Energy Impact from Gamification-Induced Behavior Change

Ingo Bensch, Evergreen Economics, Madison, WI
Ashleigh Keene, Seventhwave, Madison, WI

ABSTRACT

Behavior programs to influence energy consumption encompass a wide range of interventions. While social norms feedback programs have been studied and evaluated extensively, we know relatively little about the energy impacts of most other behavioral interventions targeting individual and household energy practices. This paper provides results from a behavior change intervention that use gamification. It suggests methodologies for studying the impact of behavioral interventions based on social dynamics and interventions among communities of people where randomized controlled trials are infeasible or impractical.

This paper is based on two assessments of Cool Choices sustainability games—one played by employees of a construction firm and one played by families of school-age children. In both cases, billing analyses showed plausible electricity savings in participating households, albeit with wide uncertainty ranges. While noisy consumption data and small sample sizes hinder precise estimates, we argue that a case can be built over time with a series of billing analyses to demonstrate the savings achieved from the game.

Furthermore, we found that the triangulation of multiple approaches to estimating energy savings increased our confidence in the results and yielded additional actionable insights that helped the program build on its achievements. Post-intervention participant interviews proved particularly insightful and yielded similar results as the billing analysis when used to estimate energy impacts.

Introduction

Energy efficiency programs are shifting their attention from the efficiency of devices to the choices and practices of people who buy and use them, thereby focusing on behaviors. While there have been long-standing efforts to influence end-users' selection of technology and efficiency levels through financial inducements, many behavior-based programs are starting to employ non-financial strategies to affect choices and practices. Examples of these strategies include the use of social norms, feedback, choice architecture, public commitment, goal setting, anchoring, small concessions, reciprocity, and loss aversion (Ashby et al. 2010).

Social science theory suggests that these interventions should be successful in affecting behavior, but the art and science of measuring the energy savings impacts is still evolving. Indeed, behavior change is used in such diverse fields as public health, environmental protection, and education. Energy efficiency programs often need to prove the cost-effectiveness of new approaches, such as behavioral strategies. This requirement necessitates quantification of their energy (and sometimes non-energy) impacts.

The practice of quantifying impacts of behavior programs in the energy efficiency field is still evolving, due in part to the variations in the types of behavioral programs offered and the challenges of isolating their effect from confounding factors. Researchers at Lawrence Berkeley National Laboratory and the State and Local Energy Efficiency Action Network have established a hierarchy among available methods for quantifying energy impacts (Todd et al. 2012). The use of randomized controlled trials (RCTs) is the recommended research method with other experimental or quasi-experimental designs presented as viable, but clearly less desirable, methods.

RCTs have been applied regularly in estimating the impacts of home energy reports provided by Opower and similar program efforts by other vendors (Smith 2014). The impacts of other behavior-oriented interventions have not been studied as extensively or as comprehensively, and it is not clear whether RCTs

can be effectively applied to program interventions that also seemingly offer plausible energy (and non-energy) impacts.

This paper explores one such intervention—the use of gamification by Cool Choices to spur more sustainable actions by participants—and presents alternative ways of estimating such a program's impacts.

Cool Choices: Not Your Standard RCT-Enabled Intervention

Cool Choices is a Wisconsin-based nonprofit dedicated to advancing climate-friendly practices by individuals and households. After researching promising interventions to affect behavior, the organization chose to employ gamification—the use of game elements in a real-life setting—as its primary strategy. Since 2011, Cool Choices has been working with companies, as well as public and nonprofit organizations, to offer a sustainability game that bears the organization's name. The game is offered primarily in a workplace setting,[1] but encourages participants to take sustainable actions in their personal lives that save energy and water at home, reduce gasoline consumption on the road, and advance sustainable practices such as reducing waste (Kuntz, Shukla & Bensch 2012).

Cool Choices' program logic relies on several typical human characteristics—the desire to have fun, the use of social cues and norms, competitive tendencies, responsiveness to intrinsic rewards from doing "the right thing," and the desire to win prizes. The core of the game comprises electronic "cards" that participants can "play" by taking sustainable actions. Each of approximately 100 cards represents a sustainable action—ranging from turning down thermostat settings to reducing one's speed on the highway to shopping with reusable bags. Players report the actions they take during the game—regardless of whether they are pre-existing practices or new ones—and receive points that are roughly commensurate with the relative impact of that action.

Cool Choices intentionally implements its game in work settings and other "communities" where people interact with peers on a regular basis. Games now last several weeks and involve daily reporting in a publicly visible way of qualifying sustainable actions that competitors are taking. Within games, participants form teams and are encouraged to interact with colleagues or peers about sustainability and the actions that they are taking as part of the game (or outside of it). The social dynamics are an important aspect of the theory of the game. In fact, one of the interim metrics by which Cool Choices measures success is increasing the sense within workplace (or other) communities that one's peers care about sustainability and, thus, that it is socially positive to be visibly sustainable. Furthermore, Cool Choices modifies details of its game to conform to the community in which it is played.

