What's the Point (of Sale)? Program Activity Impacts Efficient Bulb Sales—Proof Across 44 States and Five Years

Chris Russell, NMR Group Inc., Bryan, TX
Mike Strom, NMR Group Inc., Boston, MA
Scott Dimetrosky, Apex Analytics, Boulder, CO
Noah Leib, Apex Analytics, Boulder, CO

ABSTRACT

This evaluation examined the impacts of lighting program activity and demographic factors on efficient bulb sales across 44 continental states during the first years of EISA implementation and those immediately preceding it. Given federal regulations that ban the manufacture and import of inefficient bulbs, many have questioned the need for residential lighting programs. This paper addresses the impact lighting programs have had during EISA implementation and provides evidence for the benefits of continued lighting program intervention.

Lighting programs show great variation across the US, both in how incentives are applied and utilized and in the types of bulbs supported. The present analysis examines the influence of whether a state had a lighting program, as well as the impact of more contextual factors like program budget, age of program, and number of program-incented bulbs. This analysis is based on a point-of-sale dataset that included bulb purchases—lighting programs increased efficient bulb purchases even after EISA was implemented.

Efficient bulb sales were positively impacted by the presence of a lighting program at the state level. Further, the larger lighting program budget, the more bulbs incented by the program, the greater the number of efficient bulbs sold in that state. The evaluation argues that lighting programs are still successful in promoting efficient bulb sales over and above what would occur if there were no programs.

Research Goal

The goal of this research is to understand the influence of various predictors on the sales of efficient bulb types across the nation while federal legislation is actively phasing out less efficient bulb types—especially the impact of program activity. This research is meant to determine whether lighting program activity is still a driver of efficient bulb sales and saturation, thereby providing insight into whether states should maintain their residential lighting programs.

Introduction

The purpose of the point-of-sale (POS) modeling research was to understand the influence of various predictors on the sales of efficient bulb types across the nation, namely the impact of program activity. Given the rapid and widespread changes to the lighting market in the United States, including the increased lighting efficiency standards stemming from the Energy Independence and Security Act of 2007 (EISA) that should phase-out general service incandescent bulbs, many have begun questioning whether residential lighting programs continue to impact efficient bulb sales.[1] The introduction of general service light-emitting diode (LED) and halogen bulbs also intended to replace incandescents only complicates the matter further. To assess the continued impact of lighting programs on efficient

[1] See: Laura Morefield, *EISA and Future Residential Lighting Programs*, at www.energystar.gov for a more detailed treatment of the anticipated threat of EISA to lighting programs.

bulb sales, this research determines whether states that have lighting programs tend to sell a higher percentage of efficient bulbs than states that do not have lighting programs, while controlling for other potentially biasing factors.

Preliminary Analysis

However, the results of that preliminary exploratory analysis revealed several unusual and counterintuitive findings. For example, as shown in Figure 1, when examining non-specialty general service CFL pricing trends, states with active lighting programs (like MA, CA, and CT) did not tend to show lower CFL prices than states with no programs.[2]

