Evaluating the Cost-Effectiveness of Rebate Programs for Residential Energy-Efficiency Retrofits

Joe Maher
University of Maryland, Department of Agricultural and Resource Economics, College Park, MD 20742, jmaher@arec.umd.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC , August 4-6, 2013.

Introduction

Buildings account for 42 percent of energy use and 38 percent of CO_2 emissions in the United States (USGBC 2011). In recent years, State and Federal governments have increased funding programs that subsidize energy-efficient retrofits to existing buildings. For example, the American Recovery and Reinvestment Act of 2009 (ARRA) included $17 billion for energy-efficiency programs, which helped to initiate $54 billion in energy-related home improvements in 2009 (von Schrader 2010). In 2013, President Obama announced a new goal, "Let's cut in half the energy wasted by our homes and businesses over the next 20 years. We'll work with the states to do it."[1]

Despite the immense policy importance of these investments, surprisingly little research directly assesses the effectiveness of retrofit rebate programs. Allcott and Greenstone (2012) argue, "much of the evidence on the energy cost savings from energy-efficiency comes from engineering analyses or observational studies that suffer from a set of well-known biases." Davis et al. (2012) and Metcalf and Hassett (1999) show that engineering simulations over-predict the energy savings from retrofits by 150% to 400% compared against estimates from billing data. The literature characterizing simulation-bias is small. It only considers three retrofit options among dozens, and only verifies national-level engineering predictions, rather than preferred household-level predictions. Observational studies rely on utility-level data, rather than household data, creating numerous empirical problems, including omitted variable bias, questionable control groups, and poorly defined treatments that aggregate heterogeneous retrofit programs (see Gillingham et al. 2009 for a review of literature).

The purpose of this paper is to evaluate retrofit-specific residential rebate programs based on observed household-level billing data. I identify the energy savings from fifteen retrofit rebate programs in Gainesville, Florida using a panel dataset of electricity and natural gas consumption and building characteristics for 30,000 residences. The difference-in-difference method compares changes in energy use within a residence before and after an energy-saving retrofit intervention (treatment group) with changes in energy use within a similar residence that did not receive improvements (control group). A unique feature of the data, which is central to my identification strategy, is that I have monthly billing data combined with time-variant and time-constant characteristics of each residence.

Preliminary work makes three contributions to the literature assessing energy efficiency programs. First, this is the first assessment of a retrofit rebate program to apply difference-in-difference methods linking billing data and housing characteristics for every customer within a utility service area. Second, by assessing 9 retrofit programs, this study explores heterogeneity across a diverse range of retrofits options using billing data. Third, this study is the first to use project-level data on engineering predictions and rebate-levels to identify retrofit-specific estimates of simulation bias and cost-effectiveness, permitting a carefully matched control group.

[1] Statement delivered on February 12, 2013 in a State of the Union Speech. Obama continues to promise that "Those states with the best ideas to create jobs and lower energy bills by constructing more efficient buildings will receive federal support to help make that happen."

My results reveal important heterogeneities across retrofit programs. First, estimates of cost per kilowatt hour saved vary widely, ranging between 1 to 28 cents depending on retrofit type. Second, results indicate that engineering simulations seriously over-predict energy savings when compared against empirical estimates, though bias varies widely across retrofit types. Thus policy makers could use these results to expand cost-effective programs and eliminate ineffective ones, information that is unavailable from the utility-level program averages found in the literature.

The first section below describes my data. Then section 2 presents an overview of testable hypotheses that extend current literature. Sections 3 and 4 present the empirical model and preliminary results. Section 5 concludes.

Data

Table 1 compares descriptive statistics between two separate datasets, each identifying participants in GRU retrofit rebate programs. Due to time constraints, the results presented in subsequent sections are generated from a partial dataset, however, a complete dataset will be used in future analyses.

The University of Florida dataset is a partial sample of rebate participants for nine residential rebate programs used to generate preliminary regression results presented in Table 2. For each property address, the partial sample includes information on the retrofit installation date and type of rebate program. The Gainesville Regional Utility dataset is a full sample of rebate participant for fifteen residential rebate programs and three commercial rebate programs.

