Demand-Side Management and Energy Efficiency Revisited

Maximilian Auffhammer, Carl Blumstein** and Meredith Fowlie****

The key finding of Loughran and Kulick (2004) is that utilities have been overstating electricity savings and underestimating costs associated with energy efficiency demand-side management (DSM) programs. This claim is based on point estimates of average DSM-related savings and costs implied by an econometric model of residential electricity demand. We first argue that the chosen test statistics bias results in favor of rejecting the null hypothesis that utility-reported savings reflect true values. We also note that utility estimates of average program savings and costs are rejected based on point estimates alone. We use the same data and econometric model to estimate the appropriate test statistics. We then construct nonparametric bootstrap confidence intervals. These intervals are quite large; we fail to reject the average electricity savings and DSM costs reported by utilities. Our results suggest that the evidence for rejecting utility estimates of DSM savings and costs should be re-interpreted.

1. INTRODUCTION

As public concerns about climate change and air quality escalate, there is increasing political pressure to find ways to reduce the environmental impacts of energy use. One approach currently being pursued by policymakers involves increasing support for "demand-side management" (DSM) programs. Since the 1970s, utilities in the United States have been implementing DSM programs de-

The Energy Journal, Vol. 29, No. 3.

* University of California, Berkeley.

** University of California Energy Institute.

*** Corresponding author: University of Michigan, Department of Economics, Weill Hall, 735 S. State St. #5224, Ann Arbor, MI 48109-3091. Phone: (734) 615-9595 Fax: (734) 213-0245 E-mail: mfowlie@umich.edu.

We thank David Loughran for making the data and code of their study available to us. The authors would further like to thank the editor and four anonymous referees, Peter Berck, Severin Borenstein, Howard Chong, John DiNardo, Jeffrey LaFrance, Catherine Wolfram and the participants of the UC Energy Institute 2006 Electricity Camp for helpful suggestions. All remaining errors are ours.

signed to reduce residential and commercial electricity demand through information dissemination programs, subsidies, free installation of more efficient technologies, and other conservation related activities. Whereas program evaluations routinely find that these utility-sponsored DSM programs are highly cost effective (EPRI, 1984; Eto et al., 1996; Eto et al., 2000; Fickett et al., 1990; Jordan and Nadel, 1993; Nadel, 1992; Nadel and Geller, 1996), in the past some economists have viewed these results with skepticism (Joskow and Marron, 1992; Nichols, 1995).[1] A more recent paper by Loughran and Kulick (2004) has refueled this debate.

The stated objective of the Loughran and Kulick (LK) paper is "to test whether DSM expenditures during the 1990s succeeded in increasing the electricity efficiency of the U.S. economy" (p 21). LK fail to reject this hypothesis, however they do conclude that "DSM (has) had a much smaller effect on retail electricity sales than estimates reported by utilities themselves" (p. 19). This claim has attracted considerable attention. In the two years since its publication, this paper has been cited in a wide range of contexts, including utility revenue requirement hearings (B.C. Commission, 2005), academic papers (Gellings et al., 2007; Gillingham et al., 2006; Metcalf, 2006), policy briefs (Geller and Attali, 2005), and partisan position papers (Crane and Boaz, 2005).

Several authors have pointed out shortcomings of methods used to calculate DSM savings and costs, including the potential for free riding, unmeasured positive spillovers, and moral hazard issues.[2] LK derive their result from a novel approach to addressing the problem of free riders (that is, beneficiaries of a utility DSM program who would have saved energy even in the absence of a DSM program). While questions could be raised about whether LK have successfully dealt with the free rider issue, we do not examine such questions in this response. Instead, our objective is to demonstrate that the empirical evidence provided by the authors is consistent with (rather than contradicts) the findings of past DSM program evaluations. We use a simple hypothesis testing framework to show that DSM savings estimates reported by utilities to the Energy Information Administration (EIA) cannot be rejected even when the data and estimation approach used by LK are taken at face value.

This response proceeds as follows: Section 2 restates the question addressed by LK in terms of a hypothesis test; Section 3 uses the data and econometric models used by LK to estimate the appropriate test statistics; Section 4 reports the results of hypotheses testing; Section 5 concludes.

2. FORMULATING THE NULL HYPOTHESIS

In the past, studies demonstrating the cost effectiveness of DSM programs have relied heavily on cost and savings estimates that the utilities are re-

1. For a review of energy efficiency policies and their estimated impacts, see Gillingham et al., 2006.

2. For a review of the major criticisms and merits of DSM program evaluations, see Geller and Attali (2005).

quired to report annually to the Energy Information Administration (EIA). Each year, utilities are not only required to report their annual DSM expenditures (denoted *EE*) and electricity sales (*kWh*), but also to estimate the annual savings (*s*). LK use these data from 324 utilities over the period 1989-1999 to estimate several models of DSM electricity savings.[3]

The first aspect of the LK paper we take issue with is the statistic used to test the stated null hypothesis. In order to test the hypothesis of whether DSM expenditures increased the energy efficiency of the US economy, one needs to consider the percent change in *aggregate* US electricity consumption due to *aggregate* expenditure on energy efficiency DSM. We have verified that LK use the average percent change in electricity consumption due to energy efficiency DSM expenditures across utilities and years as their indicator. As we will show below, this choice of test statistic results in an underestimation of percent savings and an overestimation of costs, assuming that we are interested in measuring economy wide savings and costs.