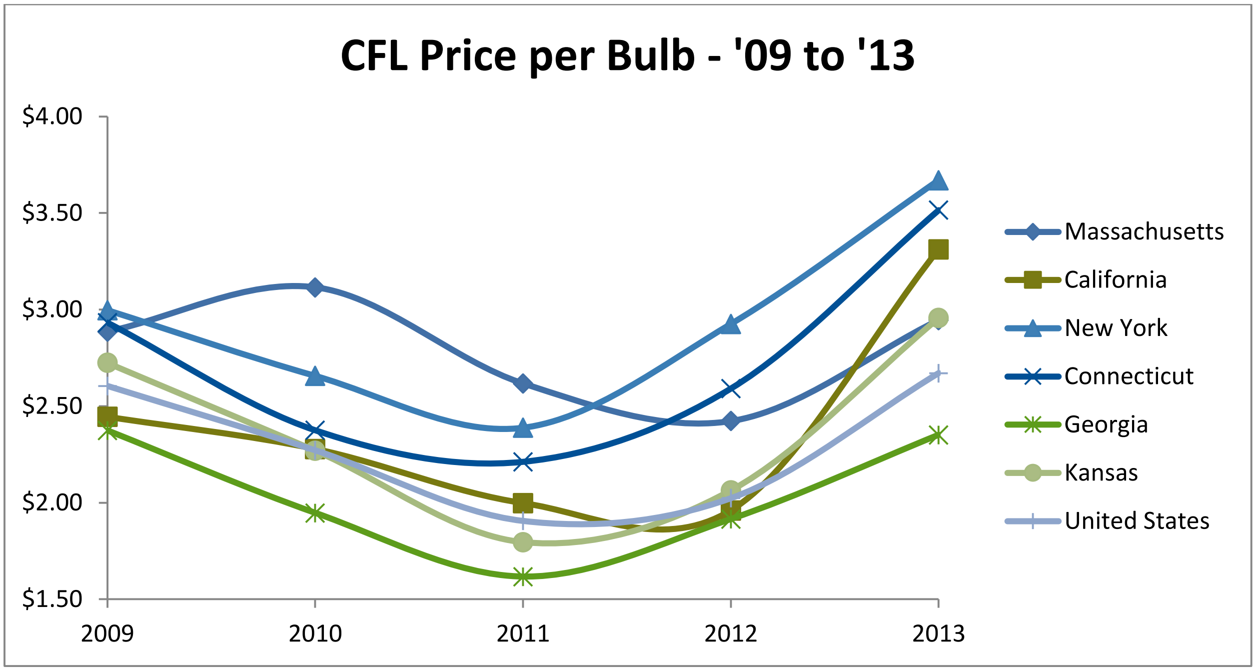

Figure 1. CFL Pricing Trends

[2]In Figure 1 and Figure 2 Kansas is a non-program state and Georgia is a recent lighting program state without a LED component.

The Team found a similarly unexpected set of results when examining LED pricing trends. As shown in Figure 2, all states considered in the initial exploratory descriptive analyses had LED prices that converged over time, despite program states specifically focusing on bringing the shelf price of LEDs (and CFLs) down to be competitive with that of incandescents.