Empirical Model

The sample of households is restricted in several ways. First, the analysis is restricted to single family households that have a single customer account during the four-year time period. Second, all treatment households are required to have at least 6 months of billing data pre-retrofit and post-retrofit to ensure consistent estimation of treatment effects. Third, the treatment group only includes households that receive a single retrofit intervention; households that receive multiple retrofits are excluded from the analysis.

Similar to Davis et al. (2012), I employ a difference-in-difference model to estimate the effect of retrofits on household energy consumption. Specifically, I use the partial dataset to estimate the following two-way fixed effects model:

$$y_{it} = \lambda_t + c_i + \tau_j \omega_{ijt} + \varepsilon_{ijt}\,, \qquad t = 1, \dots, T, \qquad j = 1, \dots, J \qquad (1)$$

where y_{it} is electricity consumption for house i in month t; λ_t are month-specific indicator variables capturing city-wide trends that affect electricity consumption over time, such as weather fluctuations; c_i are household-specific indicator variables capturing all time-constant factors of a house that affect electricity consumption; ω_{ijt} are treatment indicator variables that are equal to 1 for all months t after a retrofit installation in homes i that participate in a rebate program for retrofit type j, and equal to 0 otherwise. T is a constant equal to the maximum number of monthly bills for house i, which is 143 months on average, and J is a constant equal to the total number of rebate programs. ε_{ijt} is an error term clustered by house and represents unmeasured time-variant factors affecting electricity consumption. The coefficient τ_j is a vector of retrofit-specific ATEs, or the average monthly energy savings from retrofit type j, that is assumed to persist over time.

A two-way fixed effects model imposes very weak identifying assumptions. Specifically, my model maintains the standard fixed effect assumption of strict exogeneity of treatment, expressed as $Cov(\omega_{ijt}, \varepsilon_{is}) = 0 \; \forall \; t, s.$ Endogeneity problems violating this assumption arise from four common sources: period effects, measurement error, unobserved heterogeneity, and simultaneity.

The model directly addresses period effects. Month fixed effects eliminate time-variant period effects common to all households in Gainesville within a billing period, as non-participant households serve as counterfactuals that isolate period effects. Treatment effects are identified from remaining *within variation beyond period effects* (time trends).

To address other forms of endogenity, the preliminary model assumes: (i) non-participants in the comparison group never receive unobserved retrofit interventions (measurement error); (ii) treatments are uniform for each retrofit type (measurement error) (iii) retrofit interventions are uncorrelated with time-variant changes affecting energy consumption, such as changes in occupancy or home-remodels (omitted variable bias); and (iv) pre-retrofit shocks to energy bills never trigger retrofit interventions (simultaneity bias).

Preliminary Results

Baseline Energy Savings Estimates

Table 2 presents the main results of the energy savings from retrofit installation. Electricity usage is reported in kilowatt hours per month. Coefficients reflect average monthly treatment effects across retrofit types, represented by τ_j in equation (1). If retrofit installations increase efficiency, then energy use should decrease post retrofit for participating households. This is, in fact, the case for most retrofits.

Results confirm expectations, suggesting most retrofits reduce electricity use. At the upper extreme, a retrofit can save an average of 130 kWh per month, or 12 percent of the median household energy consumption of 1,100 kWh per month[2]. Five programs have results significant

[2] 1,100 kWh is the median energy consumption for the median house averaged across all months. Energy consumption varies seasonally. Preliminary estimates do not explore heterogeneous treatment effects across seasons, although many retrofits, such as air-conditioning systems, may cause season-specific energy savings.

at the 1 percent level (energy savings reported in parentheses); including, refrigerator buyback (49 kWh per month), pool pump replacement (129 kWh per month), super SEER air-conditioner replacement (116 kWh per month), duct leakage repair (64 kWh per month), and low-income energy efficiency upgrades (107 kWh per month). Energy savings for attic insulation installment (33 kWh per month) are significant at the 5 percent level. Three programs have no statistically significant energy savings: high-efficiency room air-conditioner replacement, air-conditioner maintenance, and low-interest energy-efficient loans.