A simple example helps to illustrate this point. Suppose utility A spends \$1 on DSM and saves 20 kWh, producing 980 kWh instead of 1,000 kWh in the counterfactual. Now consider utility B, which spends \$5 on DSM and saves 250 kWh, producing 4,750 kWh instead of 5,000 kWh in the counterfactual. Utility A saves 2% at 5 cents per kWh and utility B saves 5% at 2 cents per kWh. In order to obtain average savings, LK use the average of the (econometrically estimated) percentage savings across utilities, 3.5% to test their stated hypothesis. We argue that the appropriate savings are 270kWh/6,000kWh = 4.5%. LK use their estimate of savings to calculate average costs. For this example, using their approach one would arrive at an average cost of (\$1+\$5)/3.5% • ((1000kWh +5000kWh))=2.86 cents per kWh saved, while we argue that the appropriate number is (\$1+\$5)/(20+250)=2.22 cents per kWh saved.

Some additional notation helps to make these concepts more precise and assists in framing the statistical analysis which follows in the next two sections. Let n index utilities: n=1...N. Let t index years. The nth utility reports electricity sales, DSM related expenditures and savings in t=1...T$_n$ years.[4] The level of electricity consumption reported by utility n in year t after spending EE_{nt} on DSM programs is $kWh(1)_{nt}$. We let kWh(0) represent electricity demand in the counterfactual, unobserved situation where no DSM program is in place. Energy savings are s_{nt}. $kWh(1)_{nt}+s_{nt}=kWh(0)_{nt}$. If there are savings from DSM programs, $s_{nt}\geq 0$. If DSM results in increased electricity consumption $s_{nt}<0$. In order to obtain an estimate of s_{nt} one can either rely on the utility reported figures or use econometric estimates.

3. The sample used by LK contains only 327 of the 3,254 utilities reporting to the EIA each year. The majority of utilities do not report any DSM expenditures in the period 1992-1999. Only 119 report positive DSM expenditures throughout the study period. The energy efficiency component of DSM expenditures is reported separately by utilities beginning in 1992. LK approximate the percentage of total DSM expenditures that a utility allocated to energy efficiency for years 1989-1991 based on the percentage reported by the utility in 1992.

4. This is not a balanced sample. 156 utilities report DSM expenditures and savings in all eight years. Others do not report in all years. Several report expenditures in only one year.

As illustrated in the example above, the LK approach to summarizing savings is to take an *unweighted* average across observations. Let S_1 represent the true, unweighted average of percentage savings across all utilities and years. Let C_1 represent the measure of average cost chosen by LK, which is based on the unweighted average measure of savings S_1. Utility reported savings percentages and costs can be used to construct estimates of these two population parameters:

$$\hat{S}_1 = \frac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} \left(\frac{kWh(0)_{nt} - kWh(1)_{nt}}{kWh(0)_{nt}} \right)}{\sum_{n=1}^{N} T_n} \tag{1}$$

$$\hat{C}_1 = \frac{\dfrac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} EE_{nt}}{\sum_{n=1}^{N} T_n}}{\hat{S}_1 \left(\dfrac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} kWh(0)_{nt}}{\sum_{n=1}^{N} T_n} \right)} \tag{2}$$

In calculations of $\hat{S}_1$, savings reported by utilities who spend relatively small amounts on DSM are weighted the same as savings reported by utilities with very large DSM programs. Utilities who spend more on DSM programs report significantly larger percentage savings on average. If we are interested in the average returns per dollar spent on DSM programs, these measures will potentially be misleading.

Alternative measures of average savings and costs weight observations by electricity sales and program expenditures, respectively. Let S_2 represent total savings attributable to DSM programs divided by total electricity sales in the absence of DSM programs. Let C_2 represent the total DSM expenditures divided by total electricity savings. Utility reported data can be used to construct the following estimates of these two population parameters:

$$\hat{S}_2 = \frac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} (kWh(0)_{nt} - kWh(1)_{nt})}{\sum_{n=1}^{N} \sum_{t=1}^{T_n} kWh(0)_{nt}} \tag{3}$$

$$\hat{C}_2 = \frac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} \mathrm{EE}_{nt}}{\sum_{n=1}^{N} \sum_{t=1}^{T_n} (kWh(0)_{nt} - kWh(1)_{nt})} \tag{4}$$

These measures are more informative summaries of the overall average returns per dollar spent on DSM programs, and the average cost per kWh saved. Using utility reported savings, consumption and DSM expenditures, we construct estimates of these two sets of population parameters using the complete dataset and the five subsets of the data analyzed by LK. These summary statistics and the sample sizes used to calculate them are reported in columns (2) and (1) of Table 1, respectively. Note that estimates of percentage savings averaged across all utility-year observations ($\hat{S}_1$) are consistently smaller than average percentage savings weighted by electricity sales ($\hat{S}_2$). This is because smaller utilities (who tend to spend relatively less on DSM programs and report lower percentage savings) are weighted relatively more heavily in calculations of $\hat{S}_1$ as compared to $\hat{S}_2$.

