MEASURING THE WELFARE EFFECTS OF RESIDENTIAL ENERGY EFFICIENCY
PROGRAMS

Hunt Allcott
Michael Greenstone

Measuring the Welfare Effects of Residential Energy Efficiency Programs
Hunt Allcott and Michael Greenstone
NBER Working Paper No. 23386
May 2017
JEL No. D12,L94,Q41,Q48

ABSTRACT

This paper sets out a framework to evaluate the welfare impacts of residential energy efficiency programs in the presence of imperfect information, behavioral biases, and externalities, then estimates key parameters using a 100,000-household field experiment. Several results run counter to conventional wisdom: we find no evidence of informational or behavioral failures thought to reduce program participation, there are large unobserved benefits and costs that traditional evaluations miss, and realized energy savings are only 58 percent of predictions. In the context of the model, the two programs we study reduce social welfare by $0.18 per subsidy dollar spent, both because subsidies are not well-calibrated to estimated externality damages and because of self-selection induced by subsidies that attract households whose participation generates low social value. However, the model predicts that perfectly calibrated subsidies would increase welfare by $2.53 per subsidy dollar, revealing the potential of energy efficiency programs.

Hunt Allcott
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10012
and NBER
hunt.allcott@nyu.edu

Michael Greenstone
University of Chicago
Department of Economics
1126 E. 59th Street
Chicago, IL 60637
and NBER
mgreenst@uchicago.edu

Measuring the Welfare Effects of Residential Energy Efficiency Programs

Hunt Allcott and Michael Greenstone[*]

April 28, 2017

Abstract

This paper sets out a framework to evaluate the welfare impacts of residential energy efficiency programs in the presence of imperfect information, behavioral biases, and externalities, then estimates key parameters using a 100,000-household field experiment. Several results run counter to conventional wisdom: we find no evidence of informational or behavioral failures thought to reduce program participation, there are large unobserved benefits and costs that traditional evaluations miss, and realized energy savings are only 58 percent of predictions. In the context of the model, the two programs we study reduce social welfare by $0.18 per subsidy dollar spent, both because subsidies are not well-calibrated to estimated externality damages and because of self-selection induced by subsidies that attract households whose participation generates low social value. However, the model predicts that perfectly calibrated subsidies would increase welfare by $2.53 per subsidy dollar, revealing the potential of energy efficiency programs.

JEL Codes: D12, L94, Q41, Q48.
Keywords: Energy efficiency, program evaluation, randomized control trials, welfare analysis.

In many settings, market failures such as externalities, transaction costs, and imperfect information provide opportunities for socially beneficial policy intervention. In practice, however, it

may be difficult to identify all relevant market failures, or policymakers may be constrained in their ability to implement the first-best policy. In these "second-best" settings, well-intentioned policies may not achieve their full potential, and in some cases may even reduce welfare. This underscores the importance of carefully measuring market failures and estimating how corrective policies affect behavior.

Energy efficiency policy is a natural example of these issues. Energy efficiency decisions are widely believed to be affected by at least two types of market failures. First, most energy use causes negative environmental externalities, so if Pigouvian taxes are unavailable, subsidizing energy conservation could be beneficial. Second, imperfect information, credit constraints, and behavioral biases might cause people to forgo privately beneficial energy efficiency investments.[1] Indeed, McKinsey (2009) argues that adoption of currently available, cost-effective energy efficiency investments in the U.S. could generate $700 billion in net private cost savings. In this sense, energy efficiency closely resembles many other settings where consumers and firms appear reticent to make seemingly beneficial investments, such as better business management, migration to places with higher wages, fertilizer and high-yielding variety seeds, and preventive health care.[2]

Motivated by these potential market failures, the U.S. and other countries have dramatically expanded energy efficiency policy in recent years.[3] Indeed, virtually all credible climate change mitigation plans assign a key role to energy efficiency in reducing greenhouse gas emissions.[4] However, many energy efficiency policies are explicitly second-best: for example, subsidizing energy efficiency instead of directly pricing externalities or providing information to directly address possible information imperfections. This opens the door to unintended distortions. Furthermore, the conventional program evaluation approaches used in practice may not be well-suited to identify these unintended distortions because they rely on accounting-style approaches instead of empirically grounded economic models.

This paper formalizes an approach to modeling home energy efficiency investment decisions and uses this framework to evaluate two large energy efficiency programs in Wisconsin. Like many residential energy efficiency programs nationwide, these programs involved a two-step process. First,

---

[1] These privately beneficial but unadopted investments are often referred to as an "Energy Efficiency Gap," and they have motivated significant interest in policy circles. For example, the Alliance to Save Energy (2013) writes that "energy efficiency is increasingly recognized as the lowest cost, most abundant and cleanest "source" of energy," offering "a win-win solution for economically and environmentally sustainable growth for America." This idea dates at least to the 1970s: Yergin (1979) writes that "If the United States were to make a serious commitment to conservation, it might well consume 30 to 40 percent less energy than it now does, and still enjoy the same or an even higher standard of living ... Overcoming [the barriers] requires a government policy that champions conservation ..."

[2] See Bloom *et al.* (2013), Bryan, Chowdhury, and Mobarak (2014), Duflo, Kremer, and Robinson (2011), Dupas (2014), Dupas and Robinson (2013), Foster and Rosenzweig (2010), Suri (2011), and many others.

[3] For example, the 2009 economic stimulus provided $17 billion for energy efficiency (Allcott and Greenstone 2012). Utility "demand-side management" expenditures grew 80 percent between 2008 and 2013 (CEE 2013, 2015). Twenty-six states now have Energy Efficiency Portfolio Standards, which require utilities to run energy efficiency programs.

[4] For example, the U.S. Energy Information Administration assumes that efficiency will account for 42 percent of CO_2 reductions as part of one of its key mitigation plans (EIA 2015). The International Energy Agency (2015) projects that efficiency measures will reduce global demand growth to 2040 by 40 percent. The Clean Power Plan, the Obama Administration's flagship climate change policy, includes substantial opportunity for compliance through energy efficiency.

homeowners decide whether or not to have a home energy audit, during which they learn about a set of recommended energy efficiency investments, such as improved insulation or new heating systems. Second, they decide which of the recommended investments to undertake. The programs subsidized both audits and investments. Our model captures this two-step process, allowing both observed monetary returns and unobserved benefits and costs to affect takeup. The model accommodates the two key types of market failures introduced above: audit and investment takeup distortions such as imperfect information and behavioral biases, and uninternalized externalities that distort retail energy prices.

The model is estimated using the Wisconsin programs' administrative data, including costs and predicted energy savings for all investments recommended during the audit, paired with a 100,000-household randomized experiment. By varying the content of promotional letters across households, the experiment solves two key identification problems. First, randomly assigned audit subsidies serve as an instrument to identify self-selection—that is, the correlation between unobservables in the audit and investment takeup decisions. Second, randomly assigned informational and behavioral treatments identify the magnitudes of six specific informational and behavioral distortions thought to reduce audit takeup.

Even before quantifying welfare effects, the program evaluation process generates several important empirical results. First, in the randomized experiment, there is no evidence of the hypothesized informational or behavioral failures. Within the letter variations, only price mattered: while a $100 audit subsidy increased takeup by 32 percent relative to control, all six informational and behavioral variations had statistically and economically insignificant effects. This suggests that the informational and behavioral factors we tested are not barriers to takeup, although additional tests would be valuable. If this finding holds beyond the Wisconsin programs, this would undermine the basis for energy efficiency programs to subsidize home energy audits.

Second, in addition to observed monetary benefits and costs, energy efficiency investments entail large unobserved benefits and costs. Non-experimental investment takeup estimates imply that households that had audits were willing to pay an average of $330 for the unobserved attributes of a recommended investment, perhaps due to "warm glow" from contributing to externality reduction or from the improved comfort of a weatherized home. Furthermore, post-audit investment takeup was remarkably inelastic to monetary benefits and costs: consumers did not take up 40 percent of investments with private internal rates of return (IRRs) greater than 20 percent, and they did take up 36 percent of investments with negative private IRRs. This inelasticity implies that consumers perceive a wide dispersion in unobserved benefits and costs. These results highlight the importance of using revealed preference approaches to welfare analyses, instead of conventional accounting approaches that consider only observed monetary factors.

Third, we find strong experimental evidence of self-selection effects in program participation. Specifically, consumers that were marginal to our experimental audit subsidies were much less likely to invest than those that were inframarginal. This highlights one potential problem with policy proposals to expand residential energy efficiency programs by increasing audit subsidies: while

additional households may be induced to audit, these marginal households could be much less interested in making subsequent investments, and thus would generate less externality reduction. Similar self-selection issues arise in many other settings. For example, trading off selection effects vs. the value of experimentation, should we subsidize college enrollment or college completion (Manski 1988, Altonji 1993)? When a job training program expands to enroll more people, are the marginal participants more or less likely to eventually find a job? Our design provides a rare opportunity to identify such selection effects using experimental variation.

Fourth, we estimate that realized energy savings fell well short of predictions. Specifically, the programs' simulation models predicted that the average household that had an audit made investments that would save $153 per year at retail prices, or about 8.5 percent of baseline energy expenditures. In contrast, we estimate an average savings of $89 per year, implying a "realization rate" of 58 percent. The shortfall cannot be explained by temporary weather patterns and is far too large to be caused by a "rebound effect" (i.e. increased utilization in response to the decreased cost of energy services). Identifying potential inaccuracies in simulation predictions is crucial because these predictions are given to consumers during audits (to help decide whether to make investments) and to regulators and policymakers (to help decide whether to continue funding energy efficiency programs). This result, along with a similar result from Fowlie, Greenstone, and Wolfram (2015b), provides important new evidence on the urgency of this issue.[5]

With these results in hand, we turn to welfare evaluation and counterfactual policy analysis. In the model, the optimal investment subsidies would exactly equal the reduction in uninternalized externalities. For example, if an insulation improvement is projected to reduce climate damages and local air pollution by a present value of $500 more than is internalized into retail energy prices, the optimal subsidy would be $500. By contrast, the programs subsidized each investment by a round number multiplier on the share of the household's energy use that would be conserved. This generates several types of distortions, the most important of which results from the fact that different investments reduce natural gas, electricity, and fuel oil in different proportions, and these different fuels have vastly different uninternalized externalities. Therefore, subsidizing energy savings favors investments that reduce energy use but don't necessarily reduce uninternalized externalities.

These distortions may seem subtle, but they turn out to make a very big difference. Using mainstream estimates of the marginal damages from local air pollution and greenhouse gas emissions, the model predicts that a program with perfectly calibrated subsidies would increase welfare by $2.53 per subsidy dollar. By contrast, the model suggests that the program subsidies are so mis-targeted relative to our market failure estimates that they reduce welfare by $0.18 per subsidy

---

[5]This paper's focus and setting are very different from Fowlie, Greenstone, and Wolfram (2015b). They study a different program, the Weatherization Assistance Program, which provides energy efficiency retrofits to low-income households at no cost. By contrast, the programs we study are open to households of all incomes and require participants to pay a meaningful share of costs. Furthermore, Fowlie, Greenstone, and Wolfram (2015b) focus on estimating energy use and "rebound" effects, while the core of our paper is the theoretical framework, demand estimation, and revealed preference welfare analysis. Finally, the field experiments differ: Fowlie, Greenstone, and Wolfram's (2015b) field experiment is designed to generate a first stage for estimating the effect of weatherization on energy use, whereas our field experiment is designed to identify self-selection as well as informational and behavioral market failures.

dollar. The welfare estimates also highlight the consequences of the self-selection effects discussed above: counterfactual increases in audit subsidies generate less and less externality reduction per program dollar, because they induce audits by households that are increasingly unlikely to make energy efficiency investments. Taken together, these results highlights how potentially beneficial public programs can leave large social welfare gains on the table when they rely on heuristic judgments about the form of market failures (e.g. that energy use *per se* is a problem, instead of measuring the uninternalized externalities that vary by fuel) and do not account for empirical factors that govern individual behavior (e.g. self-selection).

As a benchmark, the paper also presents welfare estimates using the conventional "accounting approach." Consistent with the revealed preference approach, these estimates also suggest that the upfront audit and investment costs exceed the present value of reductions in energy, local air pollution, and greenhouse gases. Using the empirical estimates of energy savings, the programs' social internal rate of return is negative 4.1 percent. To help address the question of whether these results generalize outside the two Wisconsin programs, Appendix E presents a parallel analysis using data from 37 Better Buildings program sites nationwide. We find that the national programs had slightly worse IRRs than the Wisconsin programs.

This paper makes two primary contributions to the literature. First, it sets out a theoretical framework to evaluate the welfare consequences of residential energy efficiency programs in the presence of imperfect information, behavioral biases, and externalities, and shows how to estimate the key parameters using a field experiment. This approach is a departure from conventional approaches to evaluating energy efficiency programs that exclusively consider monetary net benefits, and it underscores that standard demand estimation, welfare analysis, and counterfactual simulation can be used in this setting. While researchers have long been interested to include non-monetary factors when evaluating home energy efficiency programs (see Skumatz (2008) for a list of 45 "non-energy benefits" studies), this paper demonstrates how they can be measured through standard revealed preference techniques.

Second, the bulk of the above empirical results run counter to the conventional wisdom among policymakers and practitioners about energy efficiency programs. While McKinsey (2009) and similar studies suggest that energy efficiency programs could generate large private and social benefits, the results imply that at least for the programs we study, this is not the case. This highlights the importance of additional peer-reviewed research and connects to literature in some other domains suggesting that low adoption of apparently beneficial technologies may be due to overestimated private benefits, not market failures.[6] At the same time, the results suggest that it is feasible to design socially desirable energy efficiency programs, but this may require more precise targeting of policies to market failures and more empirical knowledge of the parameters governing consumer behavior.

