



## Opower comments on NMR's second draft interim report for CL&P's home energy reports program

Date: March 23, 2012

### Introduction

Opower, a behavioral energy efficiency and smart grid software company, is the vendor that implemented the home energy reports program with Connecticut Light & Power (CL&P). We appreciate the opportunity to comment on NMR's second draft interim report on this program's impact from February through August of 2011.

Because behavioral programs are relatively new, accurate evaluation is of particular importance. Opower's impact has been independently verified by thirteen separate evaluations.<sup>1</sup> Exemplary of these evaluations, a recent article published in the *Journal of Public Economics* by Dr. Hunt Allcott of MIT evaluated nearly 22 million utility bills from Opower's 17 longest running deployments.<sup>2</sup> Allcott concluded that Opower's program generated electricity and gas savings of 1.4 – 3.3% for all targeted households, with an average of 2%, across all geographies, and that these savings persist over time.

Opower brings the experience of measuring savings from programs with over 60 utilities across 10 million US households to these recommendations.

### Summary of analysis and recommendations

Opower's internal analysis of the program's overall impact is consistent with the report's analysis. Both calculate overall savings rates through May 2011 of approximately 1.5% and through June 2011 of 1.8%. As the primary objective of this evaluation, Opower is very pleased that it has confirmed that the program produced robust, verified energy savings for CL&P's residential customers.

However, some questions still remain about the report's segmentation analysis, which differs significantly from Opower's. For example, the interim report estimates electric heat savings for the monthly treatment group at 5.9%, while Opower's analysis estimates savings for the same group to be approximately 0.8%. If left unresolved, such discrepancies could cause the EEB to draw the inaccurate conclusions about the effectiveness of the program with different market segments. It is important that these differences are resolved in the final report to ensure full confidence in the results and to inform future programs in the state.

Overall, there are four components of the report where more information is needed to reconcile differences between the evaluator’s calculations and Opower’s: use of interaction variables, normalization, control group formation, and details on the methodology used.

Opower recommends that the final report be amended to:

1. Address each of the issues identified by Opower in the segment analyses; and
2. Include a complete and transparent discussion of methodology, including algorithms used, and critical assumptions.

In advance of the report’s completion, Opower would welcome the opportunity to sit down with the NMR team to discuss our respective modeling assumptions and methodologies, and resolve these discrepancies in our analyses.

## **Discussion of analysis and recommendations**

### **The report should address questions on the use of interaction variables, normalization, control group formation, and analytical methodology**

The results of Opower’s internal analysis—using the same data and variables that the evaluator accessed—differ significantly from the report’s results. Both versions of Table 3-2 are below. For example, Opower finds that non-electric rather than all-electric households achieve the majority of savings, which is the opposite of the evaluator’s findings. In combination with the issues discussed in further detail below, such inconsistent results may lead to an incorrect interpretation of the findings and could lead to incorrect program design decisions in the future.

A few notes on the differences in the tables below:

Opower’s table does not include monthly normalization of data, as we question the validity of the methodology used to develop this data (see longer discussion below). Therefore, Opower’s table does not have a monthly savings column, and the average pre-program use column is not monthly

The savings rate and overall savings columns are consistent between both versions of Table 3-2. Opower used the same rate codes that CL&P provided NMR for their analysis of “electric” and “not all-electric” households, so these analyses should be consistent.

Opower’s internal analysis can vary slightly from external analysis. But this variance should be within the confidence interval, and should not be so significant as to lead evaluators to draw conclusions that differ from Opower’s, as is the case with this report.

**NMR Table 3-2: Estimates of Program-induced Electricity Savings, February through August 2011**

Group	Average pre-program use (kWh)	Savings Rate	Number of Households	Monthly Savings (MWh)	Overall Savings (MWh)
Monthly Treatment, not All-electric households	1,598	0.6%	10,046	96	576
Monthly Treatment, All-electric households	1,835	5.9%	4,002	433	2,598
Quarterly Treatment, not All-electric households	1,594	-0.2%	7,148	-23	-138
Quarterly Treatment, All-electric Households	1,845	5.1%	2,842	267	1,602
<b>Total savings</b>	<b>1,665</b>		<b>24,038</b>	<b>774</b>	<b>4,638</b>

**Opower Table 3-2: Estimates of Program-induced Electricity Savings, February through August 2011**

Group	Average pre-program use (kWh)	Savings Rate	Number of Households	Overall Savings (MWh)
Monthly Treatment, not All-electric households	10,484	1.7%	10,146	1,832
Monthly Treatment, All-electric households	12,432	0.8%	3,807	401
Quarterly Treatment, not All-electric households	10,482	1.1%	7,251	831
Quarterly Treatment, All-electric Households	12,373	1.0%	2,713	341
<b>Total savings</b>	<b>11,443</b>		<b>23,917</b>	<b>3,405</b>

This difference is likely a result of issues in either the design or implementation of the report's analytical methodology. Opower has identified four components of the report's segment analyses that would benefit from further explanation: use of interaction variables, normalization, control group formation, and an incomplete methodology. A brief discussion of each will provide greater insight into these issues.