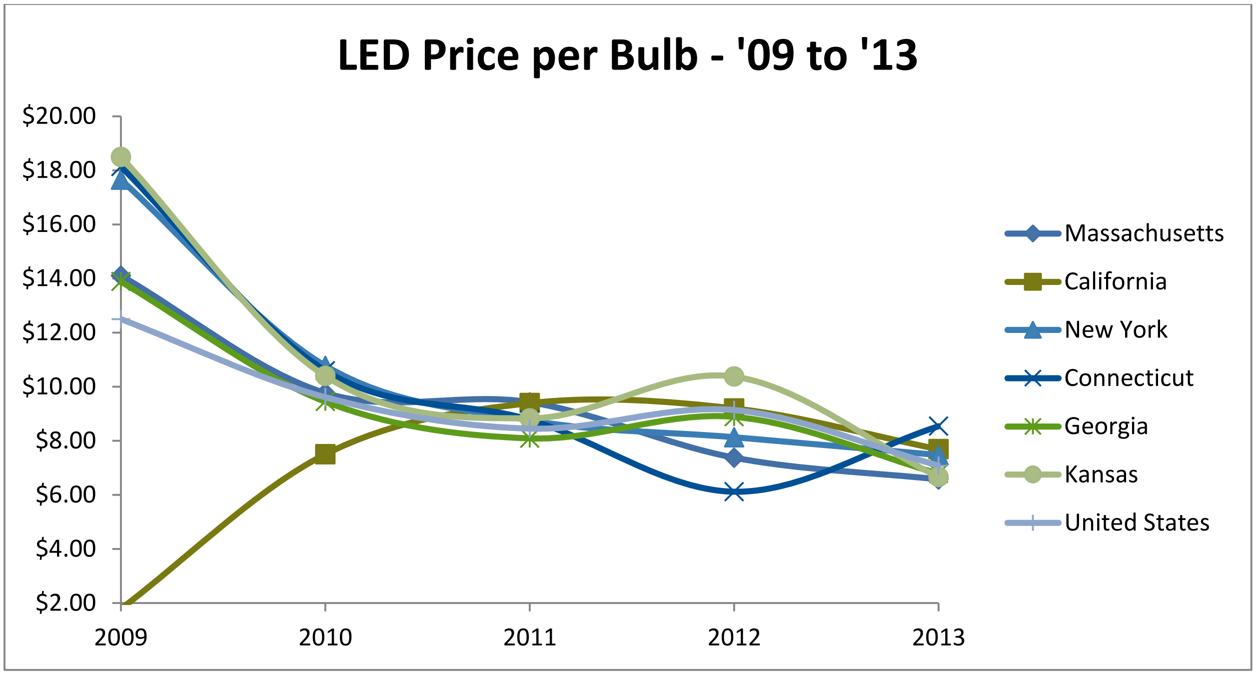

Figure 2. LED Pricing Trends

Program states did not show lower pricing trends for efficient bulbs than did non-program states. The team hypothesized that the simplistic graphs and exploratory analyses conducted on the POS data may hide the impact of intervening factors, which a larger model would be able to take into account. For these reasons, the Team turned to a more nuanced approach to explore the true influence of program activity on bulb sales. Utilizing a more sensitive and expansive modeling approach, the Team could control for external factors with the potential to bias the exploratory.

Data

The POS data were purchased through LightTracker, an initiative of the Consortium for Retail Energy Efficiency Data (CREED). The Universal Dataset (UDS) Report was determined to be the most appropriate option for the current effort because it includes data for all available territories, bulb types, and years of reporting. More specifically, this data set includes lighting sales data for grocery, drug, dollar, club, and mass market distribution channels. The POS data do not include the home improvement and hardware retail channels, which can account for a substantial amount of statewide bulb sales. Unfortunately, the POS data does not represent all light bulb sales and has been shown to represent a quarter of the sales in test states (MA and NY) and a comparable amount of program bulbs.[3] To the extent that these channels are representative of the market in each of the states then there is no bias in the results of the analyses.

Certain threats to the validity need to be addressed given the partial representation of the data to the full lighting market. Key aspects of the UDS Report include:

[3] The Team used MA and NY program bulb tracking data and matched them to the POS data to determine the saturation of program and market level sales.

- Sales volume and pricing from 2009 to 2013 for CFLs, LEDs, halogens, and incandescent bulbs for all channels combined
- Data reporting by state (with 44 states included) and bulb type
- Inclusion of all bulb styles and controls

The POS data was used to create the dependent variable for all models, defined as the percentage of all bulb sales in a particular state that were energy efficient. All model inputs described below attempted to predict the percentage of statewide efficient bulb sales defined as (CFL+LED Sales)/All Bulb Sales.

Other data feeding into the model include:

- Program activity was gathered in three ways. The first piece of information the team gathered was simply whether a particular state had a lighting program in place for any year between 2009 and 2013. If there was evidence of a program (either through internal evaluation work or a literature review), the next step was to search for more detailed program information—in particular, the lighting program budget. The Team researched and aggregated all this program activity information into two main program activity predictors to utilize as the main model inputs of interest. These included a simple yes/no predictor of whether a particular state had a program in a given year and a continuous predictor variable represented by the more detailed information the Team gathered on program budgets.
- Secondary internet research was conducted to determine the number of store locations in each state and the total square footage of those stores per state for both the retail channels reporting to the data collection vendor and those channels not reporting. The purpose of this background research, and the eventual model inputs resulting from it, was not only to assess the influence of stores' presence or absence on bulb sales, but also to utilize these data as control variables such that any significant impacts of other model inputs would not be a result of a particular state simply having greater or fewer stores in the channels whose bulb sales were reported.
- State-level demographics were gathered from the American Community Survey (ACS, www.factfinder2census.gov), including annual state-level data for the population, total number of households, household tenure, count of homes built before 1980, categorical education, median income, and average number of rooms in the home.[4] These variables were used as model inputs and, in the case of statewide populations and households, were utilized to create the per-capita and per-household square footage rates of the retailer channels that do or do not report to the data collection vendor. The rationale for gathering the demographic data was to control for non-program factors that could possible impact the number of efficient bulbs sold. The demographic variables were:
    - Median number of rooms per home-included as a proxy for house size
    - Percent of homes built before 1980-we anticipated that the age of the home would have a negative relationship with bulb purchases,
    - Percent of renters paying utilities-a greater number of renters paying utilities would have a positive relationship with efficient bulb purchases.