The paper proceeds as follows. Section I provides an overview of nationwide energy efficiency programs and our case study. Section II presents our theoretical framework. Sections III and IV

---

[6]See Hanna, Duflo, and Greenstone (2016) on cookstoves, for example.

detail the experimental design and data, while Sections V and VI present the empirical strategy and results for audit and investment takeup. Section VII estimates effects on energy use, Section VIII presents the welfare analyses, and Section IX concludes.

I Overview: A Case Study of Nationwide Programs

We focus on two energy efficiency programs in Wisconsin that were part of the national Better Buildings Neighborhood Program (BBNP). We begin by giving an overview of related programs, followed by more detail on the Wisconsin case study.

I.A Overview of Related Energy Efficiency Programs

The programs we study are called "residential energy efficiency retrofit programs": they facilitate and subsidize energy audits and energy efficiency investments, such as improved insulation and heating and cooling systems, at existing homes. Panel A of Table 1 gives an overview of related programs.[7] As shown in column 1, the Better Buildings Neighborhood Program ran from 2010-2013, facilitating approximately 119,000 energy efficiency retrofits, mostly at residential buildings. BBNP allocated $508 million through competitive grants to 41 state and local agencies, including the Wisconsin programs we study. Most of the $508 million funding came through the Energy Efficiency and Conservation Block Grant program shown in column 2. In addition to the stimulus-related programs, there are also longer-running "demand-side management" (DSM) programs shown in column 3, which help utility customers save energy through retrofits and other approaches. Finally, column 4 shows the Weatherization Assistance Program, a retrofit program available only to low-income consumers.

Because these programs are either administered by the government or overseen by regulators, many program evaluation reports have been written: Billingsley *et al.* (2014) identify 4,200 evaluations of DSM programs alone. The standard evaluation uses a straightforward "accounting approach": compare the observed investment costs to the present discounted value of energy savings. Panel B of Table 1 presents some common assumptions that these evaluations make, based on a survey by the American Council for an Energy Efficient Economy (Kushler, Nowak, and Witte 2012). Only 30 percent of programs include benefits other than reduced energy use, and we are not aware of any that measure non-monetary investment costs. Nearly all programs use simulation predictions instead of empirical analysis to estimate energy savings, and most simulation predictions are from states other than the state where the program was implemented. More than four out of five do not use empirical analysis to retroactively evaluate programs. Our paper shows how these assumptions can be relaxed and documents the importance of doing so.

These program evaluation assumptions are particularly relevant given the U.S. Environmental Protection Agency's proposed Clean Power Plan. The proposed plan allows states substantial

---

[7]The public expenditures and energy savings in this table are included only to give a sense of program magnitudes, not as a cost-benefit analysis. Public expenditures do not include any consumer investment costs, and value of predicted energy saved is based on simulation models with varying energy price assumptions.

leeway in how to comply, and it would allow compliance through energy efficiency programs and other mechanisms instead of cap-and-trade programs. Indeed, the National Association of State Energy Officials (2015) believes that "energy efficiency programs ... likely offer the most cost effective means for compliance under the pending EPA rule," and the American Council for an Energy Efficient Economy (2015) has found that "rapidly deployable energy efficiency policies can achieve nearly 70% of EPA's required greenhouse gas emissions by 2030." Compliance through energy efficiency could result in more (less) greenhouse gas abatement than expected under the Clean Power Plan if savings are evaluated through approaches that tend to understate (overstate) energy savings. Similarly, the Clean Power Plan's overall welfare effects would depend on the welfare effects of policies that the states choose to implement.

I.B   The Madison and Milwaukee Programs

We study programs called Green Madison and Milwaukee Energy Efficiency, which were operated jointly but branded separately in each city. The programs were managed by the Wisconsin Energy Conservation Corporation (WECC), a well-respected and highly professionalized program implementer, and they built on the existing design and infrastructure of the Wisconsin Focus on Energy home energy efficiency program. The two programs received part of a $20 million Wisconsin BBNP grant. They were wound down after stimulus funds were exhausted in late 2013, although similar programs continue in Wisconsin and around the country.

From a homeowner's perspective, program participation involved two steps. The first step had three sub-parts. First, a homeowner would schedule a free informational visit by an "Energy Advocate" to explain the program and discuss low-cost conservation opportunities. Second was a home energy audit by an "Energy Consultant," a state-certified independent contractor. During the audit, the Energy Consultant would often put in "direct install measures," primarily compact fluorescent lightbulbs (CFLs) and faucet and shower aerators, at no cost to the homeowner. At the end of the audit, the Energy Consultant would provide an "audit report" with a list of recommended energy efficiency investments, including projected upfront cost, simulation predictions of annual energy cost savings, payback period, and lifetime energy savings for each investment. See Appendix A for an example audit report. Third, homeowners who were interested in making investments would schedule an initial visit by a program-certified contractor to provide a formal cost estimate. In the model below, we think of these three sub-parts collectively as the "audit" step.

The second step was for a contractor to actually perform the work in the consumer's home. In some cases multiple contractors were required for different type of work, for example one for insulation and one for HVAC. After the work was complete, the Energy Consultant would return for a "post-test" to verify that the contractor had done the work properly.

Many residential retrofit programs have a similar structure. While the programs try to make participation as easy as possible, it is clear that both audits and investments require consumers' time and effort as well as money. These time and effort costs represent part of the non-monetary attributes in our model.

7

To predict energy savings for the purposes of aiding homeowner decisions and regulatory evaluations, the Wisconsin programs used an engineering simulation model called the Targeted Retrofit Energy Analysis Tool (TREAT). TREAT is one of several models commonly used by energy efficiency programs nationwide, and 28 percent of audits in the national Better Buildings Neighborhood Program data used TREAT. TREAT has repeatedly satisfied Department of Energy validation protocols, "in which results from software programs are compared to results from other software programs" (PSD 2015a). This suggests that any differences between predicted and empirically estimated savings might not be limited to this software and to the Wisconsin programs.

The programs offered large investment subsidies. The bulk of payments were tiered subsidies of $1000, $1500, and $2000, for a homeowner making investments projected to save 15-24, 25-34, or more than 35 percent of energy use, respectively. There were also less-common subsidies for correcting health and safety issues, doing air infiltration tests, "completion bonuses" for finishing projects before particular dates, and a means-tested subsidy for large retrofits. Appendix Table A.1 presents a breakdown of subsidies paid.

Program participants were also eligible for loans at 4.5 to 5.25 percent interest from a local credit union of $2,500 to $20,000 (up to 100 percent of installation costs), with terms from 3-10 years. While most people in Madison and Milwaukee probably did not know about this opportunity if they did not have an energy audit, the Energy Advocates and Energy Consultants would discuss financing opportunities during the audits, and the audit reports gave financing information in several places. Thus, for people who have had audits, credit constraints should not be a major barrier to takeup.

II Theoretical Model of Home Energy Efficiency Investments

Paralleling the program structure detailed above, we model consumers in a two-step process of audit and investment decisions. We allow for three classes of market failures that motivate energy efficiency policy, as described in overview articles by Allcott and Greenstone (2012), Jaffe and Stavins (1994), and Gillingham, Newell, and Palmer (2009). First, imperfect information or behavioral barriers might distort consumers' decisions about whether to have an audit. Second, similar distortions could affect investment decisions. Third, environmental externalities and other distortions cause an investment's private benefits to differ from its social benefits.

II.A Setup

Heterogeneous consumers indexed by i engage in a two-step process. First, they decide whether to have a home energy audit; we represent this decision with $A_i = \{0, 1\}$. Second, consumers decide whether to make each of a set of potential investments $\mathcal{J}_i$, which are indexed by j; we represent each decision with $I_{ij} = \{0, 1\}$. Consumers cannot invest without having an audit. We assume that investment opportunities are independent in the sense that adopting one does not affect the benefits and costs of adopting another.

8

Audits and investments are provided in perfectly competitive markets at prices c_A and c_{ij}, respectively, where c_A is constant but c_{ij} varies across consumers and potential investments depending on the specifics of the consumer's house. Audits and investments have net non-monetary benefits ξ_{Ai} and ξ_{ij}, which are heterogeneous and could be positive or negative. Costs such as time and hassle during the audit and construction make ξ_{Ai} and ξ_{ij} more negative, while benefits such as a more comfortable home and warm glow from reducing externalities make ξ_{Ai} and ξ_{ij} more positive. In the empirical estimates, ξ_{Ai} and ξ_{ij} are interpreted as demand unobservables, capturing non-monetary benefits as well as all sources of econometric error.

Household energy use is determined by an additional optimization problem that we do not need to model explicitly; see Dubin and McFadden (1984) and Davis (2008). The present discounted value (PDV) of baseline household energy use without the investment is e_{0i}. The investment would reduce energy use per unit of energy services and, unless utilization is fully inelastic to the price of energy services, increase utilization, for a net PDV reduction of e_{ij}. Utilization elasticity (sometimes called the "rebound effect") enters the model as more positive non-monetary benefits ξ_{ij} and lower savings e_{ij}.

The policymaker can set an audit subsidy s_{Ai} and an investment subsidy s_{ij}.[8] We assume that subsidies are funded through a lump-sum tax T, so there is no additional cost of public funds due to a deadweight loss of taxation. A consumer with initial wealth y_i has utility function

$$U_i = y_i - e_{0i} - T + A_i \cdot \left\{ s_{Ai} - c_A + \xi_{Ai} + \sum_{j \in \mathcal{J}_i} I_{ij} \cdot (s_{ij} - c_{ij} + e_{ij} + \xi_{ij}) \right\}. \tag{1}$$

Define N_P as the number of consumers in the population. To maintain a balanced budget, the lump-sum tax must equal total subsidy disbursements:

$$T = \frac{1}{N_P} \sum_{i=1}^{N_P} A_i \cdot \left\{ s_{Ai} + \sum_{j \in \mathcal{J}_i} I_{ij} s_{ij} \right\}. \tag{2}$$

II.B Audit and Investment Decisions

In a model with no market failures, consumers' audit and investment decisions would maximize Equation (1). Our model nests this possibility but also flexibly allows for market failures that might justify audit and investment subsidies.

In the second step, consumers' investment decisions maximize utility in Equation (1), except that there is a reduced-form distortion γ_{ij} that can drive a wedge between utility and investment takeup. For example, γ_{ij} might represent imperfect information that remains even after the audit is complete, or γ_{ij} could be zero if the audit fully informs consumers and there are no other distortions. The investment decision is thus

---

[8]The i subscripts indicate that both types of subsidies can vary across consumers. In the actual Wisconsin programs, the audit subsidy varied across consumers by city and due to our experimentally assigned subsidies, and the investment subsidy varied across consumers and investments depending on predicted energy savings.

$$I_{ij} = 1\left(s_{ij} - c_{ij} + e_{ij} + \xi_{ij} + \gamma_{ij} > 0\right). \tag{3}$$

Define $\lambda_i = \sum_{j \in \mathcal{J}_i} I_{ij} \cdot (s_{ij} - c_{ij} + e_{ij} + \xi_{ij} + \gamma_{ij})$ as the perceived private net benefit from investments that consumer i would make. Before the audit, consumers may be imperfectly informed about this private net benefit, receiving signal $\lambda_i + \gamma_{Ai}$. For example, in a simple rational information acquisition model in which consumers' prior is that they will receive $\mathbb{E}[\lambda]$ from investments, $\gamma_{Ai} = \mathbb{E}[\lambda] - \lambda_i$. The audit decision maximizes utility conditional on the signal of perceived private net investment benefits:

$$A_i = 1\left(s_{Ai} - c_A + \xi_{Ai} + \lambda_i + \gamma_{Ai} > 0\right). \tag{4}$$

More generally, γ_{Ai} could capture any informational or behavioral distortion affecting audit takeup. If γ_{Ai} and γ_{ij} tend to be positive (negative), this makes consumers more (less) likely to audit and invest.

II.C   Social Welfare

We allow the retail energy price to differ from social marginal cost due to uninternalized externalities and other pricing distortions. Household i's baseline energy expenditures are below social cost by a PDV of ϕ_{0i}, and investment j reduces these uninternalized negative externalities by a PDV of ϕ_{ij}. Define $\mathbf{s}$ as the vector of audit and investment subsidies across all consumers, and notice that U_i, A_i, I_{ij}, and T are all implicitly functions of $\mathbf{s}$. Social welfare is the sum over consumers of utility minus the uninternalized externality:

$$W(\mathbf{s}) = \sum_{i=1}^{N_P} \left[U_i - \phi_{0i} + \sum_{j \in \mathcal{J}_i} I_{ij} \cdot \phi_{ij}\right]. \tag{5}$$

The effect of subsidy vector $\mathbf{s}_1$ vs. $\mathbf{s}_0$ on social welfare is

$$\Delta W = W(\mathbf{s}_1) - W(\mathbf{s}_0). \tag{6}$$

The social welfare maximizing subsidies exactly offset the audit and investment takeup distortions: $s_{Ai} = -\gamma_{Ai}$ and $s_{ij} = \phi_{ij} - \gamma_{ij}$. If energy demand is fully inelastic, the equilibrium under those subsidies would be first-best.

