All the Lights We Cannot See: Estimating the Distribution of Upstream Lighting Program Bulbs

Russell M. Meyer, NMR Group, Inc., Somerville, MA
Chris Russell, NMR Group, Inc., Somerville, MA
Jerrad Powell, NMR Group, Inc., Somerville, MA
Scott Dimetrosky, Apex Analytics, Boulder, CO
Tracy Dyke-Redmond, Eversource Energy, Boston, MA
Lynn Hoefgen, NMR Group, Inc., Somerville, MA

ABSTRACT

Light bulb replacement incentives have historically been a key tool of residential energy efficiency programs. This study estimates the effects of several predictors of efficient bulb uptake in residential households in Massachusetts. To do so, this analysis makes use of a series of in-home lighting surveys performed throughout the state in which inventories of household lighting products are tabulated over a multi-year panel. This is coupled with a database of retail locations throughout the state to estimate the relationship between household proximity to retail locations and the use or purchase of energy efficient lighting technologies. In extensive testing of this relationship, the analysis finds no evidence for a meaningful association between household-to-store proximity and bulb use within the state. There is a limit to this null finding, however, since it is the case that a difference in usage rates exists between states, suggesting there is an upper bound to the distance of lighting program leakage. Within MA, even demographic differences play only a small role in predicting different rates of efficient lighting technology uptake, consistent with broad success of upstream lighting programmatic activity.

Introduction

Where do upstream program bulbs go once they leave the shelves? Do consumers who live near program-participating retailers benefit disproportionately compared to other consumers in the state? How can we estimate the distribution of program bulbs around the state? The motivation for this study is to identify the factors or combinations of factors that can be used to predict consumers' purchases of upstream lighting products supported by the program administrators (PAs) in Massachusetts (MA). This work makes use of a panel dataset of in-home lighting equipment surveys to measure energy efficient lighting technology uptake. This survey dataset has been collected in MA to report on residential lighting technology and usage patterns (see, e.g., NMR 2016). This dataset includes a count of every bulb in the household for households included in the survey, with information about each bulb's technology type and records of new energy efficient (EE, defined as bulbs of Light Emitting Diode, LED, or Compact Fluorescent, CFL, technology-types) bulb installations. Repeat annual visits to survey households allow for the identification of newly purchased lighting equipment and follow-up questions for residents to learn about those purchases and the retail channels used. Because the dataset also includes demographic information about the occupants of the household, these demographics can be used as predictors of EE bulb uptake to estimate EE bulb distribution across the state using 2015 American Community Survey tract-level data (tracts are a geographic unit used by the ACS). These tract-level statewide estimates are also useful as inputs for ongoing work in MA top-down modeling of programmatic effects and in the effort to better understand the success of the upstream lighting program in reaching hard-to-reach subpopulations. This report presents analyses of both the year-over-year changes in lighting technology usage, representing new installation (purchase modeling), and the stock of existing lighting technology installed in sockets for active use (saturation modeling).

This research shows that the PAs' upstream lighting program has been successful in spreading efficient lighting nearly evenly throughout the state, with no discernable differences attributable to proximity to different types of stores, albeit with a few small but significant differences by demographics. The very success of the program, with widespread cumulative effects (NMR, 2017a), makes it difficult to detect predictors of efficient lighting saturation and sales. Accordingly, one possible explanation could be that the program has influenced all retailers in the state; since most major lighting retailers have participated in the program at some point, the shelf stocking practices of other formerly-participating and non-participating stores may encourage efficient bulb purchases due to the wide footprint of the program on the MA lighting market. Another possible explanation is that there is substantial leakage of PA-supported efficient lighting to municipal utility territories that are in close proximity to PA territories. Due to broad distribution of EE bulbs in MA, no meaningful relationship of household proximity to retail outlets is found in the data. Further away, however, such leakage does not appear to be taking place: households in New York (NY), where there has been no upstream lighting program over the last several years, have lower levels of efficient bulb saturation and purchases than those in MA.