Estimates of $\hat{C}_1$ are larger than estimates of $\hat{C}_2$. There are two reasons for this. The first has to do with the positive relationship between DSM expenditures and percentage savings. Observations associated with smaller DSM expenditures (and thus higher per kWh costs on average) are weighted relatively more heavily in calculations of $\hat{C}_1$. Past research has found that DSM programs with larger budgets have been comparatively more successful in delivering energy savings at lower costs (Eto, 1996). In non-residential programs, utilities can use information and program replication for companies in similar industries to achieve economies of scale. One could also imagine achieving economies of scale in residential customer awareness campaigns and incentive programs if program administrative costs and transaction costs can be spread out over a larger number of households.[5]

The second reason has to do with outliers in the reported savings data. There are 1,459 utility-year observations in which utilities report both expenditures and savings, which makes it possible to calculate $/kWh saved by utility and year. In a small number of cases, very small reported savings imply extremely high costs (i.e., above $100/kWh in five cases). A closer look at the data reveals that these unusually small savings (relative to expenditures) are typically associated with the first year of reporting by utilities overseeing relatively small DSM programs.[6] These outliers are discussed in more detail in the following section.

5. We are unaware of any published studies that distinguish between economies of scale that are realized in larger DSM programs (holding utility size constant), and economies of scale realized by larger utilities. However, when we regress reported savings on DSM expenditures, utility sales, and an interaction between DSM expenditures and utility sales, all coefficients are positive and statistically significant at the one percent level. This suggests both kinds of economies of scale are present. One could imagine that larger utilities are more efficient in administering DSM programs because they have more staff and expertise to support these programs.

6. For example, utilities that report costs on the order of $400/kWh saved in the first year of their DSM program consistently report costs of $0.02/kWh in subsequent years.

Table 1. Point Estimates and 90% & 95% Confidence Intervals For Reported and Estimated Savings and Costs

	(1) Sample Size LK	(2) Utility Reported	(3) LK Estimate	(4) ABF Estimate (n=774)	(5) Hansen 90% CI Full Sample (n=774)	(6) Hansen 90% CI Big Spenders (n=205)	(7) Hansen 95% CI Full Sample (n=774)	(8) Hansen 95% CI Big Spenders (n=205)
Savings (S_1) in %								
Model 1	1815	1.51%	0.40%	1.22%	[0.55 : 1.89]	[0.79 : 2.73]	[0.41 : 2.02]	[0.60 : 2.92]
Model 2	1815	1.51%	0.30%	0.80%	[-0.21 : 1.80]	[-0.35 : 2.55]	[-0.40 : 1.99]	[-0.63 : 2.82]
Model 3	2373	1.53%	0.30%	0.43%	[-0.36 : 1.53]	[-0.56 : 2.19]	[-0.54 : 1.71]	[-0.82 : 2.45]
Model 4	774	1.99%	1.20%	1.22%	[0.55 : 1.89]	[0.79 : 2.73]	[0.41 : 2.02]	[0.60 : 2.92]
Model 5	998	2.13%	0.60%	0.43%	[-0.36 : 1.53]	[-0.56 : 2.19]	[-0.54 : 1.71]	[-0.82 : 2.45]
Savings (S_2) in %								
Model 1	1815	1.82%	0.70%	1.70%	[0.76 : 2.64]	[0.83 : 2.90]	[0.57 : 2.81]	[0.61 : 3.10]
Model 2	1815	1.82%	0.40%	1.01%	[-0.38 : 2.38]	[-0.43 : 2.62]	[-0.65 : 2.63]	[-0.73 : 2.89]
Model 3	2373	1.86%	0.50%	0.78%	[-0.54 : 2.08]	[-0.60 : 2.28]	[-0.79 : 2.32]	[-0.89 : 2.55]
Model 4	774	2.58%	1.70%	1.70%	[0.76 : 2.64]	[0.83 : 2.90]	[0.57 : 2.81]	[0.61 : 3.10]
Model 5	998	2.79%	0.90%	0.78%	[-0.54 : 2.08]	[-0.60 : 2.28]	[-0.79 : 2.32]	[-0.89 : 2.55]
Costs (C_1) in ($/kWh)								
Model 1	1815	$0.031	$0.140	$0.063	[0.011 : 0.112]	[0.010 : 0.100]	[-0.031 : 0.114]	[-0.027 : 0.103]
Model 2	1815	$0.031	$0.220	$0.097	[-0.324 : 0.536]	[-0.309 : 0.520]	[-0.756 : 0.981]	[-0.735 : 0.945]
Model 3	2373	$0.031	$0.170	$0.133	[-0.511 : 0.765]	[-0.475 : 0.708]	[-1.149 : 1.424]	[-1.074 : 1.305]
Model 4	774	$0.031	$0.060	$0.063	[0.011 : 0.112]	[0.010 : 0.100]	[-0.031 : 0.114]	[-0.027 : 0.103]
Model 5	998	$0.027	$0.120	$0.133	[-0.511 : 0.765]	[-0.475 : 0.708]	[-1.149 : 1.424]	[-1.074 : 1.305]
Costs (C_2) in ($/kWh)								
Model 1	1815	$0.026	$0.079	$0.047	[0.007 : 0.078]	[0.009 : 0.094]	[-0.022 : 0.08]	[-0.025 : 0.097]
Model 2	1815	$0.026	$0.130	$0.078	[-0.281 : 0.436]	[-0.333 : 0.519]	[-0.644 : 0.806]	[-0.764 : 0.951]
Model 3	2373	$0.023	$0.096	$0.101	[-0.387 : 0.579]	[-0.454 : 0.693]	[-0.884 : 1.061]	[-1.047 : 1.267]
Model 4	774	$0.024	$0.048	$0.048	[0.007 : 0.078]	[0.009 : 0.094]	[-0.022 : 0.080]	[-0.025 : 0.097]
Model 5	998	$0.021	$0.078	$0.101	[-0.387 : 0.579]	[-0.454 : 0.693]	[-0.884 : 1.061]	[-1.047 : 1.267]