Model Design and Inputs

The Team analyzed the POS data to determine the influence of state-level program activity, presence/absence of reporting retail channels, and state-level demographics on percentage of efficient bulb sales (also at the state level). A series of robust random-effects regression models were fit after the

[4] The Team utilized single year household ACS data from 2009, 2010, 2011, and 2012.

team determined the random-effects models to be appropriate after Hausman tests comparing the results from these models to analogous fixed-effects models failed to reject the hypothesis that the estimate obtained from the random-effects model were inconsistent.[5]

The series of robust random-effects regression models in one of the two following forms, with the difference between the two bolded:

$$\log(\%efficient.sales_{i,j}) = \alpha + \beta_{0,i} + \beta_1 \log(cr.sqft_{i,j}) + \beta_2 \log(noncreed.sqft_{i,j}) + \beta_3 avg.electric.price_{i,j} + \beta_4 cost.of.living_i + \sum_{k=1}^{p} \gamma_k dem.var_{i,j,k} + \boldsymbol{\theta prog1_{i,j}} + \tau_j + \epsilon_{i,j}$$

$$\log(\%efficient.sales_{i,j}) = \alpha + \beta_{0,i} + \beta_1 \log(cr.sqft_{i,j}) + \beta_2 \log(noncreed.sqft_{i,j}) + \beta_3 avg.electric.price_{i,j} + \beta_4 cost.of.living_i + \sum_{k=1}^{p} \gamma_k dem.var_{i,j,k} + \boldsymbol{\theta prog2_{i,j}} + \tau_j + \epsilon_{i,j}$$

Where for a given state *i* and year *j (in both models)*:

$\%efficient.sales_{i,j}$ = Proportion of total CREED-reported bulb sales that were efficient bulbs. Calculated as $(\#CFL_{i,j} + \#LED_{i,j})/(total\ bulb\ sales_{i,j})$.
$cr.sqft_{i,j}$ = Number of square feet of major CREED-reporting retailer channels.
$noncreed.sqft_{i,j}$ = Number of square feet of major non-CREED-reporting retailer channels.
avg.electric.price = Average statewide cost of electricity.
cost.of.living = Average statewide cost of living index.
$dem.var_{i,j,k}$ = One of *p* demographic variables for state *i*, at time *j*, with $k \in (1, \ldots, p)$.[6]
τ_j = Average proportion of efficient bulb sales across states that had no program activity for the entire study period, 2009 to 2013.[7]
α = Overall model intercept term
$\beta_{0,i}$ = Subject-specific deviation from overall-level intercept, α, as estimated by random-effects specification
β_1, β_2, γ_k, θ = Regression coefficients to be estimated by the model
$\epsilon_{i,j}$ = Error term

[5] The Hausman test is a specification test for panel regression that tests the assumption that each state's time-invariant characteristics are unique to the state and not correlated with other variables. When the Hausman test determines that the state's error terms are correlated it returns a non-significant p-value as proof that a random effects model is better suited to the data than is fixed effects regression.