Data

The work reported here makes use of a set of household surveys of lighting technologies and a database of retail locations in MA and NY. The in-home lighting survey data have been collected annually since 2013 from a panel of households (with newly recruited households added each year) in MA. These household lighting data collection activities have occurred in other states as well, and this study makes use of NY data (excluding NY City) as a comparison area for a portion of the results. These surveys involved site visits by trained technicians who cataloged the lighting technology used in the home, by bulb type, for each lighting socket in the home, as well as any bulbs held in storage. All bulbs were marked by the technician at the time of the site visit; any unmarked bulbs at the time of the next visit could then be assumed to be newly acquired. The most recent survey of households included 420 in MA and 150 in NY in early 2016.

Each year, several new households are added to the ongoing panel of households in the annual survey. This is, in part, to test for differences in lighting usage characteristics between households that have not previously been visited by survey technicians and the panel households who have been visited, and to ensure that the ongoing panel does not differ from the overall population as a result of the anticipation of being visited by a survey technician (i.e., the Hawthorne effect). In testing for differences between new survey participant and repeat-visit panel households no material differences are found.[1]

In addition to the cataloging of lighting equipment in use at the socket level in households, demographic information about the occupants of the households is collected as well. Of particular use in the present study are demographic characteristics that are also collected at the tract level by the American Community Survey (U.S. Census 2013) so that extrapolative predictions can be made to geographic units across the state; however, the thoroughness of data gathering during onsite visits also allows for a rich analysis of demographics that are not aligned with American Community Survey data collection which generates further insights about lighting technology uptake in the sample. Table 1 lists some of the demographic fields made available by the in-home survey data that are used in this work.

[1] For more detail on the survey collection methods see NMR 2016.

Table 1. Demographic Variable Description

Variable	Description
Single Family Home	Binary variable indicating housing construction type
Low Income Household	Binary variable indicating Low Income Home Energy Assistance Program (LIHEAP) low-income status
Mid Income Household	Binary variable indicating middle-income status defined by a household income less than $150k/year, and not classified as low-income by LIHEAP classification
Renter	Binary variable indicating tenancy status
College Grad	Binary variable indicating if at least one occupant of the household is a college graduate
Urbanization	Continuous variable indicating percentage of the household's zip code that is considered urban, according to the 2010 US Census
Number of Rooms	Count of the number of rooms in a household, top-code limited at nine

The household survey data are combined with a database containing the addresses and types of about 11,000 stores in MA and NY (not including New York City) that are expected to sell lighting products. These two data sets were used to estimate drive times for each of the household survey participants to multiple retail channels. Drive time, using road network GIS software, between household and retail outlet was estimated for each household to each store. The lowest drive time value, in minutes, for each household to a given store channel is available for inclusion in a model. This includes permutations for each retail channel of those stores that are upstream lighting program participating stores (for MA households), or those that are not program participants. Several forms of this set of variables have been included in the models, including nearest store of each channel, nearest store of each channel by program status, additional drive time to nearest program store of channel compared to non-program store of channel, and a restricted set of retail channels that includes the largest fraction of total lighting sales. Additionally, model permutations using aggregations of similar channel types were also tested (see NMR 2017b for a complete reporting of results associated with this modeling work, particularly Appendices A and B). Table 2 shows the number of stores included in the database, by channel type and state.

Table 2. Store count by channel

Store Channel	MA	NY	Total	Store Type Examples
Bargain	85	160	245	99 Cent Buy, Dollar Bazaar, Real Deals, Yankee Discount Sales
Bodega	756	751	1,507	Devita's Market, J & M Variety, Quezada Grocery
Club	37	70	107	BJ's Wholesale, Costco, Sam's Club
Corner Drug	87	243	330	Holbrook Pharmacy, Jones Drug Store
Discount	390	1,021	1,411	Big Lots, Christmas Tree Shops, Dollar Tree, Grossman's, Ocean State
Drugstore	735	1,167	1,902	CVS, Village Pharmacy, Walgreen's
Ethnic	256	375	631	Bravo Supermarket, H Mart, Nijiya Market, Super 88
Furnishings	58	87	145	Bed, Bath & Beyond; Ikea; West Elm
Grocery	255	539	794	Dover Quality Market, Harvest Coop
Hardware	374	690	1,064	ACE, Aubuchon Hardware, Tractor Supply, True Value
Home Center	75	165	240	Home Base, Home Depot, Lowe's
Lighting & Electronics	191	143	334	Batteries + Bulbs, Best Buy, Electrical Wholesalers Inc.
Mass Merchandise	124	199	323	Staples, Target, Wal-Mart
Office Supply	73	98	171	Office Depot, Office Max
Supermarket	567	1,156	1,723	Hannaford, Market Basket, Price Chopper, Shaw's, Stop & Shop
Thrift	45	51	96	Goodwill, Savers
TOTAL	4,108	6,915	11,023	