In summarizing these data and formulating a null hypothesis, LK report that average utility-estimated DSM-related electricity savings range between 1.8 and 2.3 percent and that the average per kWh program costs reported by utilities range from \$0.02-\$0.03/kWh (p.39). Although LK do not explain how they calculate these summary statistics, Table 1 shows that these ranges are more consistent with expenditure weighted averages $\hat{S}_2$ and $\hat{C}_2$ as compared to the weighted average $\hat{S}_1$ and the cost measure $\hat{C}_1$.

3. DERIVING AVERAGE SAVINGS AND COSTS FROM ECONOMETRIC ESTIMATES

When estimating energy savings from DSM programs, we want to know how the level of electricity consumption we observe after implementing a DSM program (denoted kWh(1)) differs from what electricity consumption would have been in the absence of the program (denoted kWh(0)). Of course, we can only observe the former, so we are left to construct our best estimate of what demand would have looked like had there been no DSM program in place.

LK estimate an econometric model explaining variation in the approximate percent change in observed electricity sales[7] :

$$\begin{aligned} \Delta kWh_{nt} &= \ln kWh(1)_{nt} - \ln kWh(1)_{nt-1} \\ \Delta kWh_{nt} &= \beta_0 \ln EE_{nt} + \beta_1 \ln EE_{nt-1} + \beta_2 \ln EE_{nt-2} + \gamma Z_{nt} + \varepsilon_{nt} \end{aligned} \tag{5}$$

Energy efficiency DSM expenditures enter as the current, single and double lag. The vector Z_{nt} contains, depending on the specification, a combination of the change in the number of customers, gross state product (GSP), price of electricity and substitutes, climate, share of electricity sold to different users as well as year fixed effects, state-year fixed effects or a state-specific quadratic time trend. Using results from a least squares regression of equation (5), the difference in log transformed electricity sales attributable to contemporaneous and lagged DSM expenditures is estimated as follows:

$$\hat{s}_{nt} = \beta_0 \ln EE_{nt} + \beta_1 \ln EE_{nt-1} + \beta_2 \ln EE_{nt-2} \tag{6}$$

The five different specifications control for variation in different sets of the observed utility characteristics, state characteristics, year effects, and energy prices. The samples used for each of the five models differ in the number and size of DSM program cost outliers. Table 2 reports the number of utility-year observations associated with reported costs that exceed \$10/kWh and \$100/kWh,

7. The difference in a log transformed kWh_{nt} is approximately equal to the percentage change in kWh_{nt}, provided this percentage is small. Using the log difference as the dependent variable also decreases concerns about heteroskedasticity, which would necessitate the use of a robust covariance estimator or weighted least squares.