These characteristics of the Cool Choices game and the comparatively small sizes of workplace communities that play—numbering hundreds rather than tens of thousands of people—make it difficult to implement RCTs. Holding aside a random sample of community members and isolating them from the kind of social chatter and buzz Cool Choices hopes to achieve is simply not feasible. While one could attempt to randomize at the community level—and thereby overcome sample size limitations as well—there remains the risk that differences between treatment and comparison communities would confound any comparisons.

Cool Choices hired the Energy Center of Wisconsin to determine a feasible way to measure energy savings experienced by its players as a result of their participation in the game. The savings estimates did not need to conform to the precise estimates required for many mandated energy efficiency programs, but the hope was that, over time, estimates could be developed that stand up to a similar level of scrutiny. The authors of this paper both worked for the Energy Center of Wisconsin and were involved in these measurement efforts.

[1] Other settings have included schools and churches.

Energy Savings from Cool Choices Games

Initial energy savings estimates followed a deemed savings approach, whereby each sustainable action was associated with a fixed amount of energy (or other resource) savings. Once Cool Choices had defined the actions that would comprise the card deck, the evaluation team made ex-ante estimates of the likely energy savings inherent in those actions. Those estimates were based on secondary research and assumptions about the initial state of players' households, practices, and efficiency levels. Furthermore, the estimates assumed that players would do exactly what was specified on the card when they claimed the action (i.e., there was no verification of reported actions).

For example, for the card that told players to remove or unplug a secondary refrigerator, we assumed that players who claim this card would take a full-size refrigerator out of service on a permanent basis. We also assumed that the refrigerator would be roughly 20 cubic feet in size and be vintage 1990-92, thereby using roughly 1,285 kWh per year. Similarly, we assumed that players who played the card that asked people to "replace 85 percent of their incandescent bulbs with CFLs" would do just that.

As noted earlier, players receive points for sustainable practices that they had already instituted, as well as new ones that the game spurred. We were able to tell new actions apart from pre-existing ones through a game feature that asks whether the action was newly taken or already in place. Players received the same points either way, so there was no incentive for them to game the system.

These deemed values provided an initial estimate of the savings from the early Cool Choices games, but clearly, a reality check and adjustments to true up the assumptions to actual practices were needed. Similarly, the assumptions needed to be vetted.

We chose to use a multi-pronged approach that centered on billing analyses, but also incorporated post-game interviews, reviews of individual consumption profiles, and comparison groups on a situational basis. The use of multiple empirical data sources provided a way to complement pre/post billing analyses with inherently high uncertainties due to low sample and population sizes. Thus, we were able to obtain reality checks on the billing analysis results, while also providing the program with deeper insights on what comprised the savings that Cool Choices appeared to spur among participants and actionable insights for program design.

Miron Construction: Billing Analysis with Triangulation

Cool Choices offered its inaugural game to permanent staff at Miron Construction in 2011. Miron Construction decided to bring the sustainability game to its headquarters in Neenah, Wisconsin, and remote offices as a complement to the company's focus on sustainable construction. Two hundred and twenty of about 330 Miron employees played the game between May and November 2011 and reported their sustainable actions weekly. Actions qualifying for points in the game were provided on a deck of cards. In all, players reported 3,500 unique actions, of which about half (52 percent) were reported by players as having been newly taken. The remainder of the actions comprised sustainable efforts that players were already taking prior to the game.

The deemed values that we had applied to the 58 pre-defined actions included in the game for Miron Construction suggested average annual savings of 2,100 kWh and 19 therms per active player. We realized that the actual average savings may well be lower because players may implement modified versions of the actions indicated on the Cool Choices cards or have different baseline situations than we had assumed. For example, participants who claimed the card that called on players to switch their furnace fan setting from continuous to automatic may have been running the fan all the time only during selected seasons before the game, and those who played the card that specified replacing 85 percent of incandescent bulbs with CFLs may have just added one or two CFLs.

We sought to identify the actual savings empirically, using both a billing analysis and post-game interviews to understand just what people did when they had played various cards. At the same time, we

conducted other measurement and evaluation activities to provide Cool Choices with feedback on the game's design, implementation, and logic model. Figure 1 illustrates our full efforts. This paper focuses exclusively on the impact estimates, however, which were informed by players' self-reports, the billing analysis, and the post-game interviews.