Saturation Modeling

The preferred model results from this portion of the modeling effort are shown in Table 3[2]. In the table, the coefficient estimates for each independent variable included in the model are included in the cell, with standard errors in parentheses below. The table shows that the relationship between renter status and saturation flips between the two technology types: renter households have lower LED bulb socket saturation but higher CFL bulb socket saturation (with the result, by the definition of the "Renter" indicator, reversed for owner-occupied households). Households with at least one college graduate have higher LED bulb saturation. Larger households, measured by the number of rooms in the home, have greater saturation of LED bulbs, but lower saturation of CFL bulbs. The finding that the estimates of the effects of the predictors changes sign for LED and CFL bulbs suggests that uptake rates for the two technologies are behaving quite differently and that estimating overall efficient bulb uptake without considering the differences of these two technology types could result in a misunderstanding of the relationships that influence EE bulb installations.

[2] Several model variations, in functional form and in the independent variables included, were tested in this exercise. For a complete set of results see the appendices in NMR, 2017b.

Table 3: Tobit model with EE, LED, or CFL socket saturation as the dependent variable; MA households only

	EE (LED+CFL)	LED	CFL
Renter	0.0175	**-0.0869***	**0.0647***
	(0.0339)	**(0.0317)**	**(0.0315)**
College Grad	0.0164	**0.0461***	-0.0008
	(0.0269)	**(0.0253)**	(0.0251)
Number of Rooms	**-0.0137***	**0.0151***	**-0.0159***
	(0.0075)	**(0.0069)**	**(0.0070)**
Constant	**0.5575***	-0.0886	**0.4710***
	(0.0627)	(0.0588)	**(0.0584)**
N	487	487	487
pseudo-R^2	0.0568	0.1444	0.2717

Standard errors in parentheses
* $p < 0.10$

The full report shows results from several of the alternative modeling variants generated in this analysis (NMR 2017b; Appendix A). Included in these results are a variety of models that incorporate drive times from households to lighting retail locations. These include models that differentiate program and non-program retail locations by channel, models that measure the additional drive time to program locations compared to non-program locations by channel, and models that measure the drive time to any retail location by channel. Functional forms of the models also vary: the results include linear, tobit, and probit models. Construction of the dependent variable can be either socket saturation for both types of EE bulb together or separately. Some models restrict observations to MA households only while others include all households with an indicator variable included as a predictor variable. Not incidentally, models in which MA and NY households are pooled find a statistically significant and positive saturation rate effect attributable to households located in MA, consistent with other findings of program effectiveness in MA compared to a state in which there has been little recent market intervention. Finally, the included demographic predictors in the models vary as well.

To estimate the distribution of bulbs implied by these findings, American Community Survey demographic data at the tract level are multiplied by the coefficients estimated above for each Census tract in MA. This yields a predicted saturation rate for the tract, for LEDs and CFLs separately. To estimate the number of program-supported bulbs attributable to each tract, an estimate of the number of sockets in each tract is developed. Table 4 shows average values of sockets per household by the number of rooms in the household, as estimated from the in-home lighting survey data for MA. Like the American Community Survey tract-level data, the number of rooms is top-coded at nine, and the average number of sockets in that cell represents the average of all households of nine rooms or more.

Table 4: Number of Sockets per Household, by Household Size

Household Size (Number Of Rooms)	Average Socket Count per Household
1	9.5
2	13.8
3	21.6
4	26.2
5	41.2
6	49.5
7	68.3
8	77.6
9+	114.7

Multiplying the estimated average number of sockets per household in a tract times the number of households in the tract yields a total number of sockets in the tract. The saturation rate is then multiplied by this estimate of total sockets to estimate the total number of bulbs of that technology type in the tract. The estimate of the number of program bulbs for each tract is generated by calculating the fraction of total bulbs of the technology type in the state that are in the given Census tract. Figure 1 shows predicted saturation rates by Census tract for LED and CFL bulbs.