II.D   Empirical Approaches to Welfare Analysis

We compare two approaches to measuring the social welfare effect of an energy efficiency program. The "accounting approach" counts the monetary costs and benefits, plus uninternalized externality benefits, from the entire set of investments made at subsidy $\mathbf{s}_1$:

$$\Delta W_a = \sum_{i=1}^{N_P} A_i \cdot \left[(-c_A) + \sum_{j \in \mathcal{J}_i} I_{ij} \cdot (-c_{ij} + e_{ij} + \phi_{ij}) \right]. \tag{7}$$

$\Delta W_a = \Delta W$ under two assumptions: if no investments are made at $\mathbf{s}_0$ and if non-monetary net benefits are mean-zero, i.e. $\mathbb{E}[\xi_{Ai}|A_i = 1] = 0$ and $\mathbb{E}[\xi_{ij}|I_{ij} = 1] = 0$. As discussed above, most energy efficiency programs are evaluated using variants of the accounting approach. This approach is useful because it is not very informationally demanding: ΔW_a can be calculated using administrative data on the monetary costs and benefits of investments, which most programs already record. If empirical data on average energy savings e_{ij} are available, as in our setting, then empirical estimates can be substituted in place of simulation predictions. The two required assumptions may not hold, however. In particular, it would be quite a coincidence for non-monetary benefits to be mean-zero, and we will show that this is not the case in our data. Furthermore, this approach does not allow evaluations of counterfactual subsidy structures.

The "revealed preference approach" involves using observed audit and investment takeup decisions to estimate utility function parameters. It requires the same administrative data as the engineering approach, but introduces two additional identification problems. First, we need to identify the joint distribution of unobservables in the audit and investment takeup decisions, ξ_{Ai} and ξ_{ij}. Put differently, we need exogenous variation in prices or subsidies to identify the slopes of audit and investment demand, as well as the self-selection effects that connect the two demand functions. Second, we need to identify γ_{Ai} and γ_{ij}, the wedges between takeup and utility. The randomized experiment described below helps to solve these two problems.

III   Experimental Design

III.A   Experimental Population and Randomization

We sent promotional letters by direct mail to a subset of households eligible for the Green Madison and Milwaukee Energy Efficiency programs. The experimental population included all owner-occupied single-family homes in Madison and Milwaukee that were built in 1990 or before, had no lien on the property, and had not scheduled an audit prior to June 2012. The population includes 101,881 households, of which 31,213 are in Madison and 70,668 are in Milwaukee. 79,994 households were randomly assigned to receive two identical direct mail promotional letters between June 2012 and February 2013, with the remaining 21,887 assigned to control. We used a max-min t-statistic re-randomization algorithm to ensure balance, and Appendix Table A.2 shows that this was successful.[9]

---

[9]The balancing variables were house age, property value, building area, and the Madison indicator. To ensure unbiased standard errors, we control for the balancing variables when estimating treatment effects (Bruhn and McKenzie 2009).

III.B   Letter Variations

Appendix Figures A.5 and A.6 present example letters. They were printed on 8 1/2-by-11 paper and folded in half for mailing. When opened, the top half was a picture with a short headline. The bottom half includes simple text that describes the program, lays out next steps, and gives a phone number to call to schedule the home energy audit. We varied the letters along seven dimensions, including audit subsidies and six non-price treatments that were designed to address key market failures thought to reduce takeup of home energy audits. These can be roughly categorized into three "informational" market failures and three "behavioral" failures.

III.B.1   Informational Treatments

Appendix Table A.4 details the treatments designed to address informational market failures.

Benefit Information. The Benefit Info treatments provided hard information on the private and social benefits of typical investments that could be made through the program. This was motivated by literature suggesting that imperfect information and biased beliefs could affect energy efficiency investment.[10]

Financing. The Financing treatments informed consumers that low-interest financing was available for investments made through the program. This was motivated by Berry (1984), Gillingham, Newell, and Palmer (2009), and others who propose that credit constraints could reduce energy efficiency investment.

Comparison. The Comparison treatments put the Benefit Information in context by comparing the program's energy savings to other tangible energy use decisions. We compared program non-participation to wasteful actions such as leaving the lights on all day or leaving the door wide open in the winter, in order to make participation seem like the natural choice. These treatments were designed to address the biased beliefs documented by Attari *et al.* (2010), who show that consumers tend to underestimate the savings from large energy efficiency improvements like weatherization relative to small changes like turning off lights. While we have classified this as an "informational" treatment, one could equally classify it as "behavioral."

III.B.2   "Behavioral" Treatments

The top of Appendix Table A.5 details the treatments targeted at potential behavioral failures.

Graphical Prime. We varied the pictures and headlines at the top of the letters to emphasize four different benefits of weatherization: saving money, local and global environmental protection, and a more comfortable home. The psychology literature refers to such graphical variations as "primes": activating an idea, potentially without providing any information, in a way that affects subsequent related behavior (Meyer and Schvaneveldt 1971). Prior research suggests that even subtle graphical primes can be effective. For example, Bertrand *et al.* (2010) find that showing

---

[10]See Allcott (2013), Allcott and Sweeney (2016), Allcott and Taubinsky (2015), Davis and Metcalf (2016), and Newell and Siikamaki (2013) for recent experimental analyses. See Gillingham, Newell, and Palmer (2009), Jaffe and Stavins (1994), Sanstad, Hanemann, and Auffhammer (2006) for overview articles discussing imperfect information.

a female photo increases demand for loans by as much as a two percent reduction in the monthly interest rate.

Time Frame. The Time Frame treatments varied whether the Benefit Information was framed as a one-year or seven-year total. These treatments were motivated by Turrentine and Kurani (2007), who show that consumers have difficulty aggregating savings over time, and Camilleri and Larrick (2014), who find that aggregating savings over longer periods increases stated preference for energy efficiency.

Audit Cue. The Audit Cue treatments varied whether the letter used the phrase "home energy assessment" or "home energy audit" in five different places on the page. Many energy efficiency experts suggest that using the word "audit" can reduce takeup because it cues negative associations with taxes. Program staff asked us to randomize only 1/3 of households into the "audit" condition, because they hypothesized that the word "audit" would reduce takeup.

III.B.3   Subsidy Treatments

The bottom of Appendix Table A.5 details the subsidy treatments.

Subsidy. The Subsidy treatments varied the price of the home energy audit. In the "next steps" box, the letter read: "Call to schedule a home energy [assessment/audit]. Usual cost: $400. You pay only X!" Control group households paid the standard program price, which was X=$200 in Madison and $100 in Milwaukee. Two other groups were randomly assigned to $25 and $100 additional rebates, so their listed prices were X=$175 and $100 in Madison and X=$75 and "nothing" in Milwaukee. A fourth group was presented with the standard control group price, but was also informed that they would receive a $25 Visa cash card after completing the audit. For this group, a mock Visa cash card was included in the letter, in an effort to make the money salient.

The audit subsidy information was relatively subtle, appearing once in normal font near the bottom of the letter. By contrast, the Benefit Information was in bold in a larger font, the word "audit" or "assessment" appeared in five different places, and the Graphical Primes involved the entire top fold of the letter and a headline in very large font. Thus, when we find in Section VI that the subsidy has larger effects than the non-price treatments, it is not because the non-price treatments were more subtly implemented.

IV   Data

Table 2 presents summary statistics. Panel A presents data for the 101,881 households in the Wisconsin experimental population. Of those households, 1.4 percent (1394) had a home energy audit and 0.8 percent (823) made an investment through the programs during the audit and investment sample period, which begins in June 2012 and ends as the programs were wound down in September 2013.

The Wisconsin programs' administrative data include the characteristics of each recommended and adopted investment at every household. Characteristics include investment type (e.g. insula-

tion, air sealing, etc.), unsubsidized cost, and simulation predictions of annual energy savings in physical units of natural gas, electricity, and heating oil per year. Characteristics of adopted investments can differ from the audit report as contractors refine estimates, although on average they are very similar and in many cases identical. The most common types of recommended investments are various kinds of insulation (64 percent of recommendations), air sealing (22 percent), and new heating and cooling systems (11 percent).

Panel B details the two samples of investments that we construct. "Recommended investments" comprise households' choice sets for the investment takeup estimates in Section VI. This is the set of recommendations on the audit report, plus any investments that were adopted but did not appear on the audit report. "Adopted investments" are the investments considered in the "accounting approach" to welfare analysis in Section VIII. These include all subsidized investments.

The costs and predicted savings on the audit reports and in our data assume that investments are independent. For example, a recommended new heating system will have one row in the data with one cost and one predicted savings, with no information on how these might depend on whether the household also installs new insulation. Furthermore, the programs did not retain the data to exactly reconstruct the tiered investment subsidy that a household would receive with vs. without each investment.[11] Thus, since we do not have data on complementarities and substitutabilities, we assume in the model and empirical estimates that investments are independent.

Because our study is limited to evaluating energy efficiency investments, we exclude health and safety projects (improved ventilation and fire risk reduction) and solar photovoltaics from both the "recommended" and "adopted" investment samples. The recommended investments sample additionally excludes observations with zero or negative projected dollar savings (these appear to reflect model input errors), zero-cost and direct install measures (because they are free, there is no plausibly exogenous price variation), appliances (takeup is imperfectly observed), and new hot water heaters (the program treated water heaters inconsistently across households). Our final samples of recommended and adopted investments include averages of 4.4 and 2.8 investments, respectively, per household audited.

We construct present discounted values of energy savings using standard investment lifetimes provided by the program; 95 percent of investments in our final recommendations data have a 20 year assumed lifetime. We assume a five percent annual discount rate, approximately consistent with the real post-World War II returns to the S&P 500 stock market index and with the interest rates on loans available to program participants.

We calculate energy prices to reflect averages over 2011-2014. In Madison and Milwaukee, natural gas and electricity are sold by regulated local monopolies, while heating oil is sold by multiple competing providers. We gathered retail marginal prices for natural gas and electricity from the Madison and Milwaukee utilities, and we use the Wisconsin average residential heating oil price from the Energy Information Administration (EIA). At retail prices, 76, 7.7, and 16 percent

---

[11]In particular, the program did not retain the household-specific baseline energy use estimates used to determine investment subsidy amounts.

of savings from adopted investments are from natural gas, electricity, and heating oil, respectively. For natural gas acquisition costs, we use Wisconsin wholesale ("citygate") prices from EIA. For electricity acquisition costs, we use "all-in" wholesale market prices for the MISO market, which includes Wisconsin, from Potomac Economics (2011-2014). These "all-in" electricity prices include quantity-weighted average costs for energy and capacity, plus ancillary services and uplift charges. For heating oil, we assume that retail price equals marginal cost.

Panel C presents summary statistics for the electricity and gas usage microdata. Wisconsin law prohibits utilities from sharing energy use data with researchers unless the customer consents. Customers were asked to sign release forms during the audits, and 90 percent (1258 out of 1394) agreed, but we do not have energy use data for the larger group of unaudited households in the experimental population. We drop households that installed solar photovoltaics or are recorded as having participated in another energy efficiency program for which we do not observe predicted savings; both of these factors would bias the comparison of empirically estimated savings to predicted savings from investments made through these programs. The energy use data sample begins as early as January 2006 and ends in May 2015. We do not have consistent heating oil consumption data, but only 23 households made investments that were predicted to save heating oil.

Appendix B presents additional information on data preparation, categories of investments and subsidies, and energy price and externality assumptions.

V   Empirical Strategy

We first specify the empirical analogue to Equation (4), the audit takeup equation. S_{Ei} is household i's experimental audit subsidy (either \$0, \$25, or \$100), G_i is an indicator for the \$25 gift card offer, $\mathbf{T}_i$ is the vector of indicators for informational and behavioral treatment groups, and $\mathbf{X}_i$ is a vector of the five household-level covariates from Panel A of Table 2: house age, property value, building footprint, a Madison indicator variable, and Census tract hybrid vehicle share. In the context of the model, the treatments $\mathbf{T}_i$ can affect γ_{Ai}, and household characteristics $\mathbf{X}_i$ can be associated with non-monetary preferences ξ_{Ai}, potential benefit from investing λ_i, and γ_{Ai}. Define $V_{Ai} = S_{Ei} + \varphi G_i + \tau \mathbf{T}_i + \beta_A \mathbf{X}_i + \kappa_A$ as the observed part of latent utility from auditing, where κ_A is a constant, and define ϵ_{Ai} as an econometric error. In the model, utility is money-metric, so V_{Ai} and its component parts are in units of dollars.

The empirical analogue to Equation (4) is

$$A_i = 1\left(V_{Ai} + \epsilon_{Ai} > 0\right). \tag{8}$$

We also specify an empirical analogue to Equation (3), the investment takeup equation. C_{ij} is the investment cost estimate, and E_{ij} is predicted retail energy cost savings over the assumed investment lifetime using a five percent annual discount rate. Recall that the bulk of investment subsidies were from tiered subsidies that increased by \$500 for every 10 percentage point increase in energy saved by adopted investments relative to the household's baseline. Because we do not

have the data to exactly reconstruct the tiered program subsidies, we impute for each investment a linearized subsidy with a cap at \$3500: $S_{ij} = \min\left\{\$5000\frac{E_{ij}}{E_{0i}}, \$3500\right\}$, where E_{0i} is the PDV of household i's average pre-audit energy use, discounted over the investment lifetime. Let $\xi_{ij} = \beta_I \mathbf{X}_i + \xi_j + \epsilon_{ij}$, where ξ_j is a constant and ϵ_{ij} is an econometric error, and define $V_{ij} = S_{ij} - C_{ij} + E_{ij} + \beta_I \mathbf{X}_i + \xi_j$ as the observed part of latent investment utility. The empirical analogue to Equation (3) is

$$I_{ij} = 1\left(V_{ij} + \epsilon_{ij} > 0\right). \tag{9}$$

.