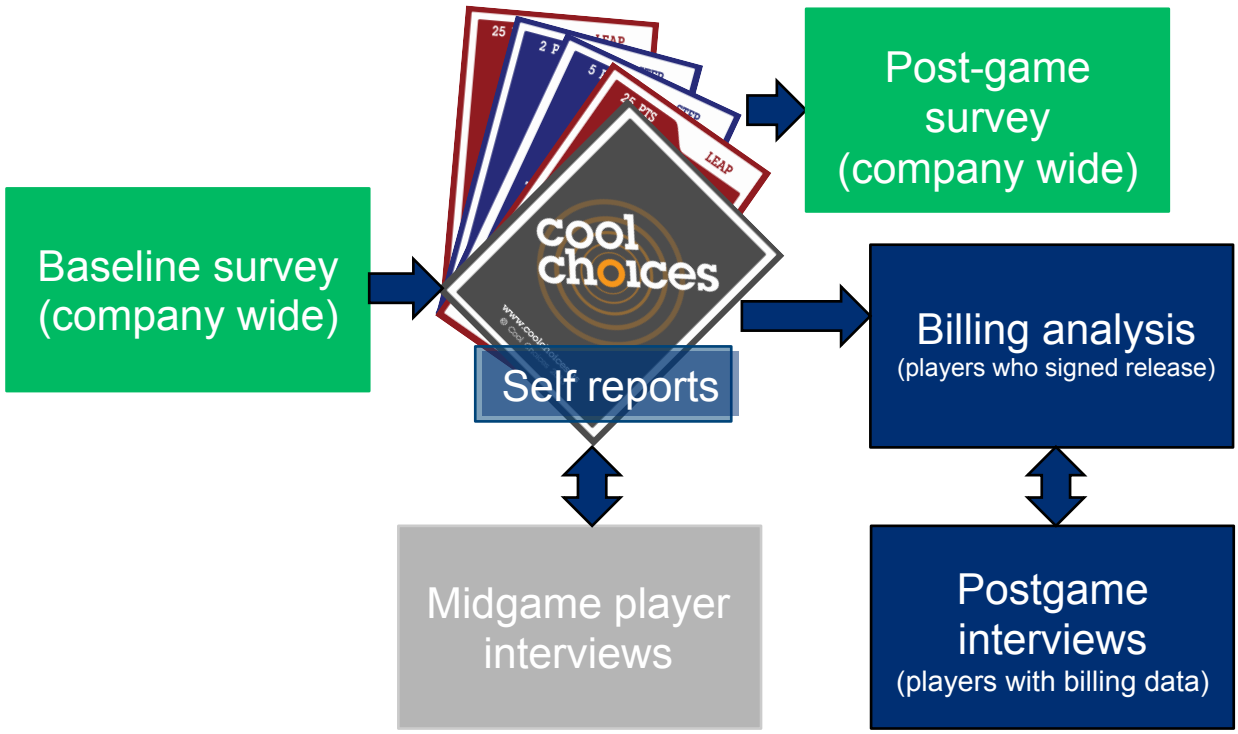

Figure 1. Evaluation activities for Cool Choices game at Miron Construction

Billing Analysis. As part of the game's implementation, Cool Choices had secured billing releases from a subset of players[2] and subsequently obtained actual electric and/or natural gas usage data directly from the applicable utilities for 70 players. The players for whom Cool Choices had billing data resembled the broader population of participants fairly well in the number of cards played (mean of 3.0 electricity-saving actions and 0.7 natural gas-saving actions for the billing analysis group and 2.6 and 0.5 for participants overall), as well as in the total savings implied by the deemed values (2,422 kWh and 24 therms versus 2,100 kWh and 19 therms), allowing us to treat the results of the analysis as representative of Miron Construction participants overall. The billing analysis group—most of whom lived in single-family homes—also resembled typical Wisconsin homes, averaging about 10,000 kWh and 800 therms of consumption annually.

Using these billing data, we compared weather-normalized electricity and natural gas consumption for the 12 months before and after the beginning of the game. The billing analysis suggests annualized

[2] Players were encouraged to sign billing releases, but efforts to obtain these releases were secondary to outreach related directly to the game. Our analysis pool was limited to those players who did sign billing releases.

electrical savings in the 100 to 800 kWh range with a point estimate of 400 kWh. Median savings, robust against the extreme ends of the savings distribution, were a bit higher at 600 kWh, or 6 percent of pre-game consumption. Estimated savings of natural gas were statistically indistinguishable from zero. The uncertainty around the point estimates was wide due to the small sample size and the naturally occurring variation in energy consumption.

As shown in Figure 2, the median electrical savings suggested by the billing analysis was 6 percent of pre-game consumption, albeit with wide variation.