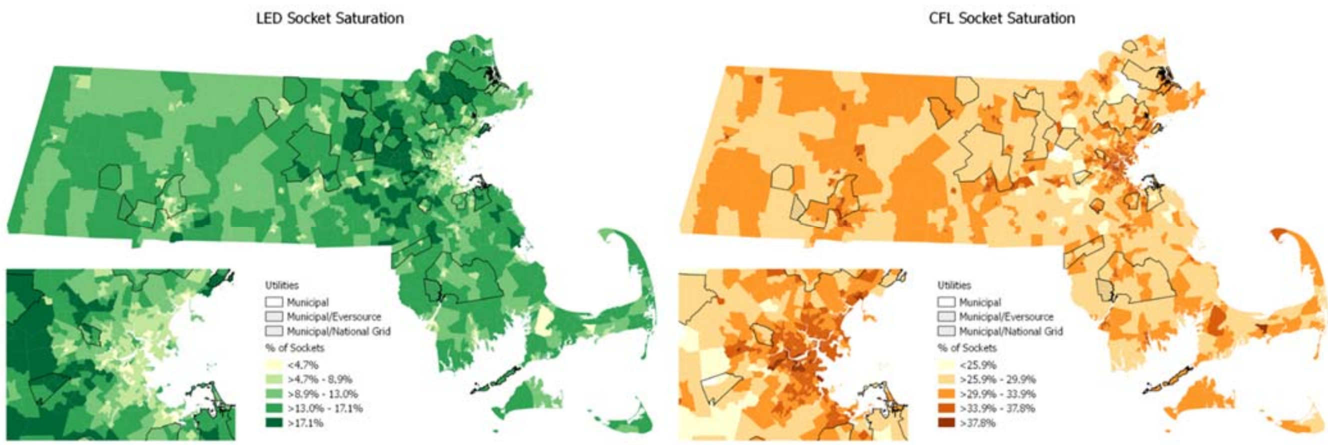

Figure 1: Socket saturation estimates by Census tract for LED and CFL bulbs in MA 2016

Purchase Modeling

As described in the data section above, the in-home surveys of lighting technology use include a panel of households that are revisited annually. This portion of the analysis makes use of the repeat in-home surveys completed by tracking changes in lighting use to infer new bulb purchases. All bulbs in the homes included in the surveys are marked by the visiting technician at the time of the site visit; therefore, any unmarked bulbs at the time of the next visit could then be assumed to be newly acquired. In MA, site visits were completed in 2013, 2014, 2015, and 2016. As a result, data for panel (repeat) visit households from which new bulb purchases can be inferred are available for 2014, 2015, and 2016. In NY, in-home surveys were conducted in 2015 and 2016, resulting in panel purchase information available for 2016. Table 5 reports the number of households with repeat visits for all years included in the resulting dataset, as well as the average lighting count values (after removing from the panel those households that have participated in an in-home installation program). The distribution of bulb purchases by households show a high degree of positive skew (median value is less than the mean). This is a result of a fraction of households that do not make an EE bulb purchase in a given year; of the households in this panel that did make an EE bulb purchase in a year, the range of values of bulb purchases spans from one to 59.

Table 5: Lighting Panel Data Count Values

Source	Number of Households	Average Sockets per Household	Average Bulbs per Household	Average Number of New EE Bulbs per Year per Household
MA 2014	103	54	80	5.9
MA 2015	188	52	74	3.0
MA 2016	246	55	83	4.0
NY 2016	85	70	104	3.7
Total	**622**	**56**	**83**	**4.0**

The preferred model results for use in modeling efforts external to the current scope of work are shown in Table 6. This relatively simple model form is the result of a stepwise model selection approach in which the available demographic predictors were excluded one-by-one (with an exclusion criterion of p>0.2). The resulting model has the two significant predictors in the LED purchase model and the one significant predictor in the CFL purchase model: the models below include all three predictors in both models. Table 7 repeats this model but includes lagged socket saturation as a predictor. In this variable, the observed fraction of sockets in the household filled with bulbs of the given technology type in the previous year is used to predict purchases of that technology type in the current year (i.e., the LED model includes lagged LED saturation while the CFL model includes lagged CFL saturation). A household's prior year socket saturation for the technology type is found to be a statistically significant predictor of new bulb purchases of that type. While that variable is not useful for extrapolating values across the state (since there is no corresponding bulb saturation data in the American Community Survey on which to draw), it is useful in understanding the nature of bulb purchases by households: the strong positive association of lagged bulb saturation with new purchases suggests there is may be some technological "stickiness" in households. That is, to the extent that a household has selected a given technology type for lighting, they continue to replace bulbs with (or cycle in) that same technology type.