In the theoretical model from Section II, ξ_{ij} represented non-monetary attributes such as comfort benefits and time costs. Empirically, these are interpreted as unobserved attributes, capturing non-monetary attributes, un-modeled monetary attributes, and any other econometric errors, including discount rates other than five percent and idiosyncratic monetary factors that are known to the consumer but unobserved on the audit reports. If consumers believe that realized energy savings will be lower (higher) than predictions presented on the audit report, this enters as lower (higher) estimated ξ_{ij}.

We assume that $\eta_A \epsilon_{Ai}$ and $\eta_I \epsilon_{ij}$ are distributed standard normal, where η_A and η_I are scaling factors. When estimating Equation (9) and Equation (10) below, standard errors are clustered by household to allow for arbitrary within-household correlation in ϵ_{ij}. Notice that because the *rescaled* errors are distributed standard normal, the estimated probit coefficients are the scaling factor times the coefficient written above. For example, the estimated coefficient on $\mathbf{T}_i$ in the audit takeup equation will be $\eta_A \tau$, and the estimated coefficient on S_{Ei} will be η_A. Dividing the estimated coefficients on $\mathbf{T}_i$ by the estimated coefficient on S_{Ei} will give $\eta_A \tau / \eta_A = \tau$, in units of dollars. Analogously, the estimated constant term in the investment takeup equation is $\eta_I \xi_j$, so dividing by the price coefficient will give ξ_j, in units of dollars.

If $\epsilon_{Ai} \perp \epsilon_{ij}$, then probit estimates of Equation (9) are relevant for the full 101,881-household population. Otherwise, probit estimates of Equation (9) are relevant only to the selected sample of households that had audits. Our primary specification allows for correlation between ϵ_{Ai} and ϵ_{ij} by jointly estimating the audit and investment takeup equations using the maximum likelihood approach of Van de Ven and Van Praag (1981). Defining $\rho = corr(\epsilon_{Ai}, \epsilon_{ij})$ and $\Phi_2(x, y, \rho)$ as the bivariate standard normal cumulative distribution of x and y with correlation ρ, the log-likelihood function is

$$\ln \mathcal{L} = \sum_i \left\{ \begin{array}{c} A_i \sum_{j \in \mathcal{J}_i} \left[I_{ij} \ln \Phi_2(V_{ij}\eta_I, V_{Ai}\eta_A, \rho) + (1 - I_{ij}) \ln \Phi_2(-V_{ij}\eta_I, V_{Ai}\eta_A, -\rho)\right] \\ + (1 - A_i)w \ln\left[1 - \Phi(V_{Ai}\eta_A)\right] \end{array} \right\}. \tag{10}$$

Inside the brackets, the first line sums over all investment opportunities for the households that did audit, while the second line sums over the households that did not audit. One feature of our

data is that there are multiple investment decisions for each audit, which causes the households that audit to appear multiple times and thus receive more weight in estimating the audit takeup coefficients. To identify the audit takeup coefficients equally from households that did vs. did not audit, we weight the non-audited households with weight w equal to the average number of recommended investments per audit, while weighting each investment observation with weight 1.

Equation (10) delivers the parameters necessary for the revealed preference welfare analysis. We can now see how our RCT helps solve the two identification problems discussed earlier. First, this is a sample selection model in the spirit of Heckman (1979) and the literature that follows: we observe characteristics and takeup decisions for recommended investments only if a household audits. The randomly assigned subsidies affect audit takeup but do not change investment incentives, which means that they act as an excluded instrument to identify the correlation between ϵ_{Ai} and ϵ_{ij}. In the literature that uses sample selection models, it is rare to have such an instrument. However, we do not have an instrument for price in the investment equation, so we need to assume that $(S_{ij} - C_{ij} + E_{ij}) \perp \epsilon_{ij} | \mathbf{X}_i$, i.e. that monetary characteristics are uncorrelated with unobservables affecting investment takeup.[12]

Second, the RCT helps to identify the audit takeup distortion γ_{Ai}. Intuitively, distortions can be measured in dollar terms by dividing the effect of removing the distortion by the effect of a price change. For example, if all consumers have $\gamma_{Ai} = -\$25$ due to imperfect information, providing full information will have the same effect on audit takeup as subsidizing audits by \$25. If instead $\gamma_{Ai} = -\$50$, providing full information will have twice the effect of the \$25 subsidy. Thus, if $\eta_A \tau$ is the effect of a treatment that fully removes an informational or behavioral distortion and η_A is the price effect, the dollar value of the audit takeup distortion is $\gamma_{Ai} = \eta_A \tau / \eta_A = \tau$.[13] Thus, our six informational and behavioral treatments test for particular sources of γ_{Ai}, although we cannot rule out the possibility of other audit takeup distortions.

VI   Empirical Results

We begin with descriptive results on the field experiment, investment takeup, and self-selection into audits. We then present formal estimates of Equation (10).

---

[12]The unobservable ϵ_{ij} would be positively correlated with E_{ij} if more energy savings bring more warm glow utility or are associated with more in-home comfort. Furthermore, ϵ_{ij} would be negatively correlated with C_{ij} if higher-cost projects also require more non-monetary effort to implement, e.g. if larger home construction jobs are both more costly and more of a hassle for the homeowner. The likely directions of these possible correlations would bias $\hat{\eta}_I$ upward, which would only reinforce our empirical finding that investment takeup is highly inelastic. We explore these issues in Appendix C.B, although we find little evidence of bias in $\hat{\eta}_I$.

[13]More precisely, Mullainathan, Schwartzstein, and Congdon (2012) show that if γ_{Ai} is homogeneous, a first-order approximation to γ_{Ai} is the ratio of the information effect to the price effect. Allcott and Taubinsky (2015), Chetty, Looney, and Kroft (2009), and Bronnenberg *et al.* (2015) use variants of this approach to identify informational and behavioral distortions in other markets.

VI.A Effects of Randomized Variations in Promotional Letters

VI.A.1 Effects of Letter and Subsidy Treatments

Column 1 of Table 3 presents probit estimates of Equation (8), the effects of the experiment on audit takeup. We first present estimates replacing $\mathbf{T}_i$ with an indicator T_i for being mailed a letter. We present marginal effects, with coefficients multiplied by 100 for readability.

Receiving a letter with zero monetary incentives increased the probability of auditing by 0.158 percentage points, or about 13 percent of the control group mean. This relatively small effect that is at the margin of conventional significance levels suggests that basic lack of awareness of the program was not a major barrier to takeup. The gift card had no effect, perhaps because of perceived transaction costs in activation.

Money does matter, however: a $100 subsidy increases the probability of an audit by 0.525 percentage points, or 32 percent of the control group takeup.[14] Even after heavy subsidies, demand for audits is remarkably low: households in the $100 subsidy group in Milwaukee (Madison) needed to pay a net-of-subsidy price of only $0 ($100) for an audit, compared to a typical market price of $400. Despite this, only 1.8 (2.2) percent of households in Milwaukee (Madison) in the $100 subsidy group had audits.

Column 2 presents analogous probit estimates of whether household made any investment. The point estimates suggest that a relatively small share of consumers that were marginal to the letters and experimental subsidies eventually invested. In the 21,887-household control group that did not receive informational letters, 64 percent of households that audited followed through with some investment. If consumers marginal to the treatments followed through at the same rate, then the ratio of estimates in column 2 to column 1 would also be 0.64. By contrast, the point estimates suggest that the average letter without experimental subsidy increased investments by about 26 percent of the increase in audits (0.041/0.158), and the experimental subsidies increased investments by 29 percent of the increase in audits (0.152/0.525). The fact that the marginal auditors are less likely to invest implies that ϵ_{Ai} is positively correlated with observable or unobservable investment attributes. We further explore this issue below.

The $\mathbf{X}$ covariates are associated with audit and investment takeup in intuitive ways. Takeup increases in house age: because building codes and construction techniques have improved and houses' efficiency depreciates over time, older houses can benefit more from weatherization retrofits. Takeup is positively correlated with hybrid vehicle share, perhaps because environmentalists benefit more due to warm glow. Takeup is also positively correlated with wealth, as measured by property value and building footprint.

---

[14]Additional estimates of both audit and investment takeup show that the effect of the $100 subsidy is not statistically distinguishable from four times the effect of the $25 subsidy.

VI.A.2 Effects of Informational and "Behavioral" Treatments

We also estimate a version of Equation (8) with the full set of $\mathbf{T}_i$ indicators; the full set of estimates are reported in Appendix Table A.7. In summary, the only treatments that affect takeup are the monetary incentives. Furthermore, Wald tests in Appendix Table A.8 show that none of the six groups of informational or behavioral treatments jointly affected audit or investment takeup.

How precisely estimated are these zero effects? Figure 1 presents the point estimates and confidence intervals for each informational and "behavioral" treatment scaled by the effect of a \$1 subsidy, i.e. $\eta_A \tau / \eta_A = \tau$. This translates the coefficient estimates into units of dollars, the same units as utility and γ_{Ai}. All 90 percent confidence intervals include zero, and the average confidence interval bounds the effect at no more than the effect of a \$30 to \$40 price change.[15]

This normalization of treatment effects into dollar terms is useful for two reasons. First, it provides an estimate of the monetary magnitude of any informational distortions under the assumption that $\gamma_{Ai} = \tau$, as discussed above. In our experiment, one should not interpret any given treatment as removing all distortions, both because it is unlikely that our treatments fully addressed the hypothesized distortions and because there are other possible distortions that we could not test. Notwithstanding, the bounds on the information effects suggest that the magnitudes of the informational and behavioral audit takeup distortions that motivated our six treatments are at least an order of magnitude smaller than the programs' \$200-\$300 audit subsidies.

Second, the normalization addresses the fact that many people don't read unsolicited mail. Imagine that share $r < 1$ of consumers who were mailed the letters actually read them, and the true treatment effect on letter readers is τ', so $\tau = r\tau'$. Then τ could be small either because the treatments had little effect on the letter readers (τ' is small) or because few people read the letters (r is small). Taking the ratio of the information effect to the subsidy effect divides out the r, giving the ratio of effects within the group of letter readers assuming that people read the entire letter. This issue is crucial to interpreting our results, as these economically small coefficient *ratios* cannot be explained by people not reading the letters: if nobody read the letters, then the experimental subsidies would also have no effect. This discussion also clarifies that all coefficients and ratios are "local" to the subset of people who read the letters, and these people could in be systematically more or less informed or "behavioral" than the people who do not read the letters.

VI.B Descriptive Evidence on Investment Takeup

Figure 2 illustrates the identification of investment takeup in Equation (9). The vertical bars are a histogram of predicted net monetary benefit $(S_{ij} - C_{ij} + E_{ij})$ in the sample of recommended invest-

---

[15]For comparison, the three statistically significant advertising treatments in Bertrand *et al.* (2010) affected demand by the equivalent of a two percent change in monthly interest rate. At a median loan size of \$150, a two-percent interest rate change is worth \$3 per month, or \$12 total for their four-month loans. This is about 2.5 percent of their population's \$470 median gross monthly income. By contrast, our population's median gross monthly income is \$4000, so 2.5 percent of median income is \$100. Thus, after accounting for population income differences, we can bound the price-scaled effects of all our non-subsidy treatments at about 30-40 percent as large as the Bertrand *et al.* (2010) estimates.

ments. We truncate the graph at $\pm\$3000$ for readability, plotting all smaller (larger) investments in the far left (right) bins. The dots illustrate the takeup rates within each bin.

Figure 2 has three striking features. First, many of the recommended investments are disadvantageous from a purely financial perspective, even at subsidized investment costs and marked-up retail energy prices and assuming that the simulated savings and 20-year lifetimes are correct. At a five percent discount rate, one quarter of recommended investments lose $\$638$ or more, the median recommendation loses $\$47$, and 53 percent do not pay back. Second, takeup rates are increasing in monetary net benefit, meaning that consumers clearly consider monetary incentives. Third, while takeup clearly increases in projected monetary benefit, the slope is quite gradual: a $\$1000$ increase in projected monetary benefit is associated with only about a five percentage point increase in takeup. Consumers did not take up 40 percent of investments with private internal rates of return (IRRs) greater than 20 percent, and they did take up 36 percent of investments with negative private IRRs. This inelastic demand implies that the estimated slope of takeup with respect to monetary benefit, $\hat{\eta}_I$, will be small: there must be wide dispersion in unobserved attributes to rationalize these takeup decisions. This finding of large unobserved benefits or costs highlights the importance of incorporating unobserved attributes into welfare analysis using the "revealed preference approach" instead of the "accounting approach." The inelasticity also implies that in Section VIII's analyses of counterfactual policies, increases in investment subsidies will generate only limited increases in investment takeup.

VI.C   Evidence of Self-Selection into Audits

Figure 3 illustrates the effects of self-selection into audits. The figure presents takeup decisions for households that received the marketing letter and were offered a $\$0$, $\$25$, or $\$100$ experimental audit subsidy; the $\$25$ gift card group is excluded. The left panel shows the audit probability as a function of the audit subsidy. The $\$100$ experimental subsidy increases audit takeup from 1.3 to 1.9 percent. The light (gray) bars on the right panel show the average investment probability by subsidy group, conditional on auditing. The $\$100$ subsidy group is about 13 percentage points less likely to invest.