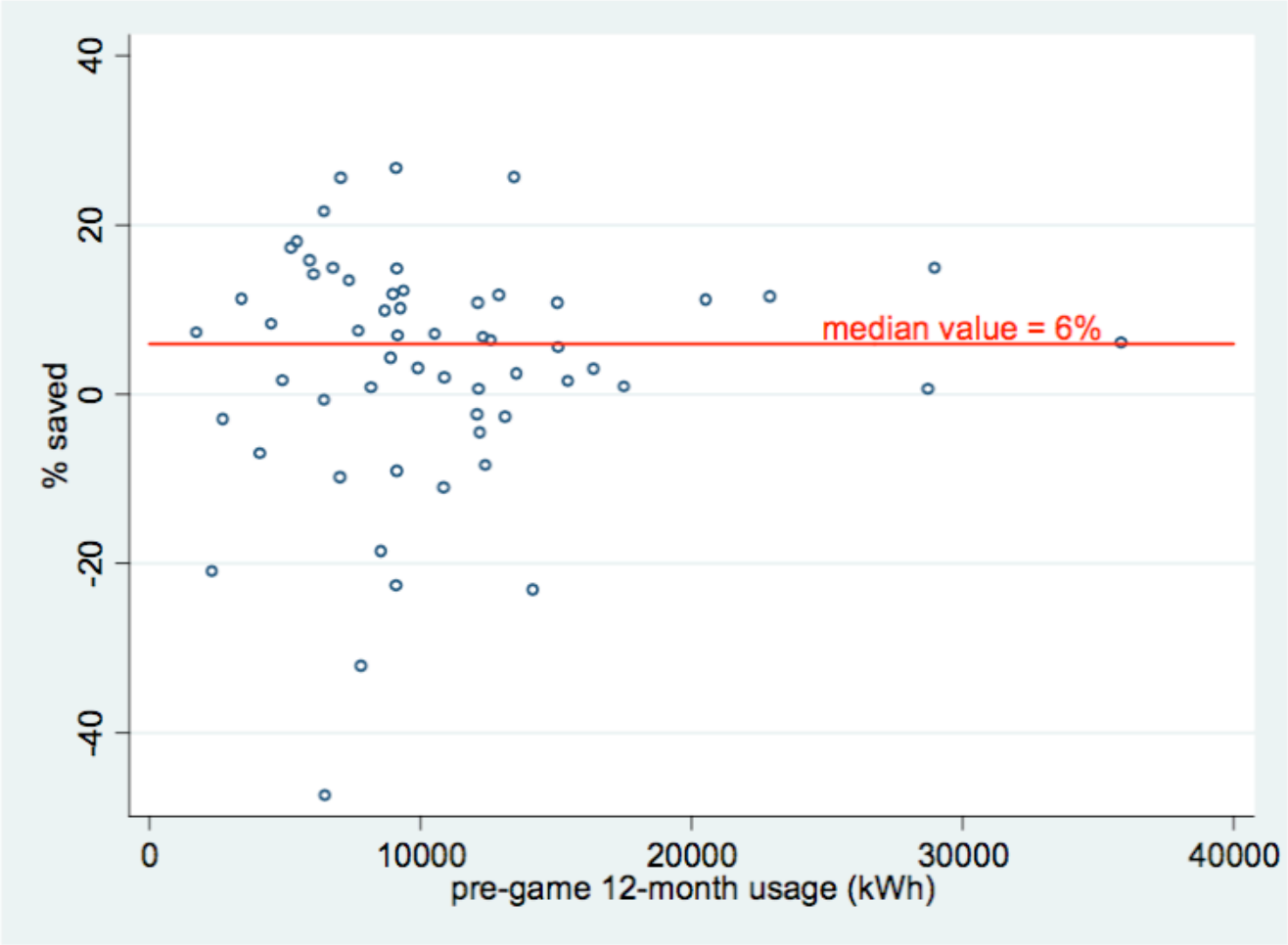

Figure 2. Median electrical savings from billing analysis of Miron Construction players

Post-Game Interviews. We conducted post-game telephone interviews (one year after the game ended) with a subset of 45 players for whom we had billing data. The purpose of the interviews was to solicit more details about the nature of the actions that people took when they claimed cards that we had found to be relatively high impact actions (based on our deemed savings assumptions and estimates). These high impact actions comprised 90 percent of the aggregate initial estimate of electricity savings and 76 percent of the initial estimate of natural gas savings for all players. These actions were:

Replacing 85 percent of incandescent bulbs with CFLs
Removing or unplugging a second refrigerator
Turning off a game console when not in use
Replacing a water heater with a more efficient model
Switching the furnace fan setting from continuous to automatic

Air sealing and insulating to recommended levels
Adjusting the thermostat from 68 to 60 degrees Fahrenheit at night or when no one is at home

In the interviews, we inquired about pre-game practices related to whichever of these actions the interviewee had claimed, changes in practices during the game, and post-game practices at the time of the interview. Where relevant, we also asked about the equipment involved. Other questions included home characteristics and non-game changes in the household (such as changes in the number of people living in the home) that could have affected energy consumption during the analysis period.