[6] The Team present results for two models in the following section, with *p* ranging from 4 to 6 across those models. The following state-level demographic variables were considered: number of households, % of homes built before 1980, and % of renters who pay their own utilities, median income, % owner-occupied households, education level, and population. The Team determined which demographic variables to include in each model by selecting the covariate pattern yielding the highest adjusted R^2.

[7] Including this term allows the model to account for naturally occurring, non-program influenced, "baseline" trends in efficient bulb sales during the study period, which in turn helps to isolate the effect of program activity on efficient bulb sales as opposed to other outside factors.

The difference between the two model specifications involves the program variable. They each capture the level of program detail the Team was able to obtain for each available state, but they differ in the following way:

$Prog1_{i,j}$ = In the first, $prog_{i,j}$ was a simple indicator (yes/no) for whether there was any program activity at all in state i in year j. In this way, the Team was able to run a model that included all 44 available states.

$Prog2_{i,j}$ was a variable indicating the lighting program budget, as gathered through published reports, internet searches, internal evaluations, or provided directly by utilities.[8] It should be noted that the program budget variable includes program activity even if it was in retail channels not represented by the POS data.[9] The square root of program-related budgets was used in the models in order to adjust for the skewedness in the distribution of that variable.

The dependent variable in the model is specified as the proportion of all bulbs sold in each state that were energy-efficient bulbs. Utilizing proportion instead of count recognizes that the state population would heavily influence absolute counts of sale, as would prior saturation levels of efficient bulbs (i.e., longer lifetimes for efficient bulbs lead to fewer bulbs of all types being sold). This allowed the model to control for factors besides the number of households, such as average hours of use, average home size (sq. ft.), and any climate-related or other regional differences across states that could also have an impact on the absolute number of state-level bulb sales.

The random-effects models account for any potential correlation among the multiple observation in each state through the random intercept terms $\beta_{0,i}$. Additionally, the models estimate the program effect using both within-state and between-state variability, using more information. Fixed-effect analyses are often used in the energy efficiency evaluation field in billing analysis given the fixed effects of a household and large variation across household. This modeling research is using store sales data and does not have the same issues concerning variation across stores and states. The random-effects model is able to use information from all available states. The main parameter of interest in all three models is θ, which quantifies the impact of state-level program activity as defined by each of the specifications of $prog_{i,j}$ discussed above. The dependent variable and several control variables in the model were transformed using natural logarithm to adjust for the skewed distributions of those variables on their original scales and reduce the impact of outlying observations.

After fitting these models with the efficient bulb proportion of sales, two additional set of models were estimated using only the proportion of CFL bulb sales and those using only the proportion of LED bulb sales as the dependent variable. In this way we can compare the impact of program activity on each of those technologies separately and determine whether either bulb type was benefiting more than the other from state-level program activity.

Model Results

The models utilize the continuous program budget variable collected by the Team with the year or "trend" variable included in the model. In other words, the Team controlled for underlying temporal trends that impact program and non-program states by including a fixed-effect term for each year in the data, represented by τ in the model above. Along with helping to control for other unobserved time-

[8] Program budget and number of bulbs supported were the only two detailed variables consistently available for all states in the POS dataset. The evaluators ran preliminary models using both of these possible indicators, and the program budget models performed better. Therefore, the evaluators and clients decided to rely on the program budget models.

[9] The degree to which this biases the results depends mostly on the percentage of budget that goes through home improvement versus other types of stores. Unfortunately, the evaluators did not have access to this level of detail, but we do know that reliance on home improvement stores varies by program administrator.

varying effects at the state level, these terms also allowed the Team to investigate how efficient bulb sales have been trending over time (for example, to see when CFL and LED sales reach their peak).

Such findings are highly informative, as they allow us to see how the market is trending. However, it is also true that, on average, program budgets increase over time. Indeed, the time and program budget variables in the model show a slight positive correlation. Although the correlation is relatively weak (r=0.19), it is still possible that the time variable in the model absorbs some of the program effects.