Comparison with Ex-Ante Engineering Models

Table 3 compares my results with GRU ex-ante predictions derived from engineering simulations. Ex-ante bias calculations report the percentage that engineering predictions deviate from my estimates of realized energy savings. Importantly, ex-ante and ex-post energy savings estimates come from different samples[3], so these comparisons are included only to illustrate likely findings from more future analysis.

Results indicate that GRU engineering simulations consistently overestimate the energy savings from residential retrofits. On average, ex-ante assessments predict more than double the realized energy savings estimated using billing data. However, bias varies considerably across retrofit types, with some engineering predictions closely matching realized energy savings. Relative bias is reported for the six programs with significant energy savings (percentage of over prediction reported in parenthesis), including, low-income weatherization (-4%), pool pump replacement (14%), duct leakage repair (171%), super SEER air-conditioner replacement (172%), refrigerator buyback (268%), and attic insulation installment (389%). Interestingly, room air-conditioner replacement and air-conditioner maintenance – two programs with zero energy savings – have considerably lower ex-ante predictions than any other retrofit types. In terms of kilowatt hours, the absolute bias is smaller for these zero-effect programs than for two-thirds of retrofits with significant observed energy savings.

Cost-Effectiveness

Preliminary results suggest that program cost-effectiveness varies considerably across retrofit types. Following standard assumptions in the literature (see, e.g., Arimura, Li, Newell, and Palmer 2011), cost-effectiveness estimates assume a 5% discount rate and 10-year product lifetime[4]. Average cost savings per kilowatt hour range from 1 cent to 27 cents among programs based on rebate levels and energy savings estimated from electricity bills. Program-specific

---

[3] Table 3 describes the samples used for ex-ante and ex-post estimates. GRU does not report expected energy savings for the low-interest loan program, so comparisons focus on the other eight rebate programs.

[4] Davis et al. 2012 assumes a 5-year treatment effect for an appliance retirement program, arguing that household would naturally replace appliances after 5 years; rebates only prompt households to replace them earlier. Since GREP rebates require participants to meet stringent energy-efficiency standards to qualify for rebates – standards that homeowners may not meet otherwise – the cost-effectiveness calculations used in this section assume a longer 10-year treatment effect suggested by others.

results are significant at the 1 percent level for six programs, including; refrigerator buyback (1 cent per kWh), pool pump replacement (3 cents per kWh), super SEER air-conditioner replacement (5 cents per kWh), duct leakage repair (6 cents per kWh), attic insulation installment (12 cents per kWh), and low-income weatherization (27 cents per kWh). Programs without significant energy savings are cost-ineffective regardless of rebate levels. Cost-effectiveness estimates change predictably under alternative assumptions. Assuming a shorter 5-year treatment effect, costs per kilowatt hour saved increase by 78 percent.

Homeowners also benefit from lower electricity bills. Based on GRU's block-rate pricing in 2009 and the distribution of monthly consumption levels observed in the data, the average marginal price that GRU charges consumers is 14.6¢. This consists of an 8.6¢ average energy charge plus a 6¢ fuel adjustment charge. These numbers imply that annual electricity savings per household range across retrofits, including, refrigerator buyback ($86.65), pool pump replacement ($225.33), super SEER air-conditioner replacement ($203.68), duct leakage repair ($111.70), attic insulation installment ($58.30), and low-income weatherization ($187.50). Programs without significant energy savings are cost-ineffective regardless of rebate levels.

Discussion

These results present estimates of energy savings, simulation-bias, and cost-effectiveness for a large number of programs compared to other studies using micro data of energy consumption. Three GRU retrofit programs can be directly compared with existing estimates from the literature, including, air-conditioner replacements, refrigerator removal, and attic insulation. Results for remaining programs provide new insights about heterogeneity between retrofit types in terms of cost-effectiveness and simulation-bias. In general, these results provide new optimism that retrofits rebates can attain cost-effective demand reduction.

Cost-Effectiveness

Unlike previous studies, my results indicate that well-designed retrofit programs can achieve similar cost-effectiveness as other leading energy-efficiency programs. In Gainesville, a portfolio of rebate retrofits targeting refrigerator buybacks, pool pump replacements, super SEER central AC replacements, and duct leakage repair can achieve an energy savings at a cost between 1 and 6 cents per kilowatt hour avoided. By comparison, Alcott (2011) reports cost-effectiveness measures for peer-comparison programs from OPOWER to range from 2 to 5 cents per kilowatt hour saved. I am not aware of any other studies using household billing data that identify rebate programs with cost-effectiveness on par with OPOWER.