Using the average investment probabilities and the share of marginal vs. inframarginal consumers implied by audit takeup in the left panel, we back out the investment probability for households marginal to each subsidy increase.[16] Even before calculating the exact numbers, we know that the marginal auditors must have markedly lower investment probabilities for higher subsidies to cause such a large decrease in average investment probability. Indeed, the dark (red) bars in

---

[16]Specifically, conditional on auditing at subsidy $S_A = s$,

$$\Pr(I = 1) = \Pr(I = 1|M = 1) \cdot \Pr(M = 1) + \Pr(I = 1|M = 0) \cdot (1 - \Pr(M = 1)), \tag{11}$$

where $M = 1$ is an indicator for being marginal to an audit subsidy increase from s_0 to s: $M = 1\left(-s < \tau\mathbf{T}_i + \beta_A\mathbf{X}_i + \kappa + \epsilon_{Ai} < -s_0\right)$. Re-arranging gives an equation for $\Pr(I = 1|M = 1)$, the investment probability for marginal consumers. $\Pr(M = 1)$ is from audit takeup rates illustrated in the left panel, and $\Pr(I = 1)$ and $\Pr(I = 1|M = 0)$ are from investment takeup rates in the light bars in the right panel.

the right panel show that marginal investment probabilities drop sharply as the subsidy increases. While 54 percent of households that audit at zero experimental subsidy make some investment, the point estimates imply that only 25 percent of households that are marginal to the \$25 subsidy do so. Remarkably, the point estimates imply that among the households that audit at a \$100 subsidy but do not audit at a \$25 subsidy, only three percent make an investment.

One policy implication is that programs interested in maximizing investments per subsidy dollar should think carefully about what to subsidize. In this setting, subsidizing audits draws in proverbial "tire kickers": households that are interested in audits, perhaps because they have a low value of time or want free CFLs, but are not interested in making energy efficiency investments. This strong self-selection will play a key role in Section VIII's counterfactual analysis of higher audit subsidies.

VI.D   Joint Estimates of Audit and Investment Takeup

The counterfactual policy analyses presented below require demand parameters relevant for the full population. The above finding that the marginal auditors are much less likely to invest makes clear that a joint model of audit and investment takeup is necessary to estimate parameters for the full population, rather than investment parameters specific to the small subset of households that audit.

Table 4 compares independent and joint estimates of the audit and investment takeup equations. Columns 1 and 2 present the independent probit estimates of Equations (8) and (9), while columns 3 and 4 present estimates of Equation (10). Columns 1 and 3 present stripped-down estimates, excluding household covariates $\mathbf{X}_i$, while columns 2 and 4 include $\mathbf{X}_i$ and replace the constant ξ_j with separate indicators for all six investment categories: air sealing, insulation, heating/cooling systems, windows, pipe and duct sealing and insulation, and programmable thermostats. The top panel of column 2 is the same as column 1 of Table 3, except that we now present probit coefficient estimates, not marginal effects.

The experimental audit subsidy S_{Ei}, gift card indicator G_i, and letter treatment indicator T_i are included in the audit takeup equation but excluded from the investment takeup equation, thus identifying the correlation between ϵ_{Ai} and ϵ_{ij}.[17] The estimated $\hat{\rho} \approx 0.94$ is remarkably high, driven by the sharp decrease in investment probability at higher subsidy levels illustrated in Figure 3. Comparing the independent and joint estimates (i.e. columns 1 vs. 3 and 2 vs. 4), the audit takeup parameters are all statistically indistinguishable. This is not surprising, because both approaches use the full sample of households.

The investment takeup parameters, however, are all statistically different across the two approaches. The starkest difference is that the "constant" terms $\widehat{\eta_I \xi_j}$ are much more negative in the joint estimates in column 3 compared to the independent estimates in column 1. This reflects self-selection, i.e. the positive correlation between ϵ_{Ai} and ϵ_{ij}. In column 1, the independent pro-

---

[17]We could also use the full set of non-price treatment indicators $\mathbf{T}_i$ as excluded variables, but they reduce precision because they do not significantly affect takeup.

bit estimate of investment takeup using data from the subset of audited households identifies the constant $\widehat{\eta_I \xi_j}$ relevant for those selected households. By contrast, the joint estimate in column 3 returns the constant relevant for the full sample of 101,881 households. The subsample of audited households have high draws of ϵ_{Ai}, and because ρ is large, they also have high draws of ϵ_{ij}. Thus, the constant $\widehat{\eta_I \xi_j}$ for the audited sample is much larger (i.e. more positive) than for the full sample. In practical terms, this again underscores that the households that audited are a self-selected group that is significantly more interested in making investments.

An illuminating way to summarize these self-selection issues is to transform the unobserved attributes into dollar terms. Dividing the estimated constant term $\widehat{\eta_I \xi_j}$ by $\hat{\eta}_I$ gives $\hat{\xi}_j$, the mean unobserved investment attribute in units of dollars. In the sample of audited households in column 1, the mean unobserved investment attribute is slightly positive: $\hat{\xi}_j \approx 0.0485/0.147 \approx \$329$. This indicates that on average, audited households positively value recommended energy efficiency investments above their financial net present value, perhaps due to warm glow or the expectation of a more comfortable weatherized home. While not dispositive, this result suggests a violation of one of the required assumptions for the "accounting approach" to welfare analysis introduced earlier: $\mathbb{E}[\xi_{ij}|I_{ij} = 1] = 0$, i.e. that the mean unobserved attribute of adopted investments is zero.

Self-selection implies that the dollarized unobserved investment attributes are much more negative in the full sample of 101,881 households than in the sub-sample of auditors. While \$329 is the mean unobserved investment attribute in audited sample, the estimates in column 3 imply that \$329 is just above the 99th percentile of unobserved investment attributes in the full sample, and the 98th percentile is approximately -\$5,000. This distribution's steep slope is pinned down by the inelasticity of investment takeup illustrated in Figure 2, and its low mean is pinned down by the low audit and investment takeup in combination with the self-selection effects illustrated in Figure 3. In Section VIII's counterfactual policy simulations, this will mean that even very large counterfactual investment subsidies cannot induce households to make more than about one percent of the total possible investments.

VII   Effects on Energy Use

VII.A   Empirical Strategy

The private and social benefits of energy saved are a crucial part of the welfare evaluation. What are the empirical estimates of energy savings, and how do they compare to the simulation predictions?

Define Y_{it} as natural gas or electricity use (in therms/day or kWh/day) for household i for the billing period ending in date t. $\mathbf{P}_{it}$ is a pair of post-audit indicators: to distinguish shorter-term from longer-term effects, we allow different coefficients on $\mathbf{P}_{it}$ for the first six months vs. later. For the billing period that includes the audit date, we pro-rate $\mathbf{P}_{it}$ on $[0, 1]$ to reflect the share of days after the audit. $\mathbf{W}_{it}$ is a vector of two weather controls: average heating degrees and average cooling degrees in the household's city (Madison or Milwaukee) over the billing period ending in t, from NOAA (2015). ν_i is a household-by-calendar month fixed effect, and μ_m is the set of

month-of-sample indicators for all months m in the sample.[18] The estimating equation is

$$Y_{it} = \alpha \mathbf{P}_{it} + \omega \mathbf{W}_{it} + \nu_i + \mu_m + \varepsilon_{it}. \tag{12}$$

Standard errors are robust and clustered by household to allow for arbitrary serial correlation in ε_{it}.

To compare empirical estimates to the simulation predictions, we also fit Equation (12) with predicted savings as the dependent variable. Denote E_{it} as the total predicted daily savings (in therms/day or kWh/day) for all observed investments made by household i as of the billing period ending in date t. ($E_{it} = 0$ before the first investment is made, and thus $E_{it} = 0$ before the audit.) If the investment install date occurs in the middle of the billing period ending in date t, we pro-rate predicted savings over the billing period.

The simulated savings are for average weather conditions. If the empirically realized weather conditions differ, this could make the empirical results differ from predictions even if predictions are unbiased in average weather. We thus weather-adjust the predictions assuming that savings scale proportionally in degree days.[19] The "realization rate" for natural gas or electricity will be the ratio of the $\hat{\alpha}$ for actual energy use to the $\hat{\alpha}$ for predicted savings.

If $\mathbf{P}_{it} \perp \varepsilon_{it} | (\mathbf{W}_{it}, \nu_i, \mu_m)$, then α is the average causal effect of the audit and ensuing investments. Even under this assumption, however, there are two reasons to interpret α carefully. First, households may simultaneously change utilization behaviors, so α would differ from causal effect of investments under constant utilization. This is important for interpreting the realization rate because TREAT and other simulation models predict savings under constant utilization. A realization rate less than 100 percent could thus reflect either a utilization increase, which the simulation models are explicitly not trying to capture, or systematic modeling bias, which the simulation models should be trying to avoid. This is also important for interpreting the accounting welfare analysis, because a utilization increase generates an increase in consumer welfare that the accounting approach does not capture.

Second, households may make unobserved investments, such as purchasing CFLs or appliances through retailers not affiliated with the programs. This is clearly important for the realization rate, because the simulation models only predict savings from observed investments, and for the accounting welfare analysis, because we do not include the costs of investments we do not observe.

---

[18]For example, there is one μ indicator variable that takes value 1 for all bills t where the midpoint of the billing period occurs in January 2012, then another μ for all bills where the midpoint occurs in February 2012, etc. Then, there is one fixed effect ν for all bills of household i with midpoint in January of any year, a second fixed effect for all bills of household i with midpoint in February of any year, etc.

[19]Specifically, we classify investments into four seasonality categories: constant (hot water and lighting), cooling (cooling system improvements), heating (heating system improvements), and cooling or heating (all others, such as insulation, air sealing, etc.). For cooling (heating) categories, $E_{it}^w = E_{it} \cdot \frac{W_{it}}{\bar{W}_i}$, where W_{it} is the mean base-65 cooling (heating) degree days for billing period t and $\bar{W}_i$ is the average cooling (heating) degrees for household i's city (Madison or Milwaukee) between 2000 and 2011. The equation for E_{it}^w for the "cooling or heating" seasonality category is identical except that W_{it} and $\bar{W}_{it}$ are the sum of heating plus cooling degrees. Appendix Table A.10 presents robustness checks without the weather adjustment; it turns out to make little difference.

We discuss these issues more below.[20]

VII.B   Empirical Results

VII.B.1   Graphical Results

Figure 4 is a standard event time graph, which illustrates effects and allows a visual test for pre-trends. To make the figure, we estimate Equation (12) replacing $\mathbf{P}_{it}$ with indicators for each two-month period within an event window extending 18 months before and after the audit. The figure combines separate estimates for natural gas and electricity, weighting each two-month period's coefficients by sample average retail prices and multiplying by 365 to transform units to annualized retail energy cost savings.

Figure 4 shows that there are no pre-audit trends in energy use. Immediately after the audit, energy use decreases. By about six months post-audit, energy use has stabilized at almost $100 less than the pre-audit annual average. Relative to a pre-audit average of about $1800 per year, this represents a five percent reduction.

The light gray lines are the simulation predictions, from using E_{it} instead of energy use as the dependent variable in the same regression. Predicted energy use decreases immediately after the audit as "direct install" measures (primarily CFLs and low-flow showerheads) are put in, then decreases gradually over the first six months post-audit as consumers make larger investments. By about six months post-audit, predicted energy use has stabilized at approximately $150 less than baseline. Appendix D presents separate figures for natural gas and electricity, as well as a series of other graphical robustness checks.

VII.B.2   Formal Results

Table 5 presents formal estimates of Equation (12). Columns 1 and 2 are for natural gas, while columns 3 and 4 are for electricity. Columns 1 and 3 use the weather-adjusted simulation predictions E_{it} as the dependent variable, while columns 2 and 4 use energy use Y_{it}. Since the typical investment is predicted to last 20 years and Figure 4 shows savings stabilizing only after about six months after the audit, we focus on the coefficients for ≥ 6 months.

After six months, the simulations predict 0.444 therms/day natural gas savings and 0.408 kWh/day electricity savings, or about 19 and 1.9 percent reductions relative to pre-audit mean gas and electricity use, respectively. By contrast, the actual savings were 0.128 therms/day and 1.013 kWh/day, each of which represents about a five percent reduction. Figure 5 summarizes these results. Natural gas savings amount to only 29 percent of predictions, while electricity savings are 248 percent of predictions. In total, the average household that had an audit saves $89 per year in retail gas and electricity costs. The total predicted savings are $153 per year, so the realization rate

---

[20]Neither of these two problems is solved by a randomized encouragement research design. Such a design credibly identifies the causal impact of the encouragement, but households could also change utilization behaviors or make unobserved investments.

is 58 percent at retail prices. We have done extensive additional analysis and robustness checks; interested readers should see Appendix D.

Because the realization rates are so different between natural gas and electricity (29 vs. 248 percent), the relative weight given to the two fuels matters for the overall realization rate. For example, because electricity has a higher retail markup over acquisition cost than natural gas, combining the two fuels at acquisition costs increases the weight on natural gas. Valuing energy at acquisition cost, the average audited household reduces gas and electricity use by \$38 per year, and the realization rate is 41 percent.

What could explain the difference between simulated and empirically estimated savings? For natural gas, additional analyses in Appendix Table A.12 show smaller or statistically zero short-falls for households that made large and common investments (insulation and new heating/cooling systems) but particularly large empirical shortfalls associated with smaller and less common investments. For electricity, the additional analyses suggest that the excess savings could result from unobserved actions such as appliance replacement that would not be recorded in the programs' administrative data. When adjusting the accounting welfare analysis in Section VIII for empirically realized savings, we assume that the excess electricity savings accrue only due to the observed investments. If there were also costly unobserved investments, this would cause us to understate total costs; addressing this would only strengthen our finding that program costs outweighed benefits.