All Efficient Bulb Types

Table 1 below shows the results of the model using the proportion of total efficient bulb sales as the dependent variable. The table summarizes the estimated regression coefficients across the model. Note that interpreting the estimated coefficients related to program activity on an actual scale requires exponentiating the coefficients presented in Table1. Increased program activity, as measured by the program budget-based $prog2_{i,j}$ variable, is positively and significantly associated with increases in efficient bulb sales at the 90% confidence level. The results demonstrate that increases in a state's lighting program budget are associated with increases in efficient bulb sales. Specifically, for every $1,000 increase in the square root of a state's lighting program budget, there is an expected increase of 5.5% in the proportion of efficient bulb sales. To simplify interpretation of the model, this relationship can also be quantified as program expenditure elasticity. Consider a $1M increase in program budget: Based on this model, such an increase would lead to a 0.36% increase in efficient bulb sales in MA, yielding an elasticity of 0.2.

Table 1. State-Level Model Results for All Efficient Bulbs

Variable	Level	Model Results†
Intercept	--	1.6392 (1.3456)
log(cr.sqft)	Continuous	0.1794* (0.0747)
log(noncreed.sqft)	Continuous	-0.1378* (0.0545)
% built pre-1980	Continuous	-1.0199† (0.5972)
% renters paying utilities	Continuous	-2.3382 (1.5922)
Median # rooms per home	Continuous	--
Electric Price	Continuous	0.0149** (0.0054)
Cost of Living Index	Continuous	-0.0083** (0.0028)
Program Budget	Continuous	5e-05*** (2e-05)
log(non-program eff. sales trend)	Continuous	0.6934*** (0.0925)

Additional Details		
Number of States		27
Number of Observations		94
R^2		0.663

†Coefficient estimates presented with standard errors in parentheses beneath them.
log(covariate) indicates the natural logarithm of that covariate.
Note: † $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

CFL-only Results

Next, the Team considered the dependent variable of CFL sales (the proportion of all statewide sales that were CFLs). Across all years and all states in the POS data, however, the average proportion of total efficient bulb sales was 17.03%, and the average proportion of CFL bulb sales was 16.91%. As such, across all states, CFLs accounted for over 99% of all efficient bulb sales between 2009 and 2013. Unsurprisingly, then, the CFL-only model behaved nearly identically to the overall CFL + LED model.

Table 2. State-Level Model Results for CFL Bulbs

Variable	Level	Model Results
Intercept	--	1.6686 (1.3443)
log(cr.sqft)	Continuous	0.1764* (0.0750)
log(noncreed.sqft)	Continuous	-0.1365* (0.0549)
% built pre-1980	Continuous	-0.9818 (0.5996)
% renters paying utilities	Continuous	-2.3068 (1.6035)
Median # rooms per home	Continuous	--
Electric Price	Continuous	0.0147** (0.0054)
Cost of Living Index	Continuous	-0.0083** (0.0028)
Program Budget	Continuous	5e-05*** (2e-05)
log(non-program CFL sales trend)	Continuous	0.7166*** (0.0877)

Additional Details		
Number of States		27
Number of Observations		94
R^2		0.666

Coefficient estimates presented with standard errors in parentheses beneath them.
log(covariate) indicates the natural logarithm of that covariate.
Note: † p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

LED-only Results

In contrast to CFLs, the proportion of LED sales nationwide remains small, thereby contributing very little to the combined CFL + LED analysis. Considering only the proportion of LED sales—as opposed to combined CFL and LED sales or CFLs sales only, then, bring about different results than were observed in earlier models. In particular, we see a stronger influence of the continuous program budget variable (prog2*i,j) on* the proportion of LED sales compared to CFL sales or all efficient bulb sales which may be compounded by the channels available in the POS data. Additionally, the final fitted models utilizing only the proportion of LED sales as the dependent measure are more parsimonious than the previously discussed models (that is, requiring fewer inputs to achieve equivalent or greater explained variation). Of note, the LEDs-only model is able to explain almost all of the variation in LED sales as a proportion of all bulbs purchases having an R^2 of over 0.9.

