effectiveness of the 65 buildings. We calculated the gross savings for the buildings by summing the savings for each building, then adjusted this figure to account for the (negative) savings seen in the comparison group. The control-adjustment was calculated as the median percentage savings for the comparison group times the total pre-treatment consumption in the treatment group. This procedure implicitly assumes that, in the absence of the program, every building in the treatment group would have experienced an increase in consumption equal to the median percentage seen in the comparison group.

Building Subgroups

In this evaluation, the sample of 65 buildings contained five buildings that did not have matching comparison buildings. The five were significantly larger than the remaining 60, with between 77 and 163 dwelling units. Energy savings, spending, and BCRs were quite different between these and the other 60 buildings. We were interested in examining the differences between the two groups.

The five very large buildings showed significantly lower per-unit energy use, ECM expenditures, and energy savings, although the percent (of gross) savings was almost identical (22.2% for the smaller buildings versus 21.7% for the very large buildings). The control-adjusted Case C' BCR was very low for this group of buildings (0.41): two of these buildings had increases, rather than reductions, in energy use after treatment. The Case C' BCR was 0.90 for the other 60 buildings.

Conclusions

Assessing multifamily evaluation issues introduces several variables not present in single-family evaluations. By making sure comparison buildings were as much like the treatment buildings as possible in size and location, a control group was a useful tool in determining true energy savings in the treatment group. After aggregating each account's heating results back to the building level, it was very important to normalize with respect to the number of units in each building. Measures of centrality should be approached with an eye for clarity, recognizing that presenting both means and medians is often a good idea. The uncertainties inherent in statistical analysis (confidence intervals and r^2) should be reported to help the evaluation user recognize the validity of the results. Control-adjusting the results gives a truer picture of energy savings by removing influences other than the program on changes in energy use.

Endnotes

[1] The Case C' benefit cost test is similar to the Total Resource Cost test as defined by the California Energy Commission. Benefits, including energy savings by all buildings in the sample times the marginal cost of gas, are discounted at 9% per year (at the time of this study). Costs include the total loans given, other costs to the building owners, and administrative costs. The following equation provides more detail.

$$BCR_{c'} = \frac{\sum_{e=1}^{n} \left\{ \frac{1+d}{d-f} \left[-1 - \left(\frac{1+f}{1+d} \right)^{t-1} \right] * FYS_e * MC \right\} * (1-FR)}{[C_1 * (1-FR)] + C_u}$$

Where:

e = an Energy Conservation Measure
f = marginal fuel escalation = 0.097 /year
d = social discount rate = 0.09 /year
t = average lifetime of ECMs = 18 years (when averaged)
FYS_e = first year savings for an ECM (therms)
MC = marginal cost of gas = \$0.246
FR = percent of free riders = 0.15
C_1 = total loans given + total other owner costs
C_u = total allocated administrative costs = \$323/unit

[2] A previous evaluation of this program was performed. See A. Evans and John Katrakis, "Chicago's Residential Energy Conservation Loan Fund: A Preliminary Evaluation of Its Impact on Multifamily Buildings," *Proceedings of ACEEE 1988 Summer Study on Energy Efficiency in Buildings*, Vol. 2, pp. 38-50, 1988.

[3] We used our own interpretation of these algorithms, rather than the PRISM software. See Ref. 1.

References

(1) Fels, M. "PRISM: An Introduction," *Energy and Buildings*, Vol. 9, Nos. 1&2, February/May 1986.

(2) Biederman, N., and J. Katrakis. "Space Heating Improvements in Multifamily Buildings," final report, Gas Research Institute, December 1989.