The simulation models assume constant utilization before vs. after the investment, and one natural explanation for the natural gas shortfall could be the "rebound effect," i.e. that consumers increase utilization in response to a decrease in the cost of energy services. Appendix D presents two calculations showing that the rebound effect is highly unlikely to explain the full gas shortfall. First, a utilization elasticity of -0.98 would be required to explain the full shortfall. By contrast, the most closely related estimates of energy utilization elasticity are much smaller: -0.06, -0.3, and -0.22 for washing machines (Davis 2008), home electricity (Dubin and McFadden 1984), and autos (Gillingham 2014), respectively. Second, a large indoor temperature change of 7.5 degrees Fahrenheit would be required to explain the full shortfall. By contrast, the Fowlie, Greenstone, and Wolfram (2015b) find no statistically significant post-weatherization temperature change in a low-income (and thus likely more price elastic) population in Michigan, and they can reject a post-weatherization temperature change of more than 1.4 degrees Fahrenheit with 90 percent confidence in a two-sided test. To the extent that utilization is not fully inelastic, the associated increase in consumer welfare is excluded from the "accounting approach," biasing those benefit estimates downward. In theory, the revealed preference analysis captures this welfare increase as part of ξ_{ij}.

A realization rate of less than 100 percent is consistent with other work. Fowlie, Greenstone, and Wolfram (2015b) find a 40 percent realization rate in their Michigan RCT. They similarly find that the simulation model overestimates natural gas savings and underestimates electricity savings, but since gas comprises a larger share of savings, the overall realization rate is far less than 100 percent. The TREAT software developers found median household-level realization rates around 60-70 percent in recent New York study (PSD 2015b), also finding that their model understated

electricity savings and overstated gas savings.

VIII   Welfare Analysis

In this section, we present welfare analyses using the "accounting" and "revealed preference" approaches. First, we detail assumptions for environmental externalities and market failure parameters ϕ_{ij}, γ_{Ai}, and γ_{ij}.

VIII.A   Assumptions for Externalities and Other Market Failures

In the model, the uninternalized externality reduction ϕ_{ij} from investment j at household i represents the present discounted value of the sum over fuels (electricity, natural gas, and heating oil) of the amount of fuel conserved times that fuel's retail price distortion – i.e., the difference between marginal social cost and marginal retail price.[21] In the U.S., there are two main reasons why marginal retail energy prices are distorted. First, energy use generates pollution and thus environmental externalities. Second, as Davis and Muehlegger (2010) point out, most utilities mark up retail marginal prices above marginal acquisition cost in order to cover fixed costs such as meter reading and overhead, and this can partially or fully offset the environmental externalities.[22]

Table 6 presents our calculation of the retail markup and environmental externality for each of the three fuels in our data, in common units of dollars per million Btu of energy use. Rows 1-3 show that under the energy price assumptions described in Section IV, marginal retail prices exceed marginal acquisition costs by about 300 percent and about 50 percent for electricity and natural gas, respectively. We assume that in the residential heating oil market, retail prices equal marginal costs.

Row 4 of Table 6 presents our estimates of the marginal environmental damages from using each fuel, calculated specifically for homes in Wisconsin. Our estimates include local air pollution (sulfur dioxide, nitrogen oxides, and particulates) and greenhouse gas emissions (both carbon dioxide and methane leakage from natural gas systems). We take standard emission rates and marginal damage estimates from existing literature, including $39 social cost of carbon (Interagency Working Group on the Social Cost of Carbon 2013) and a $6 million value of a statistical life. Natural gas is relatively clean, whereas heating oil and electricity consumed in Wisconsin both involve emissions—primarily of carbon dioxide and sulfur dioxide—that are estimated to have large marginal damages. See Appendix B.C for details on this calculation. Row 5 then presents our estimate of the per-unit *uninternalized* externality for each fuel—that is, the difference between marginal retail price and marginal social cost, taking into account both the environmental externality and the retail markup.

---

[21]Formally, $\phi_{ij} = \sum_{t=1}^{T} \sum_{f} \delta^t \cdot Energy\ Saved_{ijf} \cdot (Marginal\ Social\ Cost_f - Marginal\ Retail\ Price_f)$, where $f \in \{natural\ gas, electricity, fuel\ oil\}$, δ is the discount factor, t indexes time, T is the investment lifetime, and Marginal Social Cost is the per-unit acquisition cost plus the per-unit gross environmental externality.

[22]While rate structures that embed fixed costs in marginal retail prices may be desirable on equity grounds (see Borenstein and Davis 2012), they distort marginal prices. A more economically efficient pricing structure would be to pass through fixed costs as fixed monthly charges and then set marginal retail prices equal to marginal costs.

The table's key result is that the per-unit uninternalized externality varies significantly across the three fuels. For natural gas and especially heating oil, one household's conservation generates a net benefit for others in society, as the environmental externality reductions are not fully off-set by the loss of any retail markups. Electricity is actually the opposite—the retail markup is so large that it exceeds the environmental externality, and a household's electricity consumption thus imposes a net burden on others in society. As we shall see in detail, this means that the Wisconsin programs' approach of subsidizing energy efficiency investments based on the Btus of energy reduction (regardless of fuel) is highly inefficient relative to subsidies that scale in each fuel's uninternalized externality.

The revealed preference welfare analysis also requires estimates of γ_{Ai} and γ_{ij}, the audit and investment takeup distortions, and in our base case, we assume that $\gamma_{Ai} = \gamma_{ij} = 0$. We assume that $\gamma_{Ai} = 0$ in the base case because the RCT found statistically and economically zero treatment effects of variations designed to identify six informational and behavioral barriers to audit takeup. We assume that that $\gamma_{ij} = 0$ in the base case because, in the context of the Wisconsin programs, most of the investment takeup market failures that have been discussed in the energy efficiency literature are not relevant. The home energy audits and written audit reports provide clear information about the costs and benefits of possible investments. There is no credit constraint, as consumers have full financing available. Credit availability also reduces any potential role of present bias: allowing consumers to borrow means that investment does not require a reduction in current consumption. Further, our sample includes only owner-occupied homes, so there is no landlord-tenant distortion discussed by Davis (2012), Gillingham, Harding, and Rapson (2012), and Myers (2015). Due to the inherent uncertainty in these parameters, we will also present welfare estimates under alternative assumptions for γ_{Ai} and γ_{ij}.

VIII.B   Welfare Effects Using the "Accounting Approach"

Table 7 presents welfare effects using the "accounting approach." Costs comprise audit and invest-ment costs, while benefits comprise energy savings (valued at acquisition cost, i.e. from row 2 of Table 6) plus environmental externality reduction (i.e. from row 4 of Table 6). We assume that the unsubsidized audit cost is $c_A = \$400$, based on typical market prices. Column 1 uses the simula-tion predictions to calculate energy and externality reductions. Column 2 multiplies the simulation predictions by realization rates of 2.48 for electricity and 0.29 for natural gas and heating oil, on the basis of Table 5 and Figure 5.[23]

Under the assumptions of the accounting approach, the programs reduce welfare. Using the simulation predictions, the Wisconsin programs have a benefit/cost ratio of 0.92 at a five percent discount rate, with an internal rate of return of 4.0 percent. After adjusting for the empirical shortfall, the benefit/cost ratio is 0.43, and the IRR is negative 4.1 percent.

---

[23]The vast majority of heating oil savings are from oil-to-gas heating system conversions. Although we do not have empirical estimates of the realization rates for heating oil, we need to adjust heating oil savings by the natural gas adjustment factor to avoid predicting artificial decreases in total energy use.

Is it possible that we happened to evaluate two programs with unusually low returns? In Appendix E, we analyze data from all 37 Better Buildings Neighborhood Programs nationwide that reported data to the U.S. Department of Energy (DOE). While we were unable to run experiments or gather recommended investment data from these other sites to carry out the full revealed preference analysis, we can replicate this accounting evaluation. In aggregate, the 37 programs performed worse than the Wisconsin programs. Using the simulation predictions, the national programs have a benefit/cost ratio of 0.75 at five percent discount rates, with an IRR of 1.5 percent. After adjusting with the DOE's own estimate of the empirical shortfall, the benefit/cost ratio is 0.38 and the IRR is negative 5.2 percent. By these metrics, the result is not just that the returns are low—society would have been better off without these programs.

Very different assumptions about energy prices or environmental externalities would be required to make the programs welfare enhancing in Table 7: for empirically adjusted benefits to exceed costs, energy acquisition costs would need to be 4.2 times larger, or all environmental externality damages would need to be 3.3 times larger, or the social cost of carbon would need to be $239 per ton. While those alternative parameters seem unlikely, there is certainly scope for more moderate disagreement. For example, the U.S. Energy Information Administration (2015) Annual Energy Outlook predicts that the benchmark Henry Hub natural gas price will average $1.01 higher in real terms over 2013-2032 than it was over 2011-2014. When passed through to citygate prices, this would increase acquisition costs (and thus the value of natural gas savings) by 19 percent, slightly increasing net benefits. We could also proxy for the social marginal cost of electricity using the cost of a new combined cycle natural gas plant. Using the Annual Energy Outlook levelized cost estimates from EIA (2015) and our same per-unit externality damage assumptions, we calculate an electricity social marginal cost of $0.106 per kilowatt-hour, which is 16 percent lower than our main assumptions in Table 6. This is lower because while the levelized production cost is higher than 2011-2014 market prices, this is more than offset by the natural gas plant's relatively low local air pollution emissions. Using this lower electricity cost would slightly reduce net benefits.

VIII.C   Welfare Effects Using the "Revealed Preference Approach"

VIII.C.1   Simulation Procedure

We now evaluate the welfare impacts of a subsidy change from $\mathbf{s}_0$ to $\mathbf{s}_1$ using the "revealed preference approach." We first show how to determine the welfare effect conditional on one draw of each ϵ_{Ai} and ϵ_{ij}, and we then simulate draws over the joint distribution of ϵ_{Ai} and ϵ_{ij}. In our base case estimates, we use the simulation predictions of energy savings, but we subsequently explore the implications of the estimated energy savings shortfall.

We must first define several terms. We denote household i's vector of ϵ_{Ai} and ϵ_{ij} draws as ϵ_i, and we denote the vector of all households' ϵ_{Ai} and ϵ_{ij} draws as ϵ. We denote the vector of program subsidies as $\mathbf{s}_p$. We use "hats" to signify that a variable depends on empirical estimates, which we take from the joint estimates in column 4 of Table 4. For example, $\hat{\lambda}_i(\mathbf{s}; \epsilon_i) = \sum_{j \in \mathcal{J}_i} \hat{I}_{ij}(\mathbf{s}; \epsilon_{ij}) \cdot$

28

$(\hat{V}_{ij}(\mathbf{s}) + \epsilon_{ij}).^{24}$

The empirical analogue to the utility function in Equation (1) is

$$\hat{U}_i(\mathbf{s}; \epsilon) = y_i - E_{0i} - \hat{T}(\mathbf{s}; \epsilon) + \hat{A}_i(\mathbf{s}; \epsilon_i) \cdot \left\{ \begin{array}{c} \left(\hat{V}_{Ai} + \epsilon_{Ai}\right) + \left(\hat{\lambda}_i(\mathbf{s}; \epsilon_i) - \hat{\lambda}_i(\mathbf{s}_p; \epsilon_i)\right) \\ -\gamma_{Ai} - \sum_{j \in \mathcal{J}_i} \hat{I}_{ij}(\mathbf{s}; \epsilon_{ij}) \cdot \gamma_{ij} \end{array} \right\}. \qquad (13)$$

To understand the terms inside the brackets in Equation (13), recall that $\hat{V}_{Ai} + \epsilon_{Ai}$ is the empirical analogue to $s_{Ai} - c_A + \xi_{Ai} + \lambda_i + \gamma_{Ai}$, the perceived utility from the audit and potential investments. The λ_i in our empirical estimate of $\hat{V}_{Ai}$ is $\hat{\lambda}_i(\mathbf{s}_p; \epsilon_i)$, corresponding to existing program subsidies $\mathbf{s}_p$. Therefore, to simulate $\hat{U}_i$ for a counterfactual subsidy $\mathbf{s}$, we subtract $\hat{\lambda}_i(\mathbf{s}_p; \epsilon_i)$ and add $\hat{\lambda}_i(\mathbf{s}; \epsilon_i)$. If we did not do this, utility would be unaffected by a change in investment subsidies. The final two terms account for possible differences between observed takeup and utility from Equation (1). We subtract γ_{Ai} if consumer i audited, to account for misperceived audit benefits, and we subtract γ_{ij} if consumer i made investment j, to account for misperceived investment benefits. In our base case, where we assume $\gamma_{Ai} = \gamma_{ij} = 0$, Equation (13) provides a standard estimate of consumer surplus based on observed market decisions.

We define $\hat{W}(\mathbf{s}; \epsilon)$ as the empirical analogue to social welfare from Equation (5) conditional on a draw of ϵ. The empirical analogue to the welfare change in Equation (6) is

$$\Delta\hat{W} = \int \hat{W}(\mathbf{s}_1; \epsilon) - \hat{W}(\mathbf{s}_0; \epsilon) \, d\Phi_2(\epsilon_{Ai}, \epsilon_{ij}, \hat{\rho}). \qquad (14)$$

We simulate the integral over 250 draws of ϵ and present the mean of results across draws. Because the sample is so large, results are very similar even with many fewer draws.[25]

The next two subsections present alternative sets of welfare calculations. Because we have estimated the parameters of the utility function, we are able to estimate welfare effects under a broad array of counterfactual policies and alternative assumptions. In Section VIII.C.2, we evaluate the program subsidies and counterfactual policies under the base case assumptions for market failures ϕ_{ij}, γ_{Ai}, and γ_{ij}. In Section VIII.C.3, we evaluate the program subsidies under alternative market failure assumptions.

