COMMENTS ON EPA'S DRAFT EVALUATION, MEASUREMENT, AND VERIFICATION (EM&V) GUIDANCE FOR DEMAND-SIDE ENERGY EFFICIENCY (EE)

MICHAEL GREENSTONE (UNIVERSITY OF CHICAGO), CHRISTOPHER KNITTEL (MIT), AND CATHERINE WOLFRAM (UNIVERSITY OF CALIFORNIA, BERKELEY)[1]

The Clean Power Plan (CPP) is clear that in order for savings from energy efficiency programs to count towards compliance under rate-based approaches, programs need to be evaluated with energy efficiency (EE) evaluation, measurement, and verification (EM&V) methods. However, we have both theoretical and empirical concerns about the way that these evaluations have been done historically.

On the theoretical side, the CPP creates several conflicts of interest that all provide incentives for evaluations to overstate energy savings. Specifically, the parties involved in ordering, reviewing, and using the energy efficiency evaluations all have a strong incentives for the evaluations to show positive results. The agencies that run the projects (generally, utilities or organizations they contract with) have financial stakes in their program appearing successful; the hired evaluation teams have an incentive to deliver what their clients want to receive – a positive evaluation; and the states who are ultimately responsible for GHG reductions will also have that incentive under the CPP. Together, all three parties have incentives that point to overstating savings from EE programs. These types of incentives problems have been shown to undermine accurate reporting in environmental contexts around the world (e.g., Duflo, Greenstone, Pande, and Ryan 2013). **For the CPP to be successful, it is critical that these incentive problems are recognized and corrected so that the CPP will achieve its desired greenhouse gas reductions.**

On the empirical side, there is an increasing body of evidence that documents that standard EM&V approaches – in particular deemed savings based on engineering estimates – may have delivered claimed savings that were much larger than actual savings (Dubin, Miedama and Chandran 1986; Joskow and Marron 1992; Metcalf and Hassett 1999; Allcott and Greenstone 2012; Davis, Fuchs and Gertler 2014; Fowlie, Greenstone and Wolfram 2015). Further, many of the ex post evaluations are not entirely derived from real world data; a survey by the American Council for an Energy Efficient Economy revealed that 81 percent of programs estimate energy savings with simulation estimates instead of empirical analysis (Kushler, Nowak, and Witte 2012).

Major Comments

In our view, the core threats to the CPP achieving its goals through EE are that the current approaches to EM&V suffer from incentive problems and that the main approaches to measuring savings are unreliable. To rectify this, we suggest that the EPA institute 3 structural requirements for EM&V:

[1] Michael Greenstone, Christopher Knittel, and Catherine Wolfram are the Faculty Directors of the E2e Project, a joint research initiative between the University of Chicago, MIT, and University of California at Berkeley whose mission is to solve the energy efficiency gap.

I. Any evaluation of an energy efficiency program of sufficient size must use real world data that is ideally accessible by others and collected through a randomized controlled trial (RCT) or a quasi-experiment methodology. The evaluations must be designed ex ante and the program data analyzed ex post.

II. The EPA should hire the evaluators using funds from a central pool created with funds collected from the states relying on energy efficiency to meet CPP goals. This approach re-aligns the evaluators' incentives to produce data-driven evaluations.

III. States should only be credited for savings that are produced from studies with these two features: real world data collected through RCTs or quasi-experiment methodologies and analyses by evaluators paid by the central pool. Otherwise, calculated program savings should not count towards compliance.

Specific Comments

1) Expand the definition of randomized controlled trials

The definition of RCTs is too narrow. RCTs are defined on p. 32, section 3.1.3, as an approach that "require[s] that the program can be delivered to certain customers while excluding others from its influence, with random assignment of who receives the program incentive and who does not." It is not always feasible to either mandate or prevent participation in a program and it may be difficult to randomize who receives a treatment, but other rigorous methods can be used to evaluate these situations. For example, randomized encouragement designs can be applied when participants cannot be prevented from accessing a treatment, or implementers can randomize the order of treatment.

2) Expand definition of quasi-experimental approaches

The two quasi-experimental approaches outlined in Appendix C, p. C-11 are a "two-stage regression approach using past and future participants as the comparison" group and a "pooled fixed effects regression approach if there is insufficient data on past or future participants" to form a comparison group. Other rigorous quasi-experimental methods exist, such as difference-in-differences, instrumental variables, and regression discontinuity designs, and these can produce more credible estimates of the impact of a program. Angrist and Pischke (2008) and Gayer and Greenstone (2009) discuss the advantages and limitations of these quasi-experimental approaches. The description of quasi-experimental approaches in the EM&V guidance should thus be expanded to include these types of designs.

3) Remove "Project-Based Measurement and Verification" as a distinct evaluation approach

Three types of EM&V methods are defined as distinct categories on p. 8-9. However, footnote 50 on p. 34 acknowledges that in project-based measurement and verification (PB-MV), "the 'comparison group' is the affected facility before the energy efficiency retrofit" (p. 34). This approach, called pre-post analysis, explicitly defines two groups – facilities before a retrofit and facilities after a retrofit - and thus the PB-MV method is a form of a comparison group approach and should not be categorized as a separate EM&V method. These definitions do not conform to industry standards and should not be used.

References

Allcott, Hunt, and Michael Greenstone (2012). "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives,* 26(1): 3-28.

Angrist, Joshua and Jörn-Steffen Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Davis, Lucas, Alan Fuchs, and Paul Gertler (2014). "Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico." *American Economic Journal: Economic Policy,* 6(4):207-238.

Duflo, Esther, Michael Greenstone, Rohini Pande and Nicholas Ryan (2013). "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Quarterly Journal of Economics*, 128(4):1499-1545.

Dubin, Jeffrey, Allen Miedema, and Ram Chandran. 1986. "Price Effects of Energy-Efficient Technologies: A Study of Residential Demand for Heating and Cooling." *RAND Journal of Economics*, 17 (3): 310–25.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram (2015)."Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." E2e Working Paper 020 (June). Revision requested by *Quarterly Journal of Economics.*

Gayer, Ted and Michael Greenstone (2009). "Quasi-experimental and Experimental Approaches to Environmental Economics," *Journal of Environmental Economics and Management*, 57(1): 21-44.

Joskow, Paul, and Donald Marron (1992). "What Does a Negawatt Really Cost? Evidence from Utility Conservation Programs." *The Energy Journal,* 13(4): 41-74.

Kushler, Martin, Seth Nowak, and Patti Witte (2012). "A National Survey of State Policies and Practices for the Evaluation of Ratepayer-Funded Energy Efficiency Programs." American Council for an Energy-Efficient Economy Report Number U122 (February).

Metcalf, Gilbert, and Kevin Hassett (1999). "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data." *Review of Economics and Statistics*, 81(3): 516-528.