#### **Use of interaction variables**

The report uses interaction variables to attempt to identify the segment of the treatment group that saves the most energy. In this analysis, variables that are interacting with the treatment variable include heat usage type, report frequency, and energy usage tercile.

But methodologically, these variables must be used to stratify the eligible population during the randomization process to avoid specification searching, data mining, and bias in results.<sup>3</sup> Because these variables were not included in this process, any attempt to use them in an *ex-post* analysis can lead to biased results. The report's results are therefore potentially biased.

### **Normalization**

The report discusses a methodology for normalizing all usage data by dividing "monthly use by the average post-period control group consumption and multiplying the results by 100."<sup>4</sup> But the report does not discuss a methodology for normalizing the data to be in the same units. In each independent evaluation of Opower's program to date, the data has been normalized by days.<sup>5</sup>

This is an important step, because days are standard units that do not vary across households. To understand why, consider two households in CL&P's service territory. Household A is on a 20-day billing cycle, while household B is on a 40-day cycle. Assume that A uses 100 kWh in its cycle, and B uses 200. Both these households use 5 kWh per day, but this is not evident unless the data is normalized by day.

There is no evidence that this report performs this normalization, which implies that the analysis could be inaccurately calculating average usage and savings per household.

### **Control group formation**

This report discusses a quasi-matching and quasi-randomization methodology that is used to create control groups for the home energy report frequency analysis. First, the report discusses randomly assigning controls to the monthly and quarterly groups. Then, the report discusses how NMR, "tweaked these households within the sub-control groups until their total average use was similar [to] the use of the corresponding sub-treatment group."<sup>6</sup>

This approach is neither complete propensity score matching nor complete randomization. Rather, it is quasi-matching and quasi-randomization. Randomization is the most appropriate, precise, and unbiased approach to defining control and treatment groups. Quasi-matching or quasi-randomization approaches are less precise and more biased, and should therefore be avoided.<sup>7</sup>

### **Incomplete methodology**

This report does not include a complete explanation of its analytical model, algorithms, or critical assumptions, which makes it difficult to understand some of its findings. The most complete description of methodology is therefore Table 5-3: Population Electricity Savings Rate Estimate-Full Model.<sup>8</sup> As the full model, it would be helpful if Table 5-3 in the final report would include algorithms and an explanation for each variable.

### **The report should include a complete and transparent discussion of methodology, including algorithms used and critical assumptions**

Transparency and completeness are best practice principles that evaluations of this type normally incorporate. Other independent evaluations of Opower's program have included this information, which have provided parties with additional confidence in the findings and

allowed for further validation.<sup>9</sup> In Section 3.8 on Evaluation Characteristics and Ethics, the Department of Energy's National Action Plan for Energy Efficiency discusses completeness and transparency as the first core principle that should define any evaluation process.<sup>10</sup> Among other characteristics, the Department of Energy identifies the following components as critical for following this principle:

"Calculations are well documented in a transparent (clear) manner, with reported levels of uncertainty, in a manner that allows verification by an independent party. The scope of the documentation takes into account the relevant independent variables that determine benefits and the baseline is properly defined."<sup>11</sup>

Consistent with these best practices, the final report should include algorithms, a complete explanation of calculations, or a complete discussion of variables.

## Conclusion

This interim report accurately evaluates the overall impact of home energy reports, and concludes that the CL&P/Opower program leads to measureable energy savings for residential households. However, Opower has remaining questions on four aspects of the report's methodology, which may have led to some inaccurate conclusions about the program's impact on different segments. Opower respectfully requests that the evaluator provide a complete account of the methodologies employed in its analysis and that the issues identified in these comments be addressed in the final report.

Opower would welcome the opportunity to sit down with NMR staff to reconcile differences in our respective analyses and try to better understand the source of any discrepancies to ensure that the results in the final report are meaningful, accurate and beyond dispute.

## Endnotes

---

<sup>1</sup> See the following: (i) Davis, Matt, May 2011. "Behavior and Energy Savings: Evidence from a Series of Experimental Interventions." *Environmental Defense Fund*; (ii) Cooney, Kevin, February 2011. "Evaluation Report: OPOWER SMUD Pilot Year 2." *Navigant Consulting*; (iii) Todd, Annika, Steven Schiller, and Charles Goldman, October 2011.<sup>1</sup> "Analysis of PSE's Pilot Energy Conservation Project: "Home Energy Reports." *Lawrence Berkeley National Laboratory*; (iv) Ivanov, Chris, July 2010. "Measurement and Verification Report of OPOWER Energy Efficiency Pilot Program." *Power System Engineering*; (v) Macke, Rich, June 2010. "Measurement and Verification Report of Lake Country's OPOWER Energy Efficiency Pilot Program." *Power System Engineering*; (vi) Allcott, Hunt and Sendhil Mullainathan, March 2010. "Behavior and Energy Policy." *Science*, Vol. 327; (vii) Allcott, Hunt, February 2010. "Social Norms and Energy Conservation." *Working Paper, Massachusetts Institute of Technology's Center for Energy and Environmental Policy Research*; (viii) Ayres, Ian, et al., September 2009. "Evidence From Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage." *NBER Working Paper*; (ix) Klos, Mary, September 2009. "Impact Evaluation of OPOWER SMUD Pilot Study." *Summit Blue Consulting, LLC*; (x) October 2010. "Puget Sound Energy's Home Energy Reports Program." *KEMA*