---

[24]When fitting $\hat{A}_i$ and $\hat{V}_{Ai}$, we make two modifications. First, the latent utility from auditing did not include a term for the (non-experimental) program audit subsidy offer, as this would have been collinear in the estimation. To simulate counterfactuals at different subsidy levels, we re-write the audit takeup intercept as a function of the program audit subsidy s_{Ai} offered to household i: $\kappa_A = \tilde{\kappa}_{Ai} + s_{Ai}$. Second, to reflect conditions in the absence of our RCT, we set $S_{Ei} = G_i = T_i = 0$.

[25]When simulating $\Delta\hat{W}$, any household in the sample could be predicted to audit when it receives high simulation draws of ϵ_i, but we do not have characteristics of recommended investments for households that did not audit. For these households, we simulate monetary characteristics of recommended investments with a random draw of the $\mathcal{J}_i$ for a household that did audit. We find no evidence against this approach: as shown in Appendix Table A.15, the costs, energy savings, and count of a household's recommended investments are uncorrelated with the household's experimental audit subsidy. The simulation results are effectively identical when we simulate investment characteristics for 100 percent of households, instead of only the 99 percent that did not audit.

VIII.C.2 Results: Program Subsidies and Alternative Policies Under Base Case Assumptions

Table 8 reports the simulated effects of alternative residential energy efficiency programs with different audit and investment subsidies. These counterfactual policy exercises rely on the joint estimation results reported in column 4 of Table 4, and they extrapolate along the audit and investment demand curves, which assume a particular functional form for the joint distribution of ϵ_{Ai} and ϵ_{ij}. Panel A details the specific program subsidy assumptions, Panel B reports the impacts on households and uninternalized externalities, and Panel C presents welfare impacts. Throughout the table, we use the base case assumptions for market failures, i.e. $\gamma_{Ai} = \gamma_{ij} = 0$ and ϕ_{ij} from the uninternalized externality assumptions from Table 6.

Column 1 reports on the case where audit and investment subsidies are equal to zero. This counterfactual indicates that the market for home energy audits and retrofits would almost entirely disappear in the absence of government intervention. For example, the results suggest that over the sample period, only 0.008 percent of households would have an audit, and only 0.005 percent of possible energy efficiency investments would be adopted. Across all 101,881 households in the sample, the mean investment expenditures are $0.20, and the mean value of uninternalized externality benefits ϕ_{ij} from adopted investments is $0.02.

Column 2a presents the revealed preference welfare evaluation of the Wisconsin programs at the subsidy levels employed in these programs. Panel A restates that the audit subsidies were $200 and $300 in Madison and Milwaukee, respectively, and that we use the linearized investment subsidy of $5000\frac{E_{ij}}{E_{0i}}$, i.e. 5,000 times the predicted share of savings in household energy use (up to a cap of $3500). The estimates in Panel B suggest that the program subsidies effectively create this market, increasing audit and investment takeup rates to 1.17 and 0.65 percent, respectively. The costs of creating the market are also evident: $14.48 per sample household to fund audit and investment subsidies, or about $1.5 million in total across all 101,881 sample households. Under these subsidies, the average sample household spends $40.34 on energy efficiency investments, which sums to $4.1 million in the full sample. The investments are projected to produce around $9.67 of uninternalized externality reductions over their lifetimes per sample household, or about $0.67 per dollar of subsidies.

Panel C presents the simulated social welfare effects of subsidies compared to the zero-subsidy counterfactual in column 1. The first two rows of Panel C consider the two components of social welfare. The first row reports the change in consumer utility. Per Equation (13), this depends on subsidy amounts, the monetary benefits and costs of investments, draws of ϵ_{Ai} and ϵ_{ij}, and the resulting takeup decisions and lump-sum taxes to fund the subsidies. In this scenario, the existing Wisconsin programs reduce consumer utility (ignoring externalities) by $12.25 per sample household. Any scenario that uses the base case assumption that $\gamma_{Ai} = \gamma_{ij} = 0$ will naturally involve a decrease in consumer utility, because households are being taxed to induce themselves to make decisions that they would not otherwise make, thus distorting choices from the private optimum. Of course, these losses can be offset with social gains from uninternalized externality

reduction. The second row of Panel C shows that relative to the no subsidy counterfactual, the program subsidies induce investments that reduce uninternalized externalities by an average of $9.65 per sample household.

The social welfare change is the consumer utility change plus the uninternalized externality reduction. In total, the program subsidies generate a social welfare loss of $2.60 per household, or about $265,000 in aggregate in our sample. This amounts to a social loss of $0.18 for every $1 in subsidy. Thus, from the perspective of our model, the program subsidies are worse than having no subsidies at all.

Large increases in residential energy efficiency investments play an important role in virtually all climate mitigation plans. Column 2b probes the feasibility of achieving this goal by tripling both audit and investment subsidies. To be clear, this scenario is relatively extreme. It involves $600 and $900 audit subsidies in Madison and Milwaukee, respectively, so given a market price of $400, households would *receive* $200 and $500 to have an audit. Thus, it is not surprising that this causes 9.94 percent of households to audit over the sample period, or about nine times more than the program subsidies in column 2a.

However, this increase in audits is not matched by an increase in investments: only about one percent of possible investments in the sample are adopted, and the average investment across all households is just shy of $69. This finding of a much smaller investment response is largely explained by the earlier evidence that audit subsidy increases draw in households that are increasingly negatively selected in their interest in making investments. This alternative policy would cost more than $140 per household in subsidies and would produce only 30 cents of externality reduction per subsidy dollar, significantly lower than in column 2a. Overall, our model implies that such a subsidy would reduce welfare by about $39 per sample household, meaning that this sort of program expansion is not justified by this metric.

Columns 3a and 3b explore a different approach to structuring the investment subsidies. Specifically, column 3a considers investment subsidies set to exactly offset the uninternalized externality ϕ_{ij}, while column 3b considers investment subsidies of $2\phi_{ij}$. (To parallel the program subsidies, we also cap these subsidies at $3500 per investment.) In both columns, there is no audit subsidy. Except for the $3500 cap, the column 3a subsidies are socially optimal in our model if $\gamma_{Ai} = \gamma_{ij} = 0$ and energy use is fully price inelastic.

There are three differences between the programs' investment subsidies and the "socially optimal" subsidies in column 3a. First, and most importantly, the program subsidies scale in energy savings (in physical energy units such as million British thermal units (mmBtu)) instead of uninternalized externality reductions. As we saw in Table 6, the per-unit uninternalized externality varies substantially between the three different fuels, with heating oil priced substantially below social marginal cost and electricity actually priced above social marginal cost. Thus, the program subsidies induce investments that save more electricity and less heating oil compared to subsidies that are precisely calibrated to the uninternalized externality. Put differently, the subsidies induce investments that save energy but do not necessarily reduce uninternalized externalities.

31

Second, the program subsidies scale as a percentage of household energy consumption, rather than the level of energy savings. The consequence is that low-usage and high-usage households receive very different subsidies for investments that reduces energy consumption by the same amount; this distorts investment decisions away from the social optimum. Third, the program subsidies are relatively generous: they cover an average of 24 percent of cost, compared to 16 percent in column 3a. This induces more energy efficiency investment than would be optimal under our uninternalized externality assumptions.

These three factors matter a lot. The subsidies in column 3a generate a social welfare gain of $21.93 per sample household, or $2.2 million across the full Wisconsin sample. This welfare gain is accomplished with only $8.68 per household in subsidies, and there is $3.05 in uninternalized externality benefit per dollar of subsidy. The primary driver of this improved performance is the stronger targeting of subsidies toward investments that reduce more heating oil and less electricity. Even in column 3b, with subsidies that are double the "social optimum" in our model, the welfare effect is positive, although this is naturally less positive than in 3a because subsidies exceeding ϕ_{ij} induce some investments with social costs larger than social benefits.

Columns 4a and 4b focus on the effects of audit subsidies in isolation. Specifically, they present results for counterfactual programs with zero investment subsidies and audit subsidies equal to the actual program amounts (column 4a) and twice the program amounts (column 4b). At both amounts, audit subsidies substantially reduce social welfare because we have assumed that $\gamma_{Ai} = 0$. Subsidizing audits is an inefficient way to reduce uninternalized externalities compared to simply subsidizing uninternalized externality reduction.

Comparing column 4b to column 4a sharply illustrates the impact of self-selection—that is, the correlation of ϵ_{Ai} and ϵ_{ij}. The higher subsidy increases audit takeup by a factor of 13, from 0.17 percent to 2.26 percent. However, the investment takeup rate and investment expenditures increase by only a factor of five. The bottom row of Panel B shows that as a result, this subsidy increase generates only four cents of externality benefit per subsidy dollar, compared to 17 cents in column 4a. This underscores a core problem that affects efforts to expand energy efficiency programs by drawing in households with higher audit subsidies: the households that are marginal to these subsidies are more likely to be "tire-kickers" who are unlikely to make efficiency investments.

These simulations have limitations. For example, the joint normal functional form assumption determines how demand responds to counterfactual subsidies, so we must be especially cautious in interpreting these out-of-sample predictions. Additionally, the investment takeup elasticity is estimated from non-experimental data and is lower than our *ex ante* expectations. Furthermore, there is considerable uncertainty in the market failure parameters γ_{Ai}, γ_{ij}, and ϕ_{ij}, which we explore in the next section.

Even noting these limitations, two main qualitative conclusions are likely to be robust. First, it is important to calibrate energy efficiency program subsidies to the market failures that might justify them, and failing to do so can leave surprising amounts of money on the table. This message may be even more important outside of the Wisconsin programs: many other programs across the

country offer investment subsidies that scale as a percentage of investment *costs*, thus incentivizing high-cost investments with no regard for either energy savings or environmental benefits. Second, subsidizing audits can be an ineffective way to reduce externalities, both because audits don't directly generate externality reduction and because audit subsidy increases draw in consumers who are less and less likely to invest.

VIII.C.3   Results: Program Subsidies Under Alternative Assumptions

Table 9 presents welfare evaluations of the existing program subsidies (i.e. those in column 2a of Table 8) under alternative assumptions for γ_{Ai}, γ_{ij}, and ϕ_{ij}. Panel A describes the alternative assumptions for each scenario, including the assumed γ_{Ai} and the average assumed γ_{ij} and ϕ_{ij} across all recommended investments. Panel B presents the welfare effects of the program subsidies, relative to no subsidies, paralleling the bottom panel of Table 8.Column 1 re-prints the base case assumptions, giving the same welfare results as column 2a of Table 8. Columns 2-5 take into account the results in Section VII that the simulation predictions overstate natural gas savings and understate electricity savings.[26]

In column 2, we empirically adjust only the uninternalized externality ϕ_{ij}. This is not the same as simply reducing all ϕ_{ij} by the 58 percent realization rate calculated earlier. Instead, the empirically adjusted average ϕ_{ij} of recommended investments happens to be almost exactly zero (i.e., -\$3), meaning that on average, the energy conserved by recommended investments is priced very close to social marginal cost. This happens because as Table 6 demonstrated, the retail marginal price is *above* social marginal cost for electricity, while it is below social marginal cost for natural gas. The empirical adjustment weights electricity substantially more, and gas substantially less, so this re-weighting reduces the average uninternalized externality. Naturally, this reduction in uninternalized externality benefits worsens the program's estimated welfare effects.

Column 3 additionally assumes that consumers took the audit report's energy savings predictions at face value, whereas only the empirically adjusted savings were realized. This misperception implies deadweight loss, because consumers made investments that they would not have made if they knew that the true savings were lower. We accommodate this in the model by calculating the "actual savings" for investment ij (by multiplying the simulation predictions of savings on the audit report by the fuel-specific realization rates from Table 5), then setting γ_{ij} equal to the simulation predictions minus the "actual savings." In practice, this means that the mean γ_{ij} across all recommended investments is \$489, so consumers overestimate the present discounted retail energy savings from the average investment by \$489. Under this scenario, the program subsidies generate a large social welfare loss of \$1.92 per subsidy dollar, or \$2.8 million for the full sample. This starkly

---

[26]As another alternative, we could put aside our assumptions for γ_{Ai}, γ_{ij}, ϕ_{ij}, and energy savings realization rates, and we instead assume that the program subsidies were optimally calibrated: that is, γ_{Ai} equals \$200 (\$300) in Madison (Milwaukee), and the total investment distortion ($\phi_{ij} - \gamma_{ij}$) equals the program investment subsidy. Under these assumptions, the existing program subsidies would be socially optimal, and they would achieve the first best in our model if home energy use were fully price inelastic. The welfare gains amount to \$2.17 per sample household, or \$221,000 for the full sample. While this shows the possible welfare gains if market failures happen to equal the program subsidies, our calibrations in Table 6 imply that this counterfactual is not realistic.

highlights how important it is for energy efficiency programs to accurately inform consumers about how much energy they can expect to save.

Column 4 repeats column 3, except after doubling the social cost of carbon. This somewhat improves the welfare results, although the improvement is attenuated by the fact that the carbon savings are attenuated by the empirically estimated realization rates.