My research has unusual policy relevance and external validity, as GRU employs best practices that demonstrate the potential of well-designed retrofit rebate programs. By contrast, any retrofit rebate can be cost-prohibitive if rebates are too generous, or energy-savings are too low. For example, Davis et al. (2011) acknowledge that the Mexican rebate program they evaluate has a flawed program design that allows participants to purchase virtually any new appliances regardless of energy-efficiency[5]. As a result, Davis et al. (2011) find refrigerator replacement rebates are costly, averaging 17 cents per kilowatt hour removed[6]. At the opposite extreme, GRU is an award-winning national leader in energy conservation[7] requiring households and contractors to report detailed product information to verify that installed retrofits meet stringent program standards for energy-efficiency. Unlike the inefficient Mexican program, GRU operated a refrigerator buyback program averaging 1 cent per kilowatt hour removed. Thus, my study makes and important contribution to literature by demonstrating the potential of best-practices, thereby rebutting literature that focuses on under-preforming programs.

However, my results agree with some claims in the literature that select retrofits are cost-prohibitive, including room air-conditioning and attic insulation. Hypothetically, a GRU program focused only on attic insulation, low income grants, room air-conditioning replacements, and central air-conditioning maintenance would be inefficient with cost between 12 to 27 cents, or more, for each kilowatt hour saved. This claim echoes findings by Davis et al. (2011) that room air-conditioning replacements do not provide any energy savings; making any amount of rebate cost-prohibitive. Similarly, Metcalf and Hassett (1999) also argue that attic insulation interventions are not cost-effective, although they lack detailed information about project costs.

Simulation-Bias

For the three previously studied retrofit types, simulation-bias is comparable to estimates available in the literature. Davis et al. (2012) find engineering simulations over-predict savings by 150% to 250% for refrigerator replacements; while GRU simulations over-predict energy saving by 168% for refrigerator buybacks[8]. In the same study, Davis et al. (2012) find that room air-conditioner replacements actually *increase* in energy consumption; GRU estimates also suggest that room air-conditioner replacements may increase energy use, though estimates are not significant. Metcalf and Hassett (1999) suggests that engineering simulations over predict returns from attic insulation by as much as 400%; while GRU estimates suggest an over-

[5] The Mexican program rules require participants to purchase refrigerators and air-conditioners that are 5% more efficient than minimum energy-efficiency standards established in 2002. The rebate program was administered from 2009 to 2011, during which time almost any new appliances available on the market met or exceeded program requirements based on outdated energy-efficiency requirements.

[6] 17 cent cost-effectiveness is based off of a 10-year treatment effect and 5% discount rate, which parallel assumptions used in Table 3.

[7] Awards include the 2005 Green Power Beacon Award presented by the U.S. Environmental Protection Agency, the U.S. Department of Energy and the Center for Resource Solutions. (http://www.epa.gov/greenpower/documents/2005awards.pdf; https://www.gru.com/AboutGRU/NewsReleases/Archives/Articles/news-2005-10-28.jsp)

[8] Refridgerator treatments differ between the two studies, so direct comparisons of kilowatt hour saved are not appropriate. Davis et al. (2012) is a replacement program where an old refridgerator is removed after a new one is purchased. The GRU buyback program simply removes an old refridgerator, which in many cases is a second refridgerator in the house that is not replaced. However, comparisons of simulation-bias between the studies are more valid as engineering models are adjusted to account for specific program requirements.
refrigerator buyback[8] despite a higher level of energy savings of 594 kilowatt hours per year.

prediction bias of 289%[9]. Although previous studies only evaluate basic country-level engineering simulations, GRU project-specific engineering simulations do not improve the accuracy of predictions.