Table 1: OLS Model of Demographic Predictors of Efficient Bulb Purchases in MA, 2016

	LED	CFL
Single Family Household	0.479	**1.301***
	(0.993)	**(0.557)**
Renter	**-2.942***	0.324
	(0.972)	(0.545)
College Grad	**1.993***	0.389
	(0.884)	(0.496)
Constant	1.778	0.136
	1.128	(0.633)
N	240	240
Adjusted R^2	0.070	0.016

Standard errors in parentheses
* $p < 0.10$

Table 7: OLS Model of Demographic Predictors of Efficient Bulb Purchases in MA, 2016 Including Lagged Saturation

	LED	CFL
Single Family Household	0.433	**1.247***
	(0.991)	**(0.558)**
Renter	**-2.659***	0.178
	(0.980)	(0.552)
College Grad	**1.985***	0.419
	(0.884)	(0.497)
Lagged Saturation	**7.566***	**1.508***
	(3.806)	**(0.865)**
Constant	1.350	-0.402
	(1.147)	(0.708)
N	239	239
Adjusted R^2	0.081	0.040

Standard errors in parentheses
* $p < 0.10$

Figure 2 shows the estimate distribution of bulb purchases by Census tract implied by the model in Table 6. The results of the purchase model are applied in a way similar to the way they are applied to the saturation model, with some differences due to the difference in the units of the model and to account for negative predicted values for some Census tracts (for detail on the translation of the coefficients to tract-level bulb estimates please see Appendix D of NMR 2017b).

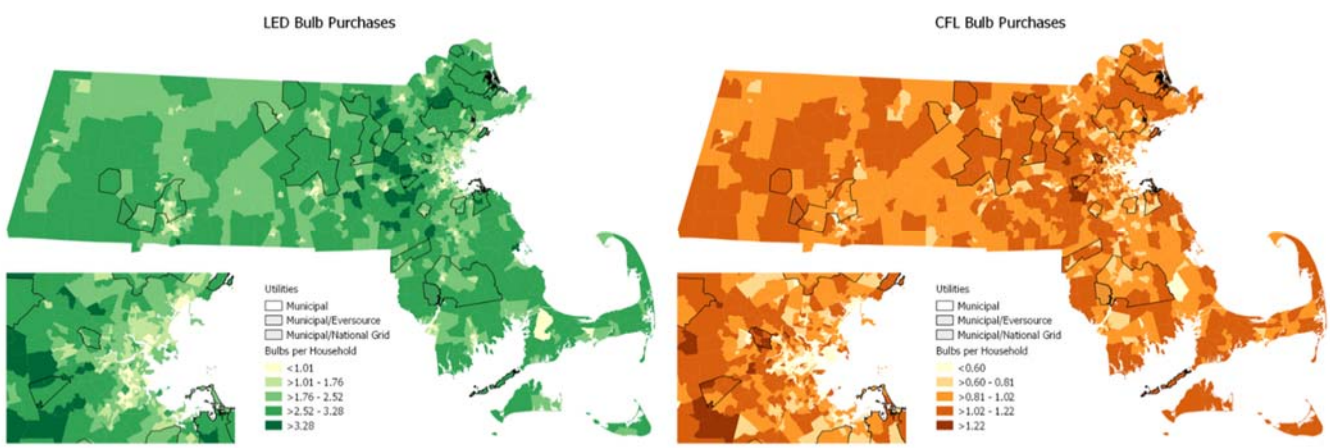

Figure 2: Bulb purchase rates per household estimated by Census tract for LED and CFL bulbs in MA 2016

Effect of Drive Time on Energy Efficient Lighting Use

In the preceding two sections that describe the findings of saturation and purchase modeling, respectively, predictors estimating the effect of drive times from households to retail locations are excluded from the results presented. This is because extensive testing of alternative model forms and variable constructions consistently failed to find a meaningful association between the proximity of households to retail locations and household EE socket saturation or new EE bulb purchases. Table 8 shows an example saturation model in which drive-time variables are included. The coefficient estimates show that there is very little relationship between proximity from households to retail locations. What relationship there is appears to be a spurious finding: furnishing stores (e.g., Bed, Bath & Beyond and Ikea), for example, have a negative association between their drive time proximity to households in the sample and the saturation rate of CFL bulbs but there is no intuitive reason or literature to support why this should be the case.