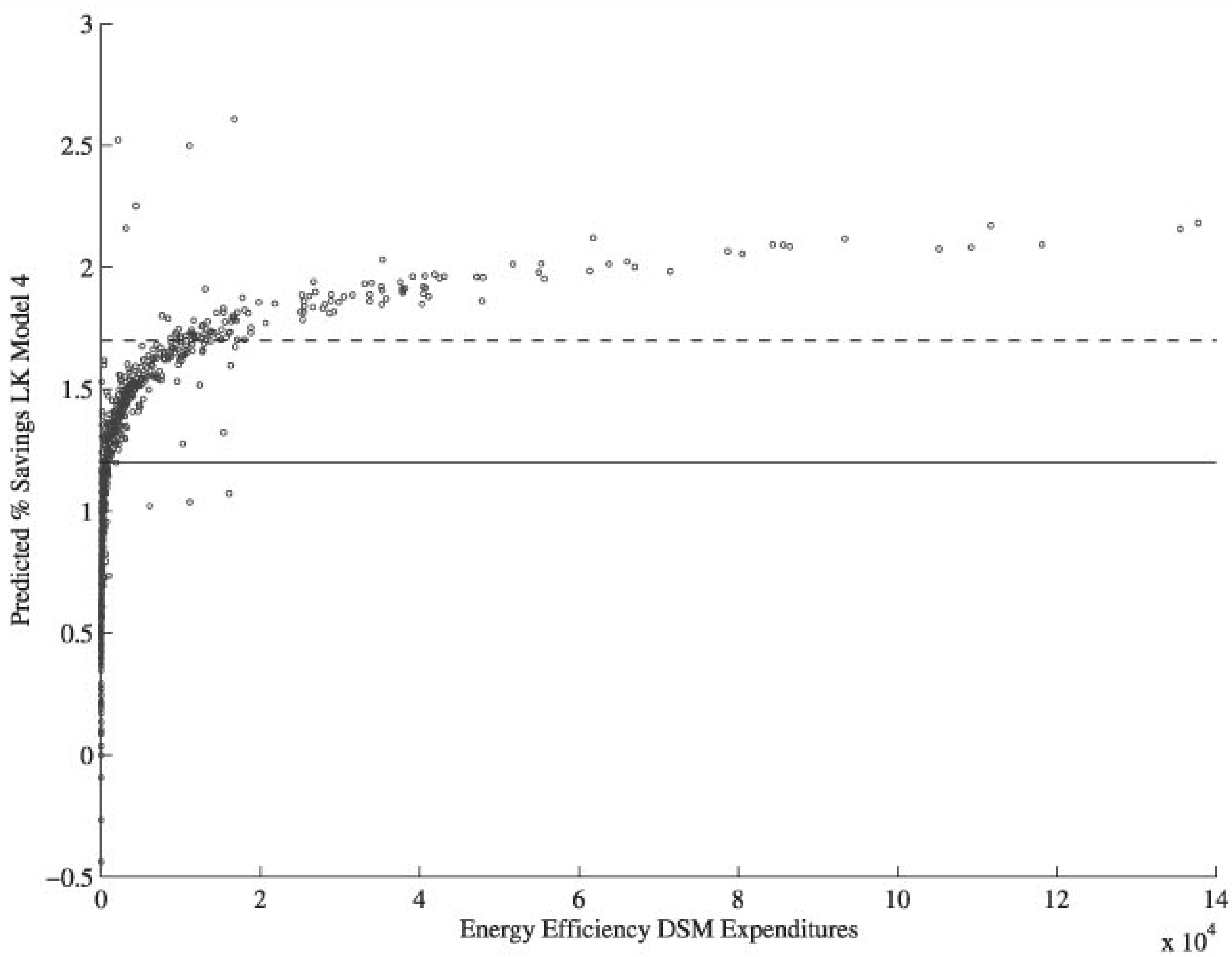
Table 2. Reported Costs per kWh Saved Outliers

Model	# Observations	# Observations above $10/kWh	# Observations above $100/kWh
1	1815	24	3
2	1815	24	3
3	2273	26	3
4	774	7	0
5	998	26	3

respectively. The subset of the data used to estimate Model 4 contains the fewest outliers. It is also the specification that directly controls for the most confounding factors (including energy prices and GSP).

The sample used for Model 4 corrects an additional issue not raised so far. LK treat DSM policy as exogenous, which is similar to a random assignment assumption. In reality a utility has to decide whether to engage in DSM activities and if so, how much to spend on DSM. Utility specific fixed effects do not account for this two step process. An instrumental variables estimation procedure would be required here, yet in that absence of valid instruments, this is not feasible. One

Figure 1. Energy Efficiency DSM Expenditures and Predicted Savings



alternate approach would be to estimate the effect of a dollar spent on DSM conditional on program participation. The sample used for Model 4 does exactly that, since it omits any utility year with zero expenditures. We restrict our analysis to this sample. We leave the investigation of the impacts of accounting for selection in program effectiveness to future research.[8] Subsequent analysis in this response will emphasize the sample used for Model 4, which leaves us with three distinctly different specifications, since Models 1 and 4 as well as Models 3 and 5 have identical specifications and differed only by the number of observations used.

Estimated regression coefficients, together with reported DSM expenditures, are used to generate estimates of electricity savings attributable to DSM expenditures. LK use these utility-year specific estimates of percentage electricity savings $\hat{s}_{nt}$ to generate an estimate of the population parameter S_1. For each regression model, LK report the average percent savings across utilities:

$$\hat{S}_1 = \frac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} \hat{s}_{nt}}{\sum_{n=1}^{N} T_n} \tag{7}$$

This econometrically estimated *average* is used to compute an estimate of average costs per kWh saved at the mean of the data, which uses $\hat{S}_1$ above to obtain the predicted kWh saved for the cost calculation:

$$\hat{C}_1 = \frac{\dfrac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} EE_{nt}}{\sum_{n=1}^{N} T_n}}{\hat{S}_1 \left(\dfrac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} kWh(0)_{nt}}{\sum_{n=1}^{N} T_n} \right)} \tag{8}$$

We have argued that the weighted averages S_2 and C_2 are the preferred test statistics, since they relate more closely to the stated null hypothesis. We construct estimates of these parameters using LK's econometric estimates of $\hat{s}_{nt}$. Table 1 presents the estimates of S_1 and C_1 that LK report in the paper as well as the point estimates of S_2 and C_2, which we argue are the preferred estimates.

The sales-weighted savings implied by the econometric estimates are consistently larger than the unweighted average $\hat{S}_1$. Figure 1 helps to illustrate why this is so. The percentage savings estimates $\hat{s}_{nt}$ constructed by using estimated regression coefficients from Model 4 are plotted against reported expenditures EE_{nt}. The DSM-related energy savings implied by the regression model are larger among utilities that spend more on DSM programs. The solid line corresponds

8. Since the statistic of interest is a weighted average of $\hat{s}_{nt}$, which depends on three estimated coefficients, the sign of the bias from failing to account for selection is indeterminable *a priori*.

to the unweighted average ($\hat{S}_1$ =1.2%). The broken line corresponds to the sales weighted average $\hat{S}_2$ = 1.7%). Because of the strong positive relationship between DSM expenditures and electricity savings, $\hat{S}_1 < \hat{S}_2$.