Our interviews suggest that players took real actions when they claimed points for various actions. The interviewees had claimed 77 actions related to the high impact activities that we inquired about, and the interview responses suggest that they had taken relevant actions in 65 of these cases. However, what they did varied and deviated from our pre-game assumptions substantially. As shown in Table 1, actions fit the initial assumptions closely in only 26 cases.

Table 1. Comparison of actions claimed and taken (Miron Construction interviewees)

Action	**# who claimed an action***	**# who described a newly taken action**	**# whose action fit initial assumption**
Replace 85 percent of incandescent bulbs with CFLs	13	13	10
Remove or unplug second refrigerator	17	17	7
Turn off game console when not in use	18	12	1
Replace water heater with more efficient model	6	6	0
Switch furnace fan setting from continuous to automatic	15	10	4
Air seal and insulate to recommended levels	8	7	4

* Only includes interviewees who claimed the action to be newly taken when initially reported during the game.

The details that we gathered about what people did allowed us to adjust the assumptions and deemed savings estimates, truing them up to the reality of players' real actions. At the same time, they provided context for a component of the billing analysis in which we examined consumption profiles for signatures that might corroborate or enhance our understanding of the actions taken.

Removal or unplugging of refrigerators serves as an illustrative example of the types of insights that we gathered through the interviews. All 17 interviewees who had claimed the refrigerator action indicated retrospectively that they had either unplugged or removed a refrigerator during the game, and more than four-fifths of these indicated that they had not replaced any refrigerator they removed or permanently plugged in any refrigerator that they unplugged since the game. However, half of the refrigerators (or freezers) the players had unplugged or removed were smaller devices. Furthermore, in some isolated cases, the refrigerators were removed from the interviewee's home and given to someone else to use (and thus not removed from service).

Using these insights from all of the high impact electricity and natural gas-saving actions included in the game, we revised our action-specific assumptions and savings estimates and recalculated whole-game savings based on these trued-up values. Overall, we reduced the savings for the most impactful actions by

more than 50 percent for electricity and even more for natural gas, and we excluded the less impactful actions from the final impact calculation.

Recomputing the savings based on an understanding of what people actually did when they took their sustainability-oriented actions and played the corresponding Cool Choices cards, we estimated savings of 900 kWh and 4 therms per player, as opposed to our initial estimates of 2,100 kWh and 19 therms. Again, the electricity savings were meaningful, while the aggregate natural gas savings were not, although we did hear from a small number of people who took seemingly meaningful steps to reduce their homes' heat loss. The adjusted deemed electricity savings amounted to a median savings of about 9 percent of typical pre-game usage.

Interestingly, our two approaches—the billing analysis and the interview-adjusted savings calculations—yielded similar estimates: medians of 6 and 9 percent in electricity savings and negligible overall natural gas savings. However, comparisons of the *household-specific* savings estimates by our two methods indicate substantial variation, as shown in

Figure 3. The wide range of change in household energy consumption from the pre-game period to the post-game period—as indicated by the billing analysis—highlights the degree to which natural variation over time introduces noise that makes the true savings difficult to disentangle. Presumably, most or all of the "negative savings" shown are due to factors outside the game (such as changes in household size, hours spent at home, and acquisition of new appliances), and there is probably a similar influence from external factors in the positive direction as well.