Table 9 reports results of tests of the joint significance (*F*-tests) of drive time variables for several purchase model implementations. There are three constructions of the drive time variables included in the table. The first, “drive time to nearest of channel,” is a set of drive time variables that represent the drive time, in minutes, from the household to the nearest location for each of the several retail channels, regardless of program/non-program distinction of the retail channel—similar to that shown in Table 8. The second group of drive time variables is “drive time to nearest program store of channel and drive time to nearest non-program store of channel.” These models include a set of drive time variables that estimate the drive time from the household to the nearest location of each channel type separately for program and non-program stores (resulting in nearly twice as many total drive time variables). The third set of models includes drive time variables constructed as the difference between the drive time to a program store and a non-program store, for each retail channel type. This represents the additional drive time required for a household to reach a program store in a given retail channel in comparison to a non-program location (a value that can be negative if the household is nearer to a program store than a non-program store for a retail channel type). Each set of drive time predictors is used in models for each year of purchase data and for each purchase of any EE bulb, LED bulbs, or CFL bulbs. In no case for any of the years of data do drive time predictors jointly provide statistically significant explanatory power to both the LED and CFL models. In the models that do show joint significance among the drive time predictors, the same predictors (i.e., store types) are not consistently significant, nor are they directionally consistent with intuitive expectations: several show that proximity to retail locations is associated with lower household EE bulb purchase rates. The inconsistency of the individual predictors coupled with the directionally counter-intuitive effects of those that are individually significant supports the conclusion that these statistically significant findings are likely spurious, and not representative of a causal relationship.

Table 8: Tobit model with EE, LED, or CFL socket saturation as the dependent variable; MA households only

		EE (LED+CFL)	LED	CFL
Demographics	Renter	0.01825	**-0.08861***	**0.06821***
		(0.03411)	**(0.03176)**	**(0.03195)**
	College Grad	0.01096	0.03325	-0.00127
		(0.02794)	(0.02611)	(0.02617)
	Number of Rooms	-0.01277	**0.01759***	**-0.01734***
		(0.00788)	**(0.00725)**	**(0.00738)**
Drive Time to Store Type	Bargain	0.00030	0.00123	-0.00045
		(0.00131)	(0.00117)	(0.00122)
	Bodega	0.00633	0.00548	0.00423
		(0.00521)	(0.00463)	(0.00489)
	Club	0.00473	-0.00010	0.00427
		(0.00324)	(0.00301)	(0.00304)
	Discount	-0.00386	-0.00421	-0.00178
		(0.00549)	(0.00501)	(0.00516)
	Ethnic	0.00417	0.00291	0.00328
		(0.00351)	(0.00319)	(0.00330)
	Furnishings	**-0.00588***	-0.00148	**-0.00531***
		(0.00291)	(0.00273)	**(0.00273)**
	Grocery	-0.00148	0.00017	-0.00197
		(0.00359)	(0.00329)	(0.00337)
	Hardware	0.00256	0.00696	-0.00031
		(0.00620)	(0.00560)	(0.00581)
	Home Center	-0.00300	-0.00043	-0.00182
		(0.00417)	(0.00379)	(0.00390)
	Lighting	0.00214	0.00253	0.00057
		(0.00413)	(0.00377)	(0.00388)
	Mass Merchandise	-0.00110	0.00121	-0.00135
		(0.00374)	(0.00343)	(0.00351)
	Office Supply	-0.00286	-0.00451	-0.00026
		(0.00384)	(0.00354)	(0.00360)
	Supermarket	0.01215	0.00438	0.00496
		(0.00885)	(0.00795)	(0.00830)
	Thrift	0.00053	-0.00230	0.00114
		(0.00234)	(0.00215)	(0.00220)
	Corner Drug	-0.00062	-0.00287	0.00078
		(0.00253)	(0.00232)	(0.00237)
	Drugstore	**-0.01563***	-0.01309	-0.00357
		(0.00906)	(0.00808)	(0.00850)
	Constant	**0.57342***	-0.05463	**0.47461***
		(0.06540)	(0.06110)	**(0.06126)**
	N	487	487	487
	pseudo-R^2	0.1569	0.2103	0.3692