4. HYPOTHESIS TESTING

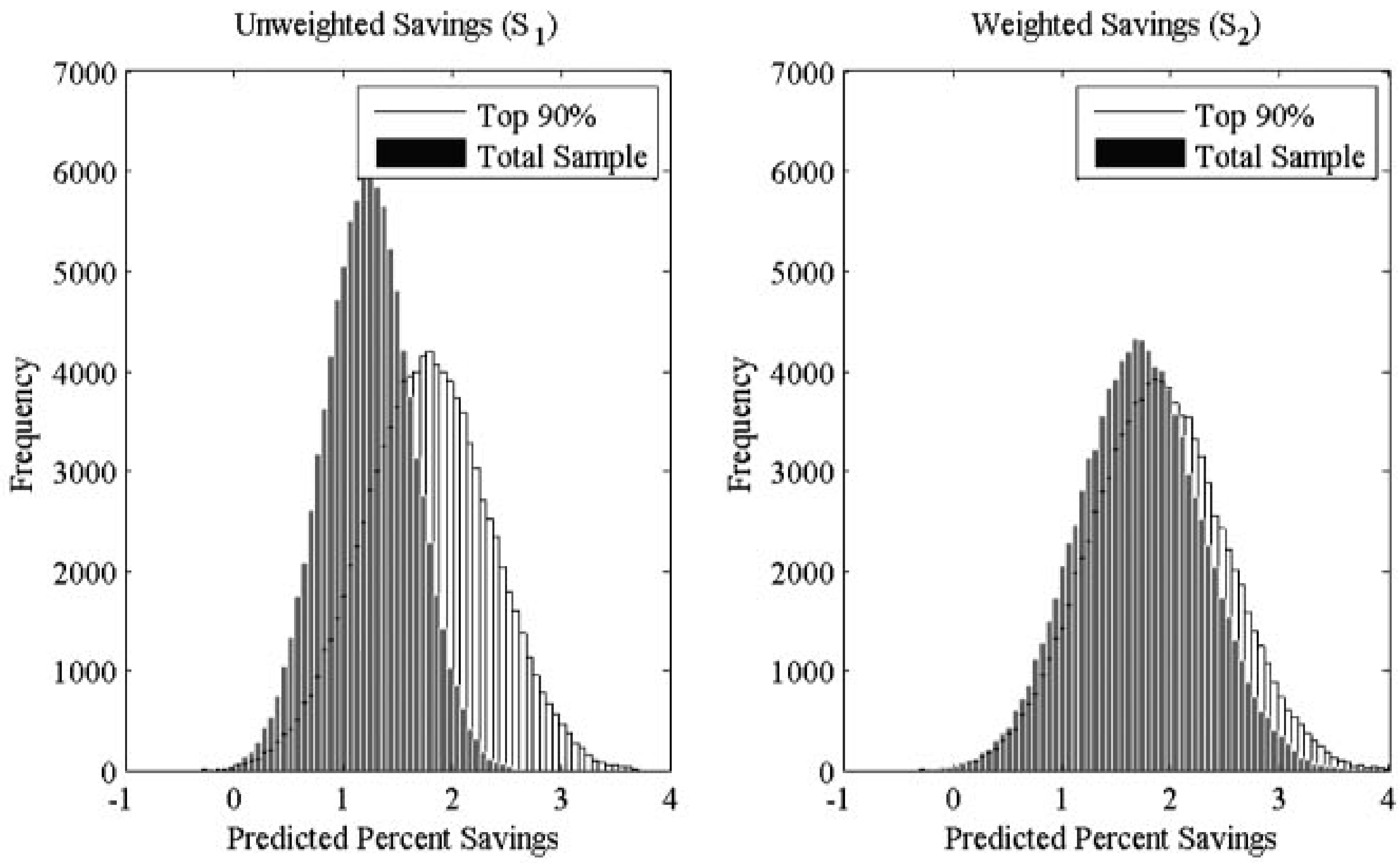
LK correctly note that the effects of DSM program expenditures implied by their regression coefficient point estimates are small relative to the estimate of average utility-reported savings. Because the former are arguably less contaminated by selection bias, they conclude that the true effects of DSM programs on retail electricity sales were much smaller than those reported by utilities. They further conclude that true average program costs per kWh saved exceed those reported by utilities.

Our second criticism pertains to these provocative conclusions. In order to formally reject a null hypothesis, it is not sufficient to establish that the point estimate of a test statistic is not equal to the value of the statistic under the null hypothesis. A null hypothesis can only be rejected if we are sufficiently certain that the observed value of the test statistic would not occur if the null hypothesis were true. In this section, we estimate confidence intervals around the point estimates of average savings and costs in order to explicitly account for the variance of the estimated regression coefficients. Column (4) of Table 1 reports the point estimates of weighted and unweighted savings and costs using the sample LK use for Model 4. As expected, the estimates of savings are slightly higher for most models, since the observations used here are all DSM program participants.