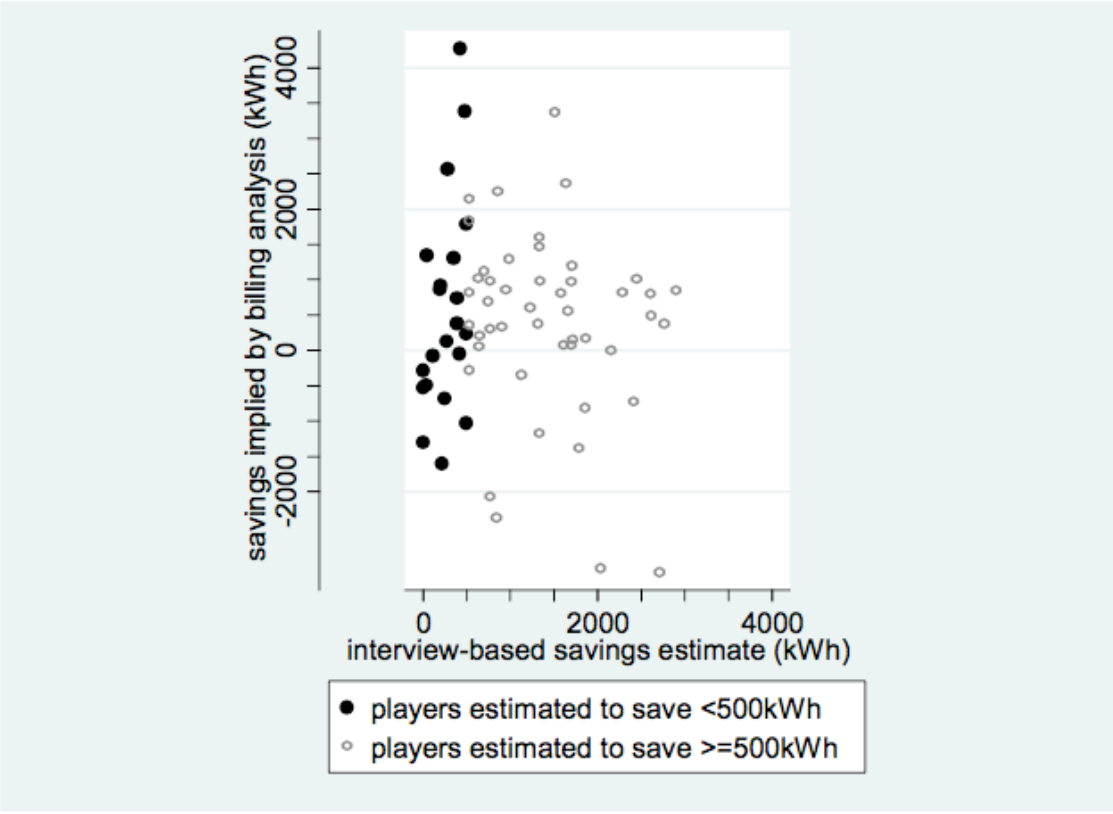

Figure 3. Comparison of usage change (billing analysis) to estimated savings (interview-based)

Because we were analyzing only 70 households, we also examined households' consumption patterns individually to look for usage signatures that could inform us further about the actions that the households took. For impactful measures, such as the removal of a full-sized refrigerator or a permanent change in a furnace fan setting, we thought we might see a corresponding drop in the post-game usage

during summer, winter, and shoulder seasons after the game started (compared to the corresponding pre-game periods).

While interesting, this analysis did not corroborate the billing analysis or interview-based results or otherwise provide clarity. In many cases, there was too much change in usage from varied household activities (associated with the game and external to it) to see the signatures we were expecting. Our comparison of savings expected from self-reported actions to the results of the billing analysis (

Figure 3) already alluded to the differences between expected and observed savings, even though their totals matched reasonably well in the aggregate.

However, we were generally not able to identify the expected signature even when expected and observed savings aligned and usage data appeared to show the drop in consumption.

Figure 4 shows the usage pattern for one such player. This participant's claimed electricity-saving actions were limited to the elimination of a full-size refrigerator. While the billing analysis showed the expected savings associated with the removal of a refrigerator, our analysis of the usage pattern for this player revealed an abnormally big drop in consumption during the winter after the game that probably accounted for a substantial share of the savings identified by the billing analysis. (The refrigerator was removed from the home early in the game, so we would have expected to see a drop in consumption that begins near the first of the red vertical lines that demark the start and end of the game in the left part of the graph.) Several other households with billing analyses that matched expected savings showed similar results—consumption patterns that showed signatures other than the ones we were expecting.