Unlike existing literature, my study finds that engineering simulation-bias varies widely across retrofit types. Despite reasonable agreement with other studies on room air-conditioner, refrigerator buyback and attic insulation, these three retrofit types have unusually large simulation-biases. Broad claims about simulation-bias for other retrofit interventions would be misleading. In fact, my results that engineering-bias can be very low for certain retrofit types. For example, simulations are surprisingly accurate for pool pump retrofits and low-income grant programs, with average predictions within 4% and 14% of actual energy savings[10]. Thus, an important contribution of this study is to reveal that engineering simulations are much more accurate for some retrofits missing from existing literature.

Given the reliance on simulations for retrofit program evaluations, it is worthwhile to qualify my assessments of simulation-bias using on observed energy savings. Simulation-bias captures unpredicted exogenous shocks, such as weather events, in addition to bias caused by simulation error. Weather shocks may particularly affect retrofits related to space conditioning, such as central air-conditioner replacement, duct leakage repair, and attic insulation. Adjusting simulation to account for actual weather conditions could improve identification of underlying-bias of engineering predictions. In fact, empirical studies that rely on a small sample of post-retrofit billing data may identify energy savings that are unrepresentative of long-term treatment effects, especially if weather is abnormal. By assuming a constant 10-year treatment effect for cost-effectiveness calculations, this study implicitly assumes that observed post-retrofit weather accurately reflect future weather conditions. Thus, one way of minimizing simulation-bias from exogenous shocks is extending the length of post-retrofit billing data until post-weather trends converges to historical baselines used by engineering software. [* By including retrofits as early as 2007, our estimates use a much longer post-estimation period than other studies. By contrast, Davis uses only 2 years of billing data in total *]

Conclusion

Despite the widespread implementation of retrofit rebate programs and calls for increased investment in demand side management programs, surprisingly little is known about whether energy-efficiency retrofits are an effective way to reduce energy consumption. Engineering simulations provide most of the evidence, but simulated predictions, even if based on sound

[9] Metcalf and Hassett (1999) find the attic insulation reduces total household energy consumption by 9%, while contemporary engineering simulations predicted energy savings as high as 50%. GRU attic insulation retrofits achieve about a 3.5% reduction in household energy consumption, with simulations predicting energy savings of about 13%.

[10] I have not run hypothesis tests to determine predictions are significantly different from point estimates. Hypothesis test may require more information about the error reported for engineering simulations.

models, do not account for installation quality or behavioral responses. Hence there is an important and timely need for empirical research that uses field data to more fully evaluate the effects of energy-efficiency retrofits on energy consumption.

The primary contribution of this paper is the evaluation of retrofit-specific residential rebate programs based on actual billing data. Using retrofit programs in the city of Gainesville, I find an immense variation in energy savings and cost-effectiveness across retrofit types. Three out of nine programs fail to achieve any energy savings. For the remaining six programs, average realized energy savings are less than half of predicted savings from engineering simulations. Though direct comparison with engineering simulations is challenging, my estimates reveal a systematic over-prediction bias that raises new doubts about the reliability of simulated predictions. I also estimate the costs and benefits of the retrofit rebate program. Cost-effectiveness varies widely across retrofit programs, varying between 1 and 27 cents per kilowatt hour saved. Five out of nine programs achieve energy savings below 18 cents per kilowatt hour, or the average cost of new electricity generation in Gainesville. The six effective retrofit types benefit homeowners with reductions in annual electricity expenditures between $58 and $225.

These contributions provide new policy insights about the effectiveness and cost-effectiveness of retrofits. First, results inform policymakers about the relative efficacy of different retrofit rebates, allowing inefficient programs to be terminated and efficient programs to be expanded. Second, results also provide new empirical evidence about the over-prediction bias of engineering models, suggesting a need for future research to validate ex-ante program evaluations. Third, these evaluations empower homeowners to make informed decisions about energy-efficiency investments using credible information on the expected cost-savings from various retrofit options.