In the simplest of cases, the standard error of an estimate of a population parameter can be calculated analytically based on standard assumptions about the distribution of the observed data in the population. In the current context, however, analytical approaches to constructing confidence intervals are very complex because the predicted savings (i.e., the $\hat{s}_{nt}$) are not independent within utilities by construction. We thus resort to a heteroskedasticity robust nonparametric residual bootstrapping technique called the wild bootstrap (Liu, 1988). We first estimate each model and record residuals for each utility. We transform the residuals using the transformation suggested by MacKinnon and White (1985). As suggested by Davidson and Flachaire (2001), we then resample by utility and with replacement from the transformed residuals multiplied by a Rademacher variable. We repeat this for 100,000 bootstrap replications. For each replication, the unweighted average $\hat{S}_1$ and the weighted average $\hat{S}_2$ are estimated and recorded for the entire sample and for the subsample that accounts for 90% of total DSM expenditures, which we call the "big spenders."[9] Figure 2 represents these four bootstrap distributions. The shaded histogram in each panel corresponds to the full sample, the clear framed histogram corresponds to using only observations responsible for 90% of total program expenditures. The shaded distribution in the left panel is the

9. The findings are robust to specifying the cutoff at 80% and 95% as well. At 90% this accounts for 205 utility/year observations or roughly 26.5% of the observations used to estimate Model 4.

Figure 2. Bootstrap Distribution of Average Savings



one underlying the LK results. For the S_1 measure, the distributions for the full and 90% sample are quite different. For the S_2 measure, as expected, the two distributions are quite similar since the large spenders are weighted more heavily.

Quantiles of these bootstrap distributions can be used to construct percentile confidence intervals when the underlying distributions of the test statistics are symmetric, as is roughly the case for the savings distributions. If the underlying distributions are asymmetric, as is the case for the average costs distribution, these percentile intervals can perform poorly. Hansen (2007) suggests an alternative approach to constructing confidence intervals that have proper coverage probability when test statistic distributions are not symmetric. These two types of 90% and 95% confidence intervals are constructed for the point estimates of S_1, S_2, C_1 and C_2. Results are reported in columns (5) and (7) of Table 1. The corresponding estimates using the "big spenders" are reported in columns (6) and (8) of Table 1.

Based on their point estimates of S_1, LK conclude that the true average electricity savings attributable to DSM are less than 1.8%. As Table 1 shows, the 90% and 95% confidence intervals for the weighted savings include savings significantly greater than 1.8%. For our preferred test statistic, the 95% confidence interval includes savings up to 2.81% for the full set of utility years and up to 3.10% for the top DSM spenders.

Similarly, LK note that their estimates of costs per saved kWh are higher than the average costs reported by utilities ($0.02-$0.03/kWh). Due to the asymmetry of the cost distribution, the appropriate confidence interval is given by

Table 3. Comparison of Bootstrapped Confidence Interval by CI Type and Bootstrapping Technique for Model 4

Variable	CI Type	Bootstrap Type	Weighted 90%	Weighted 95%	Unweighted (LK) 90%	Unweighted (LK) 95%
Savings	Quantile	NP	[0.44 : 1.93]	[0.30 : 2.07]	[0.32 : 1.38]	[0.22 : 1.48]
Savings	Hansen	NP	[0.43 : 1.92]	[0.29 : 2.05]	[0.32 : 1.38]	[0.22 : 1.49]
Savings	Quantile	NP Wild	[0.77 : 2.65]	[0.59 : 2.84]	[0.54 : 1.89]	[0.42 : 2.02]
Savings	Hansen	NP Wild	[0.76 : 2.64]	[0.57 : 2.81]	[0.55 : 1.89]	[0.41 : 2.02]
Costs	Quantile	NP	[0.04 : 0.17]	[0.037 : 0.235]	[0.055 : 0.237]	[0.051 : 0.327]
Costs	Hansen	NP	[-0.007 : 0.123]	[-0.071 : 0.126]	[-0.009 : 0.173]	[-0.099 : 0.177]
Costs	Quantile	NP Wild	[0.029 : 0.101]	[0.027 : 0.130]	[0.040 : 0.141]	[0.038 : 0.183]
Costs	Hansen	NP Wild	[0.007 : 0.078]	[0.022 : 0.080]	[0.011 : 0.112]	[-0.031 : 0.114]

Hansen (2007). Costs are very imprecisely estimated as indicated by the large confidence intervals. The 90% confidence interval for costs includes costs of 0.7 cents as well as costs of up to 7.8 cents. Given the imprecision of these estimates, we fail to reject costs close to zero as well as some rather large costs. It should be noted that the models for which LK estimated the highest costs are also the models with the largest confidence intervals. Table 3 provides the confidence intervals for Model (4) using regular quantile based confidence intervals as well as the Hansen (2007) confidence intervals by bootstrap technique. It shows quite clearly that using the confidence intervals with the correct coverage probability combined with a heteroskedasticity robust bootstrap has a significant impact on the estimated confidence intervals.