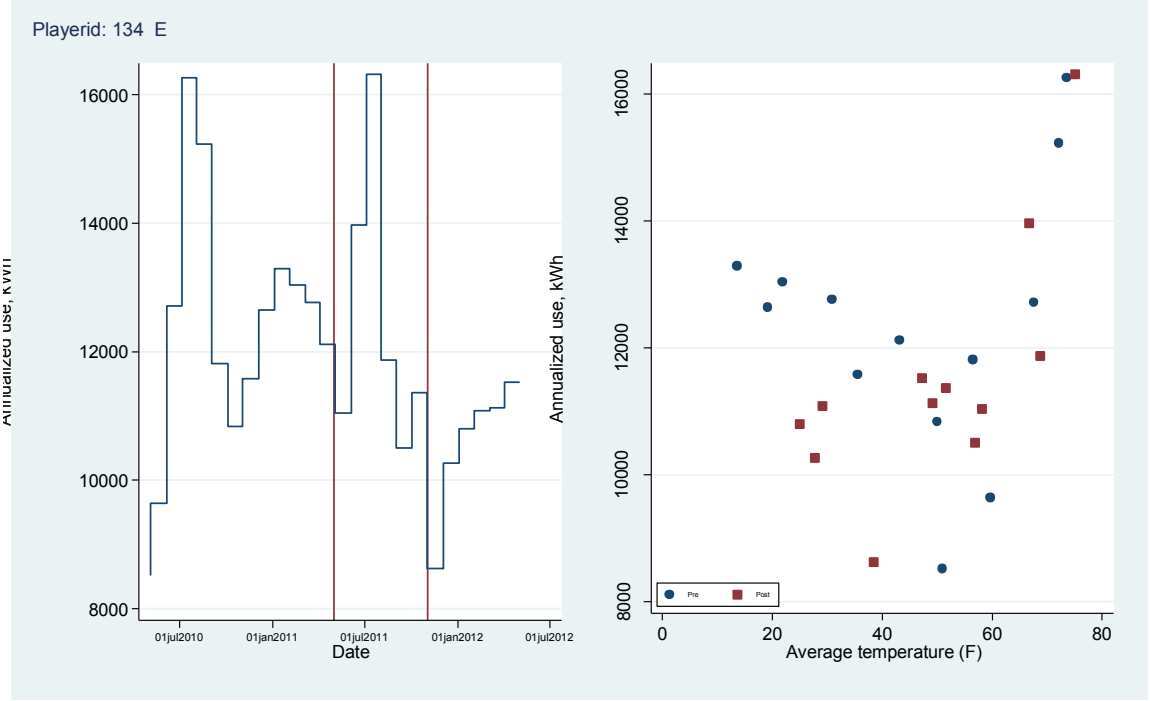

Figure 4: Consumption profile for one player who removed a refrigerator

Sustainable Fox Valley: Billing Analysis with a Comparison Group

Two years later (2013), Cool Choices launched a school-based game in partnership with Sustainable Fox Valley in the cities of Kaukauna and Menasha, Wisconsin. This game resembled the one at Miron Construction, except that it was promoted in several of the local schools and relied on teachers to introduce the game in their classrooms and students to serve as a conduit to their households. As such, there were some differences in the game dynamics that seemed likely to reduce the game's impact. On the other hand,

a partnership with the municipal utilities that serve Kaukauna and Menasha provided the opportunity to obtain usage data for a non-participant comparison group.

Two hundred and twenty-four players participated in the game and reported 4,743 unique Cool Choices actions, 41 percent of them newly taken. The game's duration was shorter than for Miron Construction, spanning from February into May. (Miron players had provided feedback that playing Cool Choices for six months, as originally envisioned, was too long.)

Billing data were available for 103 players representing 75 different households. These households represented the full population of participants served by the two municipal providers for whom we were able to match player data with households. We compared household consumption during the 12 months before and after the game. During this span, 39 households experienced reductions in electricity consumption, while 34 experienced increases. Average annualized consumption for this group decreased by 2.1 percent of pre-game consumption (or 248 kWh), but with a 95 percent confidence interval of -0.9 percent to 5.1 percent. (In other words, these results were statistically not significant at the 95 percent confidence level.)

We used the entire residential populations of Kaukauna and Menasha as comparison groups to assess the degree to which non-game factors influenced electric consumption during the analysis period. As shown in Table 2, residential electric accounts served by Kaukauna Utilities reduced their electric consumption by 0.3 percent, while residential electric customers of Menasha Utilities increased their consumption by 0.3 percent. The table also shows that pre-game consumption by game participants was higher than that of the average Kaukauna and Menasha households, probably because households with school-age children are likely to be above average in household size and more likely to live in single-family homes than the full population of households in these communities.

Because we used a local comparison group, we did not weather-normalize the consumption data. The Cool Choices participants and residents at large would have been exposed to the same weather effects during the analysis period.

Table 2. Electricity usage among participating households and comparison groups: pre-game (February 2012 to January 2013) and post-game (June 2013 to May 2014)

	Number of households	**Pre-game consumption**	**Savings (reduction in consumption)**		**95% confidence interval for % savings**	
		kWh/yr	kWh/yr	Percent	low	high
Participating households	75	11,820	248	2.1%	-0.9%	5.1%
Residential population: Kaukauna	13,106	9,193	-30	-0.3%	n/a	n/a
Residential population: Menasha	8,249	7,385	25	0.3%	n/a	n/a

Conclusions

Some promising behavioral interventions require measurement approaches other than the highly prized RCTs. Program efforts whose program logic relies on social or community-level interaction prevent the sort of experimental isolation required for RCTs, and some behavioral pilots (and non-behavioral pilots) are implemented on a smaller scale than is needed to obtain savings estimates with narrow uncertainty ranges that prove a positive impact. Cool Choices' game-based intervention is one example of a behavioral intervention that requires an alternate measurement strategy.