Standard errors in parentheses
* $p < 0.10$

Table 9: Tests of joint significance (*F*-test) for drive time variables

Sample	EE (LED+CFL)	LED	CFL
	Drive time to Nearest of Channel		
MA 2014	0.67 (0.8154)	0.63 (0.8513)	0.58 (0.8890)
MA 2015	1.10 (0.3574)	0.48 (0.9550)	1.48 (0.1142)
MA 2016	1.48 (0.1089)	1.50 (0.1034)	0.89 (0.5762)
NY 2016	0.87 (0.6048)	1.08 (0.3964)	0.53 (0.9179)
	Drive time to Nearest Program Store of Category and to Nearest Non-Program Store of Channel		
MA 2014	0.88 (0.6408)	0.90 (0.6098)	0.89 (0.6352)
MA 2015	1.35 (0.124)	0.65 (0.9176)	**1.68 (0.0237)**
MA 2016	1.30 (0.1502)	1.27 (0.1735)	1.13 (0.309)
	Additional Drive time to Nearest Program Store of Channel		
MA 2014	0.44 (0.9561)	0.99 (0.471)	0.69 (0.7746)
MA 2015	**1.68 (0.066)**	0.59 (0.8695)	**1.95 (0.0252)**
MA 2016	**1.79 (0.0421)**	**1.64 (0.0707)**	1.20 (0.2777)

p-value in parentheses

Study Limitations

The findings of this study related to the effects of driving time are limited by the evaluation team's inability to capture effects of home-to-work commuting, modes of transportation other than driving (i.e., walking, biking, and public transportation), and regular non-lighting shopping patterns. This work uses road-network drive time estimates from a household to retail locations as a proxy for retail availability (Moore et al. 2014 and Boscoe, Henry, and Zdeb 2012). But it is not necessarily the case, nor even likely, that each purchase of a lighting product originates with the purchaser at their home and driving directly to the retail location. It is likely that lighting purchases are made on shopping trips with other, non-lighting, primary purposes; and that these trips may not originate from home but may start from a work location, a point convenient along a commuting route, or any number of other possible locations a resident may find themselves. An additional analysis of total spending per shopping trip in which a lighting purchase was made (using a separate data set), for example, found that lighting expenditures were only a small fraction of total expenditures on the trip—indicating that the lighting purchase was not necessarily the primary purpose for the excursion (NMR 2017c).

Another limitation of this work relates to the selection of the sample from which lighting purchase patterns are being inferred. In the in-home sample, on which results are reported in sections 2 and 3, the household occupants are aware that lighting technicians from an energy-efficiency evaluation firm are planning to visit their homes to catalogue their lighting equipment. It is possible that this knowledge would systematically change the average behavior of those occupants with respect to lighting purchase decision-making in a way that makes the sample different from the population of interest at large. However, that survey research effort annually recruits new households to the panel in order to compare the characteristics of new households to repeat visit households and finds that this effect is, at most, quite modest (NMR 2016).

Since this work is being used, in part, to inform ongoing modeling efforts in MA, the recommendation of which lighting distribution model to use as an input to that work is constrained by the availability of data to extrapolate those findings across the state. For example, a finding that prior year (lagged) socket saturation for an efficient bulb type is a useful predictor of current year lighting purchases is not a result that can be applied to a statewide model, absent a census on prior year socket saturation. In the purchase model reported, a recommendation is made for a simplified model that excludes prior year saturation for this reason. This does not preclude using these pieces of analysis to generate insights to better understand the distribution of lighting throughout the state, however.

Finally, the results of this analysis include a considerable amount of uncertainty. The model forms that are have been selected retain relatively low overall predictive power, resulting in a correspondingly high degree of remaining unexplained variability in the distribution of lighting products. Additionally, extrapolating the model predictions across the state results in several out-of-range predictions that must be adjusted for the purposes of input to other state-wide modeling exercises. As these results apply to those exercises, appropriate caution should be taken in the degree of sensitivity of those models to these inputs.