Perhaps the most striking finding across the LED-specific model, however, is the steadily increasing trend over time in the proportion of LED sales. The LED-specific model shows that the proportion of LED sales was over 22 times higher in 2013 than it was in 2009. This contrasts with the trend for CFLs, for which the proportion of total bulb sales peaked between 2011 and 2012, and by 2013 had a proportion of sales no different than that in 2009 showing that CFL sales are no longer increasing in proportion of total bulbs sold while the proportion of total bulbs sold accounted for by LED bulbs continues to increase.

Table 3. State-Level Model Results, LEDs Only

Variable	Level	Model Results
Intercept	--	-1.1557 (0.9101)
log(cr.sqft)	Continuous	0.3886* (0.1677)
log(noncreed.sqft)	Continuous	0.2062 (0.1997)
Median income	Continuous	0.0196* (0.0090)
log(# of households)	Continuous	-0.5984** (0.2066)
% own	Continuous	--
% built pre-1980	Continuous	--
Electric Price	Continuous	-0.0047 (0.0104)
Cost of Living Index	Continuous	-0.0012 (0.0059)
Program Indicator	Yes	--
	No	--
Program Budget	Continuous	7e-05* (3e-05)
log(non-program LED sales trend)	Continuous	1.0607*** (0.0278)

Additional Details		
Number of States		27
Number of Observations		94
R^2		0.956

Coefficient estimates presented with standard errors in parentheses beneath them.
log(covariate) indicates the natural logarithm of that covariate.
Note: † p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Threats to Validity

The model results provided in this report are not without limitations. First and foremost is the issue of generalizability. As discussed, the sales data that serve as the dependent measure for all models do not represent full, market-level sales nationwide. Although many program and non-program bulbs sell through the retail channels included in the POS dataset, the absence of home improvement and hardware channels means that many bulb sales are not accounted for in the models. Based on the assessment of market-level bulb sales in MA calculated during the most recent onsite saturation study, the we estimate that the POS data represents roughly one-quarter of all sales.

This is not to discount the importance and quality of the data that are available—residential lighting program evaluators and implementers have been working for years to obtain actual bulb data captured at the point of sale for any retail channels, and the current POS data set represents the best of what is available. However, it should not be viewed as perfectly representative of the entire lighting market.

Conclusions

The results of the modeling efforts conducted in the present research suggest that lighting programs continue to have an influence on the lighting market, even in the years following EISA implementation. Across the three separate bulb-proportion dependent variables (all efficient bulbs, CFLs only, and LEDs only) the model demonstrated the positive and significant influence of program activity on the percentage of energy-efficient bulbs purchased statewide. Results suggest that as the lighting market continues to progress, programs focusing on LEDs are likely to have greater relative impacts.

The modeling also reveals that more simplistic approaches to understanding the lighting market, considering only factors such as bulb pricing trends or the number of efficient bulbs sold, often fall short of being able to explain or account for the many interceding dynamics in the market. The models presented here provide evidence that lighting programs matter, but the preliminary exploratory analyses hide the impact of intervening factors.

Yet, while the models and the results are important, this effort also represents a victory for program administrators and evaluators of upstream lighting programs. With some exceptions, administrators of upstream lighting programs have struggled to gain access to market-level bulb sales data—that is sales of both program and non-program bulbs of all types. As a result, previous evaluations have struggled to describe market share as the sole estimates available were limited to program-supported sales or customer or supplier self-reports, each of which suffers from measurement bias (e.g., recollection error or even intentional gaming of the estimates). While the LightTracker data are not perfect—critically, they lack any estimates from the home improvement and hardware channels that have historically served as the base for many residential lighting programs—the dataset contains market-level sales for the reporting retail channels comprising both program and non-program sales for CFLs and LEDs as well as halogen and incandescent bulbs. As such, they have provided evaluators with the ability to assess trends in market share and determine the continued impact of programs on efficient bulb sales across the nation.