Appendix: Residential Retrofit

Table 1. Descriptive Statistics for Rebate Participation Data

	Partial Sample[1]	Complete Sample[2]		
Rebate Program	Frequency	Frequency	Rebate	Ex-Ante Savings
Duct Leakage Repair	222	2,702	$359	1.3
AC Maintenance	401	3,158	$55	0.5
High Efficiency Room AC	130	812	$161	0.2
Super SEER Central AC	393	792	$550	2.4
Attic Insulation	393	2,033	$361	1.6
Refrigerator Buyback (I)	391	3,131	$64	1.6
Refrigerator Buyback (II)		544	$51	1.2
Pool Pump (I)	113	797	$307	1.8
Pool Pump (II)		284	$251	1.8
Low Income Grants (LEEP)	213	695	$2,703	1.2
Low-Interest Loan	61	211	$606	
Central AC (I)		2,430	$549	1.9
Central AC (II)		1,182	$294	0.5
Natural Gas Water Heater		1,657	$392	2.8
Home Performance Energy Star		3,398	$377	1.6
Solar Water Heater		129	$477	1.5
Solar Photovoltaic		122	$6,207	3,504.3
Window Replacement/Film		159	$184	0.7
Customized Business		511	$8,724	243.2
LED Exit Signs		365	$469	0.2
Vending Machines		305	$287	1.8
Total	2,317	25,558		

1 University of Florida dataset is a partial sample of rebate participants for nine residential rebate programs used to generate preliminary regression results presented in Table 2.
2 Gainesville Regional Utility dataset is a full sample of rebate participant for fifteen residential rebate programs and three commercial rebate programs. Average rebate amounts (dollars) and average ex-ante engineering predictions of energy savings (megawatt hours per year) by rebate program are aggregated from project-level rebates and energy saving predictions.

Table 2. Two-Way Fixed Effect Estimates of Electricity Savings[1]

Dependent variable: kilowatt hours per month		
Duct Leakage Repair	-63.75	***
	(17.95)	
AC Maintenance	0.04	
	(11.70)	
High Efficiency Room AC	24.74	
	(18.91)	
Super SEER Central AC	-116.26	***
	(11.84)	
Attic Insulation	-33.27	**
	(16.90)	
Refrigerator Buyback	-49.46	***
	(10.78)	
Pool Pump	-128.61	***
	(28.46)	
Low Income Grants	-107.02	***
	(20.61)	
Low-Interest Loan	-39.71	
	(33.97)	
Constant	912.44	***
	(3.93)	
Fixed Effects		
MonthxYear	YES	
House	YES	
No. Observations	412,292	
No. Houses	9,450	
R^2	0.341	

[1]Variables are program specific treatment indicators equal to 1 for billing months after retrofit installation for houses participating in that specific program and equal to 0 otherwise. Coefficients represent the electricity changes caused by retrofits (average treatment effect). Results generated using partial sample including 2,317 treatment homes and 7,133 control homes (Table 1. includes program-specific observation counts). Numbers in parentheses are standard errors clustered by house. Asterisks denote statistical significance: ** denotes significance at the 5% level and *** denotes significance at the 1% level.

Table 3. Ex-Ante Energy Savings Bias and Rebate Cost-Effectiveness

	Energy Savings[1]			Cost Savings[2]	
Retrofit Program	Ex-Post	Ex-Ante	Ex-Ante Bias	Rebate	$/kWh
Duct Leakage Repair	0.8	1.3	71 %	$ 359	0.06
AC Maintenance		0.5		$ 55	
High Efficiency Room AC		0.2		$ 162	
Super SEER Central AC	1.4	2.4	72 %	$ 550	0.05
Attic Insulation	0.4	1.6	289 %	$ 361	0.12
Refrigerator Buyback	0.6	1.6	168 %	$ 64	0.01
Pool Pump	1.5	1.8	14 %	$ 307	0.03
Low Income Grants	1.3	1.2	- 4 %	$ 2,703	0.27

[1] Energy savings are reported in megawatt hours per year. Ex-Post energy savings calculated using the estimates in Table 2 multiplied by 12 to reflect annual changes and divided by 1,000 to covert units to megawatt hours. Ex-Ante engineering simulations are program averages of project-level GRU predictions from the full dataset. Reported ex-post and ex-ante estimates are based on different treatment samples. Ex-Ante bias represents the percentage that ex-ante engineering predictions overestimate realized energy savings.

[2] Cost-savings are reported in dollars per kilowatt hour saved. Cost-savings calculations use average rebate amount and average ex-post energy savings assuming a 10-year treatment effect and 5% annual discount rate for all programs.