Billing analyses provided substantial insights about the impact that Cool Choices had on players' in-home energy usage and could conceivably be applied to estimate water savings as well. The advantage of billing analyses is that they are empirical measures of actual changes in consumption. Gasoline savings,

waste reduction, and other environmental impacts achieved by sustainability-focused interventions that go beyond energy require other techniques, however.

Billing analyses of participants alone do not distinguish between program-induced effects and naturally occurring ones, including those induced by economic, demographic, market, and social forces. While an RCT may not be feasible for community-oriented and social interventions like Cool Choices, we found a community-based comparison to be helpful by providing a meaningful comparative context. If external factors affect energy consumption during the analysis period, one would expect the comparison group to reflect those trends and allow for an appropriate adjustment in the savings estimates from the pre/post billing analysis for the participants in the intervention.

Nevertheless, while billing analyses offered insights in our analysis of Cool Choices' impacts, they were not definitive because of the (apparent) single-digit savings percentages achieved and the comparatively large uncertainty ranges associated with natural variations in usage data and game sizes in the hundreds of players. One billing analysis seemed to indicate positive savings in a statistical sense, while the other pointed to possible savings, but was statistically indistinguishable from zero. In such cases, it seems worthwhile to compile impact evidence over time and across games (or pilots). Consistent point estimates in the single digits—regardless of the uncertainty ranges for any one analysis—would point toward a positive effect.

We also found two types of supplemental analyses to be useful. The first was post-game interviews to understand the actual practices and behaviors that may have led to energy savings. Triangulating energy savings estimates based on both billing analyses and a deeper understanding of the behaviors that led to those savings enhances one's confidence that the savings computed are both real and due to the intervention (*if* the two estimates actually point toward similar savings). Furthermore, the details obtained from interviews can provide actionable insights that the program can use for adjustments and improvements. For example, learning that players unplug or remove different types and sizes of refrigerators and freezers and sometimes give them to others for reuse allows Cool Choices to adjust its in-game messaging and the information presented on its action card for second refrigerators. The program can then incorporate information for players that provides savings estimates for smaller units, too, or seek to discourage giveaways of inefficient appliances.

The second supplemental information source was a review of household-specific consumption profiles based on the billing data. In theory, these should have shown usage patterns with signatures that are consistent with the major behavioral changes reported by players. Had we seen such signatures consistently, we would have had even more confidence in a bottoms-up savings calculation (i.e., the interview-adjusted deemed savings calculations). In our analysis, however, the noise in usage data or remaining inaccuracies in our action-specific savings estimates (or both) resulted in inconclusive results. Nevertheless, such deep looks at individual cases can supplement the savings estimates from billing analyses or corroborate action-specific estimates.

Overall, we believe that the impact estimates being applied by Cool Choices offer a path forward for estimating energy (and water) impacts from behavioral interventions that do not lend themselves to RCTs.

References

ADM Associates, Inc. 2014. *Home Performance Program: Evaluation, Measurement, and Verification Report 2013.* Prepared for FirstEnergy Ohio Companies.

Ashby, K.V., M. Nevius, M. Walton, and B. Ceniceros. 2010. "Behaving Ourselves: How Behavior Change Insights Are Being Applied to Energy Efficiency Programs." *In Proceedings of the ACEEE 2010 Summer Study on Energy Efficiency in Buildings,* 7:13-25. Pacific Grove, Calif: American Council for an Energy-Efficient Economy.

Bensch, Ingo. 2013. *Identifying the impacts of Cool Choices' game at Miron Construction: Energy savings from player actions.* Madison, Wisc.: Energy Center of Wisconsin.

Bensch, Ingo. 2014. *Impacts of Cool Choices' Game: Sustainable Fox Valley.* Madison, Wisc.: Energy Center of Wisconsin.

DVN-GL. 2014. *2013 Impact Evaluation. Home Electricity Report Program, January 2012 through December 2013 Study Period*. Prepared for Seattle City Light.

DNV-GL. 2014. *Home Energy Report Program: 2013 Impact Evaluation.* Prepared for Puget Sound Energy.

Kuntz, K., R. Shukla, and I. Bensch. 2012. "How Many Points for That? A Game-Based Approach to Environmental Sustainability." *In Proceedings of the ACEEE 2012 Summer Study on Energy Efficiency in Buildings,* 7:126:137. Pacific Grove, Calif: American Council for an Energy-Efficient Economy.

Smith, B. A. 2014. Business Energy Reports: First Year's Evaluation Results. Presented at the Behavior, Energy and Climate Change Conference, Washington, D.C., October 21.

State and Local Energy Efficiency Action Network. 2012. *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations.* Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. http://behavioranalytics.lbl.gov.