5. CONCLUSIONS

DSM programs have the potential to play an important role in mitigating the environmental impacts associated with meeting increasing demand for electricity end-uses. Past program evaluations and utility-reported data have indicated that these programs are highly cost effective. In some respects, Loughran and Kulick (2004) offer empirical evidence that is broadly consistent with the earlier literature. They find that DSM expenditures during the 1990s succeeded in increasing the electricity efficiency of the U.S. economy. However, the finding that has attracted the most attention since this paper was published is that the effects of DSM are "small relative to what the utilities themselves report" (p. 38), implying that programs are not as cost effective as previously thought. In this response, we identify two issues which, when properly addressed, suggest that this provocative claim cannot be supported by the data.

First, we contend that LK's choice of test statistic biases results in favor of rejecting the null hypothesis that utility-reported electricity savings reflect true values. Utilities who spend more on DSM programs report significantly larger percentage savings on average. An unweighted average of reported savings weights smaller utilities with small DSM program expenditures and large utilities with significant program expenditures equally. This summary statistic underestimates the electricity savings per dollar spent on DSM programs. It is worth noting that the LK estimates imply that there may exist economies of scale in implementing DSM. This is important since it implies that DSM (or at least the energy efficiency component of DSM) could be more effectively administered at the state level rather than by individual utilities.

Second, we note that LK reject utility estimates of DSM savings and costs based on point estimates alone. A null hypothesis can only be rejected if we are sufficiently certain that the observed value of the test statistic would not occur if the null hypothesis were true. We estimate confidence intervals around the point estimates of average savings and costs implied by LK's regression coefficient estimates. Using both the weighted and unweighted measure of savings, we fail to reject the null of (utility reported) savings between 2% and 3%, which LK rejected based on a point estimate alone. The appropriate bootstrapped confidence interval for costs per kWh saved contains values significantly lower and values significantly higher than those reported by the utilities.

This analysis, which uses the same data and the same econometric models used by these authors, implies that utility-reported savings and costs cannot be rejected on the basis of these econometric results. The estimates of average savings and costs implied by the regression coefficient estimates are consistent with the average effects reported by the utilities themselves over the same study period.

REFERENCES

British Columbia Commission (2005). "British Columbia Commission in the matter of the Utilities Commission Act S.B.C. 1996, chapter 473 and British Columbia Hydro and Power Authority 2004/05 and 2005/06." Revenue Requirements Application and An Application by British Columbia Transmission Corporation for Approval of an Application for Deferral Accounts. Vol. 20 Proceedings. Vancouver, B.C.

Crane, E.H. and D. Boaz (2005). *Cato Handbook on Policy*. 6th Edition. Cato Institute. Washington DC.

Davidson, R. and E. Flachaire (2001). "The wild bootstrap, tamed at last." *Working paper, IER#1000*, Queen's University.

Electric Power Research Institute (EPRI) (1984). "Demand Side Management. Vol. 1: Overview of Key Issues." EA/EM-3597. Palo Alto, CA.

Eto, J., E. Vine, L. Shown, R. Sonnenblick, and C. Payne (1996). "The Total Cost and Measured Performance of Utility-Sponsored Energy Efficiency Programs." *The Energy Journal* 17(1): 31-51.

Eto, J., S. Kito, L. Shown, and R. Sonnenblick (2000). "Where Did the Money Go? The Cost and Measured Performance of the Largest 1992 Commercial Sector DSM Programs." *The Energy Journal* 21(2): 23-49.

Fickett, A.P., G.W. Gellings and A. Lovins (1990). "Efficient Use of Electricity." *Science*. 263: 64-74.

Geller, H. and S. Attali (2005). "The Experience with Energy Efficiency Policies and Programmes in IEA Countries: Learning from the Critics." IEA Information Paper. Paris, France. IEA.

Gellings, C.W., G. Wikler and D. Ghosh (2006). "Assessment of U.S. Electric End-Use Energy Efficiency Potential." *The Electricity Journal* 19(9): 55-69.

Gillingham, K., R. Newell and K. Palmer (2006). "Energy Efficiency policies: A Retrospective Examination." *Annual Review of Environment and Resources* 31: 161-192.

Hansen, Bruce E. (2007). "Econometrics." Mimeo. Department of Economics, University of Wisconsin.

Jordan, J. and S. Nadel (1993). "Industrial Demand-Side Management Programs: What's Happened, What Works, What's Needed." ACEEE Rep. U931. Washington, DC.

Joskow, P.L. and D.B. Marron (1992). "What does a negawatt really cost? Evidence from utility conservation programs." *The Energy Journal* 13(4): 41-74.

Liu, R.Y. (1988). "Bootstrap Procedures under some Non-IID Models" *The Annals of Statistics*. 16(4): 1696-1708.

Loughran, D. and J. Kulick (2004). "Demand side management and energy efficiency in the United States." *The Energy Journal* 25(1): 19-41

MacKinnon, J.G. and H.L. White (1985). "Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties." *Journal of Econometrics* 21: 53-70.

Metcalf, G.E. (2006). "Energy Conservation in the United States: Understanding its Role in Climate Policy." NBER Working Paper No. 12272. June 2006.

Nadel, S. and H. Geller (1996). "Utility DSM: What have we learned? Where are we going?" *Energy Policy* 24(4): 289-302.

Nadel, S. (1992). "Utility demand-side management experience and potential: a critical review." *Annual Review of Energy and the Environment* 17: 507-35.

Nichols, A.L. (1994). "Demand-side management: overcoming market barriers or obscuring real costs?" *Energy Policy* 22(10): 840-47.