Conclusions

This analysis used a survey of in-home lighting inventories coupled with a database of lighting retail locations to assess the relationship of household proximity to retail outlets and in-home efficient lighting technology uptake. This analysis is novel in that it applies a method distinct from the geospatial approach described for estimating cross-service area sales in the Uniform Methods Protocol (Dimetrosky, Parkinson, and Lieb 2014). In the end, we conclude that drive time to retail stores does not have a discernable statistical effect on efficient lighting purchases or in-home efficient lighting saturation. In several modeling variations, the form of the model, the construction of the dependent variable, and the construction of the drive time predictor variables resulted in either no statistical effect, or weak and directionally inconsistent statistical effects. This finding is not an assessment of the efficacy of program activities; indeed, models including a pooled sample of MA and NY households find that MA households have statistically higher EE socket saturation.

It is worth noting the relationship between drive time and straight-line distance. The two measures of proximity are strongly correlated in the values generated in this analysis. Estimating straight-line distance is far more straightforward to estimate from a computational and data-preparation strand-point however (Jones et al. 2010). To the extent that proximity to retail locations is of interest to future work, it is likely that much of the relationship will be found using the easier to estimate straight-line distances (Phibbs and Luft 1995). If a significant relationship is found with a measure of straight-line distance, a subsequent more detailed analysis using drive times may then be warranted.

Some demographic factors have a small, but statistically significant, effect on efficient bulb purchases and saturation. LED and CFL bulb saturation is higher in MA households compared to NY households, renters have fewer LED bulbs installed but more CFL bulbs installed compared to non-renters, and larger homes tend to have a higher fraction of LED bulbs but a lower fraction of CFL bulbs. Households in which at least one member is a graduate of a four-year college purchase more LED bulbs, and single family households purchase more CFL bulbs (compared to households in multifamily buildings).

For future work that will rely on these modeling efforts as an input, we recommend utilizing the saturation model based on demographic predictors, but excluding drive time, as an input. The results of the purchase model can be used to test this input’s sensitivity. We also recommend ignoring PA service territory boundaries when considering the distribution of program-subsidized bulbs to households in the state since there is no evidence of an effect of retail proximity provided by the estimates of drive time from households to retail locations.

References

Boscoe, F.P, K.A. Henry, and M.S. Zdeb, 2012. *A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals*. Prof Geogr. Apr 1; 64(2). Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835347/

Dimetrosky, S., K. Parkinson, and N. Lieb, 2014. *Chapter 6: Residential Lighting Evaluation Protocol.* The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. February 2014. NREL/SR-7A30-53827

Jones, S.G., A.J. Ashby, S.R. Momin, and A. Naidoo, 2010. *Spatial Implications Associated with Using Euclidean Distance Measurements and Geographic Centroid Imputation in Health Care Research.* Helth Serv Res. 2010 Feb; 45(1): 316-327. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2813451/

Moore, S.F., G.M. Braman, G. Stiles, and B. Stull, 2014. *Conquering Leakage, Breakage and Equitable Allocation by Dialing-Up Big Data.* 2014 ACEEE Summer Study on Energy Efficiency in Buildings. Available at http://aceee.org/files/proceedings/2014/data/

NMR, 2016. *2015-16 Lighting Market Assessment Consumer Survey and On-site Saturation Study*. August 2016. Available at http://ma-eeac.org/studies/residential-program-studies/

NMR, 2017a. *RLPNC 16-7: 2016-17 Lighting Market Assessment Consumer Survey and On-site Saturation Study.* April 2017. Available at http://ma-eeac.org/studies/residential-program-studies/

NMR, 2017b. *Lighting Distribution Modeling.* March, 2017. Available at http://ma-eeac.org/studies/residential-program-studies/

NMR, 2017c. *RLPNC: 16-3 Lighting Decision Making.* December 2016. Available at http://ma-eeac.org/studies/residential-program-studies/

Phibbs, C.S. and H.S. Luft, 1995. *Correlation of Travel Time on Roads Versus Straight Line Distance*. Med Care Res Rev. 1995 Nov; 52(4): 532-542.

U.S. Census Bureau, 2013. *Median household income, American Community Survey 5-year estimates.*