<sup>2</sup> Allcott, Hunt, October 2011, "Social Norms and Energy Conservation," *Journal of Public Economics*, available here: <http://web.mit.edu/allcott/www/Allcott%202011%20JPubEc%20-%20Social%20Norms%20and%20Energy%20Conservation.pdf>

<sup>3</sup> For more on appropriate stratification, see, e.g.: Duflo, Esther, Rachel Glennerster, and Michael Kremer, January 2007, "Using Randomization in Development Economics Research: A Toolkit," *Centre for Economic Policy Research*, Discussion Paper No. 6059, available here: <http://www.cepr.org/pubs/new-dps/dplist.asp?dpno=6059.asp>

<sup>4</sup> See Second Draft Interim Report, p. 22

<sup>5</sup> See algorithms from the following evaluations: (i) Davis, Matt, May 2011. "Behavior and Energy Savings: Evidence from a Series of Experimental Interventions." *Environmental Defense Fund*; (ii) Cooney, Kevin, February 2011. "Evaluation Report: OPOWER SMUD Pilot Year 2." *Navigant Consulting*; (iii) Todd, Annika, Steven Schiller, and Charles Goldman, October 2011.<sup>5</sup> "Analysis of PSE's Pilot Energy Conservation Project: "Home Energy Reports." *Lawrence Berkeley National Laboratory*; (iv) Ivanov, Chris, July 2010. "Measurement and Verification Report of OPOWER Energy Efficiency Pilot Program." *Power System Engineering*; (v) Macke, Rich, June 2010. "Measurement and Verification Report of Lake Country's OPOWER Energy Efficiency Pilot Program." *Power System Engineering*; (vi) Allcott, Hunt and Sendhil Mullainathan, March 2010. "Behavior and Energy Policy." *Science*, Vol. 327; (vii) Allcott, Hunt, February 2010. "Social Norms and Energy Conservation." *Working Paper, Massachusetts Institute of Technology's Center for Energy and Environmental Policy Research*; (viii) Ayres, Ian, et al., September 2009. "Evidence From Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage." *NBER Working Paper*; (ix) Klos, Mary, September 2009. "Impact Evaluation of OPOWER SMUD Pilot Study." *Summit Blue Consulting, LLC*; (x) October 2010. "Puget Sound Energy's Home Energy Reports Program." *KEMA*

<sup>6</sup> See Second Draft Interim Report, p. 25

<sup>7</sup> For more on randomization, see, e.g.: Duflo, Esther, Rachel Glennerster, and Michael Kremer, January 2007, "Using Randomization in Development Economics Research: A Toolkit," *Centre for Economic Policy Research*, Discussion Paper No. 6059, available here: <http://www.cepr.org/pubs/new-dps/dplist.asp?dpno=6059.asp>

<sup>8</sup> See Second Draft Interim Report, p. 27

<sup>9</sup> See algorithms from the following evaluations: (i) Davis, Matt, May 2011. "Behavior and Energy Savings: Evidence from a Series of Experimental Interventions." *Environmental Defense Fund*; (ii) Cooney, Kevin, February 2011. "Evaluation Report: OPOWER SMUD Pilot Year 2." *Navigant Consulting*; (iii) Todd, Annika, Steven Schiller, and Charles Goldman, October 2011.<sup>9</sup> "Analysis of PSE's Pilot Energy Conservation Project: "Home Energy Reports." *Lawrence Berkeley National Laboratory*; (iv) Ivanov, Chris, July 2010. "Measurement and Verification Report of OPOWER Energy

---

Efficiency Pilot Program." *Power System Engineering*; (v) Macke, Rich, June 2010. "Measurement and Verification Report of Lake Country's OPOWER Energy Efficiency Pilot Program." *Power System Engineering*; (vi) Allcott, Hunt and Sendhil Mullainathan, March 2010. "Behavior and Energy Policy." *Science*, Vol. 327; (vii) Allcott, Hunt, February 2010. "Social Norms and Energy Conservation." *Working Paper, Massachusetts Institute of Technology's Center for Energy and Environmental Policy Research*; (viii) Ayres, Ian, et al., September 2009. "Evidence From Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage." *NBER Working Paper*; (ix) Klos, Mary, September 2009. "Impact Evaluation of OPOWER SMUD Pilot Study." *Summit Blue Consulting, LLC*; (x) October 2010. "Puget Sound Energy's Home Energy Reports Program." *KEMA*

<sup>10</sup> National Action Plan for Energy Efficiency (2007). *Model Energy Efficiency Program Impact Evaluation Guide*. Prepared by Steven R. Schiller, Schiller Consulting, Inc. Available here: [http://www.epa.gov/cleanenergy/documents/suca/evaluation\\_guide.pdf](http://www.epa.gov/cleanenergy/documents/suca/evaluation_guide.pdf)

<sup>11</sup> *Id.*, p. 3-13