Column 5 also modifies column 3, considering a scenario in which energy efficiency improvements are not capitalized into home resale prices due to asymmetric information problems between home sellers and buyers. While we do not formally model the details of this asymmetric information problem, we can capture its implications: homeowners that improve energy efficiency effectively impose positive externalities on the future buyers of their homes, as the buyers receive a more energy efficient home without paying a higher price. As a benchmark, we add a positive externality of 40 percent of the present discounted value of empirically adjusted retail energy cost savings; that increases the mean investment uninternalized externality by $84.[27] Relative to column 3, this increases the net social welfare gain by about $0.59 per subsidy dollar spent, although the net effect of the program subsidies remains substantially negative.

VIII.D   Discussion

Both the "accounting" and "revealed preference" welfare approaches ignore important additional issues. First, we have used a static instead of dynamic framework. In reality, consumers know that if they do not audit or invest now, they can do so later, and their ξ_{Ai} and ξ_{ij} could vary over time—for example, as they get a raise or do a major renovation. As a result, some consumers who did not audit or invest during our June 2012-September 2013 sample period might do so in the future, although this is less likely in the absence of the program subsidies. Our framework is appropriate for modeling takeup over our specific sample period, but it would not be appropriate for modeling takeup over longer or shorter time periods.

Second, the calculations exclude administrative and marketing costs. In total, the Wisconsin Better Buildings programs (of which the Milwaukee and Madison programs are just a subset) facilitated $25.4 million in retrofits and also incurred $528,000 in marketing and outreach and $11.2 million in other program expenses (DOE 2015b). This suggests that overhead costs were roughly 46 percent of the total retrofit costs. Applying this 0.46 ratio of overhead/retrofit costs to the investment costs in our experimental sample implies $2.1 million of overhead costs that could be allocated to our sample. When these overhead costs are included in the accounting welfare approach, the benefit/cost ratios in Table 7 are now 0.65 and 0.30 with the simulation predictions and empirical adjustment, respectively. This compares to 0.92 and 0.43, respectively, without accounting for overhead costs. Analogously, including overhead costs also substantially worsens the revealed preference welfare calculations; for example, $2.1 million is just slightly less than the

---

[27]This is roughly consistent with a homeowner who sells after 10 years and capitalizes none of the investment in the sale price, giving the buyer another ten years of a 20-year investment lifetime at a five percent annual discount rate.

welfare gains from the "socially optimal" subsidies in Column 3a of Table 8.

Third, part of the goal of the Better Buildings programs was to make investments that would support an economically sustainable home retrofit market. On the basis of their budget records, WECC staff report that more than $7.9 million of WECC's $20 million grant for these programs was spent in ways that continued to support local retrofit markets after the grant ended, including a loan loss reserve, program design efforts, technical training and mentoring for contractors, and funding for cities to continue parts of the programs (Curtis 2017). This does not affect the interpretation of Tables 7 and 8, as those include only direct audit and investment costs, but it means that the overhead costs discussed in the previous paragraph have additional benefits that we have not quantified.

Fourth, the Better Buildings programs were funded with economic stimulus dollars, and job creation was an important motivation.[28] Our analysis is designed to ask whether these programs would be welfare enhancing in the absence of a macroeconomic stimulus benefit.

IX Conclusion

This paper lays out a framework to evaluate the welfare impacts of residential energy efficiency programs in the presence of imperfect information, behavioral biases, and externalities, and implements the framework with the help of a 100,000-household randomized field experiment. The exercise demonstrates that standard revealed preference analysis can be used to evaluate residential energy efficiency programs and that the results can be used to assess a wide range of counterfactual policies.

The empirical results are remarkable in that they run counter to many aspects of the conventional wisdom about energy efficiency programs. First, the data provide no statistically or economically significant evidence of any of the six potential informational and behavioral failures that were tested. Second, consumer decisions imply large unobserved benefits and costs that conventional accounting-style evaluations do not measure. Third, there is evidence of strong self-selection in these programs, such that marginal participants are less likely to make externality-reducing investments; this makes audit subsidies less socially desirable because they attract participants who are unlikely to undertake energy efficiency investments. Fourth, the estimated energy savings from energy efficiency investments are only 58 percent of predicted savings. Fifth, the programs reduced welfare. In the accounting-style evaluation, the programs have an internal rate of return of negative 4.1 percent. In our revealed preference model, the programs reduce welfare by $0.18 per subsidy dollar. The welfare results demonstrate the significant implications of even subtle-seeming policy design issues and underscore both the feasibility and importance of using revealed preference techniques to understand the consequences of energy efficiency programs.

While these results may not be encouraging for the program structure we evaluated, the counterfactual simulations suggest that if Pigouvian taxes or other more direct approaches are infeasible,

---

[28]DOE (2015c) finds that BBNP created or retained net 10,191 full-time-equivalent jobs, or about one job for every $44,000 in federal outlays.

there are significant opportunities for energy efficiency programs to increase welfare. Specifically, we find that a hypothetical program with perfectly calibrated subsidies could increase welfare by $2.53 per subsidy dollar. It is apparent that residential energy efficiency programs have significant potential, but reaching this potential requires restructuring policies to better target the market failures that motivate them.

References

[1] Allcott, Hunt (2013). "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy*, Vol. 5, No. 3 (August), pages 30-66.

[2] Allcott, Hunt, and Michael Greenstone (2012). "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives*, Vol. 26, No. 1 (Winter), pages 3-28.

[3] Allcott, Hunt, and Richard Sweeney (2016). "The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment." *Management Science*, forthcoming.

[4] Allcott, Hunt, and Dmitry Taubinsky (2015). "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market." *American Economic Review*, Vol. 105, No. 8 (August), pages 2501-2538.

[5] Alliance to Save Energy (2013). "Energy Efficiency and the President's Climate Action Plan." Available from http://www.ase.org/resources/energy-efficiency-and-presidents-climate-action-plan. Accessed September 13, 2016.

[6] Altonji, Joseph (1993). "The Demand for and Return to Education When Education Outcomes Are Uncertain." *Journal of Labor Economics*, Vol. 11, No. 1 (August), pages 48-83.

[7] American Council for an Energy Efficient Economy (2015). "Section 111(d) of the Clean Air Act." Available from http://aceee.org/topics/section-111d-clean-air-act. Accessed June 8, 2015.

[8] Attari, Shahzeen, Michael DeKay, Cliff Davidson, and Wandi Bruine de Bruin (2010). "Public Perceptions of Energy Consumption and Savings." *Proceedings of the National Academy of Sciences*, Vol. 107, pages 16054-16059.

[9] Berry, Linda (1984). "The Role of Financial Incentives in Utility-Sponsored Residential Conservation Programs: A Review of Customer Surveys." *Evaluation and Program Planning*, Vol. 7, No. 2, pages 131-141.

[10] Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman (2010). "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *Quarterly Journal of Economics*, Vol. 125, No. 1 (February), pages 263-306.

[11] Billingsley, Megan, Ian Hoffman, Elizabeth Stuart, Steven Schiller, Charles Goldman, and Kristina LaCommare (2014). "The Program Administrator Cost of Saved Energy for Utility Customer-Funded Energy Efficiency Programs." Lawrence Berkeley National Laboratory Report No. 6595E. https://emp.lbl.gov/sites/all/files/lbnl-6595e.pdf

[12] Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, Vol. 128, No. 1 (February), pages 1-51.

[13] Borenstein, Severin, and Lucas Davis (2012). "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets." *Journal of Law and Economics*, Vol. 55, No. 1, pages 75-128.

[14] Bronnenberg, Bart, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse Shapiro (2015). "Do Pharmacists Buy Bayer? Sophisticated Shoppers and the Brand Premium." *Quarterly Journal of Economics*, Vol. 130, No. 4 (November), pages 1669-1726.

[15] Bruhn, Miriam, and David McKenzie (2009). "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, Vol. 1, No. 4 (October), pages 200-232.

[16] Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, Vol. 82, No. 5 (September), pages 1671-1748.

[17] Camilleri, Adrian R., and Richard P. Larrick (2014). "Metric and Scale Design as Choice Architecture Tools." *Journal of Public Policy & Marketing*, Vol. 33, No. 1 (Spring), pages 108-125.

[18] CEE (Consortium for Energy Efficiency) (2013, 2015). "State of the Efficiency Program Industry." Boston MA: Consortium for Energy Efficiency. Available from http://www.cee1.org/annual-industry-reports. Accessed May 20, 2015.

[19] Chamberlain, Gary (1980). "Analysis of Covariance with Qualitative Data." *Review of Economic Studies*, Vol. 47, pages 225-238.

[20] Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence." *American Economic Review*, Vol. 99, No. 4 (September), pages 1145-1177.

[21] Curtis, Monica (2017). Personal communication. Janary 13th.

[22] Davis, Lucas (2012). "Evaluating the Slow Adoption of Energy Efficient Investments: Are Renters Less Likely to Have Energy-Efficient Appliances?" in The Design and Implementation of U.S. Climate Policy, Don Fullerton and Catherine Wolfram, Eds. Chicago, IL: University of Chicago Press.

[23] Davis, Lucas, and Erich Muehlegger (2010). "Do Americans Consume Too Little Natural Gas? An Empirical Test of Marginal Cost Pricing." *RAND Journal of Economics*, Vol. 41, No. 4, pages 791-810.

[24] Davis, Lucas, and Gilbert Metcalf (2016). "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels." *Journal of the Association of Environmental and Resource Economists*, Vol. 3, No. 3 (September), pages 589-625.

[25] DOE (U.S. Department of Energy) (2014). http://energy.gov/eere/wipo/energy-efficiency-and-conservation-block-grant-program. Accessed May 14, 2014.

[26] DOE (U.S. Department of Energy) (2015a). "Better Buildings Neighborhood Program: History." http://energy.gov/eere/better-buildings-neighborhood-program/history. Accessed June 8, 2015.

[27] DOE (U.S. Department of Energy) (2015b). "Wisconsin Data Dashboard." http://energy.gov/eere/better-buildings-neighborhood-program/downloads/wisconsin-data-dashboard. Accessed June 13, 2015.

[28] DOE (U.S. Department of Energy) (2015c). "Evaluation of the Better Buildings Neighborhood Program." http://www1.eere.energy.gov/analysis/pdfs/bbnp_volume_1_final_evaluation_072215.pdf. Accessed August 18, 2015.

[29] Dubin, Jeffrey, and Daniel McFadden (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica*, Vol. 52, No. 2 (March), pages 345-362.

[30] Duflo, Esther, Michael Kremer, and Jonathan Robinson (2011). "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review*, Vol. 101, No. 6, pages 2350-2390.

[31] Dupas, Pascaline (2013). "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment." *Econometrica*, Vol. 82, No. 1 (January), pages 197-228.

[32] Dupas, Pascaline, and Jonathan Robinson (2013). "Why Don't the Poor Save More? Evidence from Health Savings Experiments." *American Economic Review*, Vol. 103, No. 4 (June), pages 1138-1171.

[33] EIA (U.S. Energy Information Administration) (2015). "Annual Energy Outlook." Available from http://www.eia.gov/forecasts/aeo/. Accessed September 3, 2015.

[34] Foster, Andrew, and Mark Rosenzweig (2010). "Microeconomics of Technology Adoption." *Annual Review of Economics*, Vol. 2, pages 395-424.

[35] Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram (2015a). "Are the Non-Monetary Costs of Energy Efficiency Investments Large? Understanding Low Take-up of a Free Energy Efficiency Program." *American Economic Review, Papers and Proceedings*, forthcoming.

[36] Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram (2015b). "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." Working Paper, University of Chicago (June).

[37] Gillingham, Kenneth (2014). "Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California." *Regional Science and Urban Economics*, Vol. 47 (July), pages 13-24.

[38] Gillingham, Kenneth, Matthew Harding, and David Rapson (2012). "Split Incentives and Household Energy Consumption." *The Energy Journal*, Vol. 33, No. 2, pages 37-62.

[39] Gillingham, Kenneth, Richard Newell, and Karen Palmer (2009). "Energy Efficiency Economics and Policy." *Annual Review of Resource Economics*, Vol. 1, pages 597-619.

[40] Greenstone, Michael, Elizabeth Kopits, and Ann Wolverton (2013). "Developing a Social Cost of Carbon for US Regulatory Analysis: A Methodology and Interpretation." *Review of Environmental Economics and Policy*, Vol. 7, No. 1, pages 23-46.

[41] Hanna, Rema, Esther Duflo, and Michael Greenstone (2016). "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves." *American Economic Journal: Economic Policy*, Vol. 8, No. 1 (February), pages 80-114.

[42] Heckman, James (1979). "Sample Selection Bias as a Specification Error." *Econometrica*, Vol. 47, No. 1 (January), pages 153-161.

[43] Interagency Working Group on the Social Cost of Carbon (2013). "Technical Support Document - Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866." Available from http://www.whitehouse.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update.pdf

[44] International Energy Agency (2015). "World Energy Outlook 2015." Available from http://www.worldenergyoutlook.org/weo2015/. Accessed August 18, 2016.

[45] Jaffe, Adam, and Robert Stavins (1994). "The Energy-Efficiency Gap: What Does it Mean?" *Energy Policy*, Vol. 22, No. 10, pages 804-810.

[46] 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).

[47] Manski, Charles (1988). "Should We Subsdize Enrollment in or Completion of Postsecondary Schooling?" Available from http://files.eric.ed.gov/fulltext/ED315523.pdf.

[48] McKinsey & Co. (2009). "Unlocking Energy Efficiency in the U.S. Economy." Available from http://www.mckinsey.com/clientservice/electricpowernaturalgas/downloads/US_energy_efficiency_full_report.pdf. Accessed June 25, 2014.

[49] Meyer, David, and Roger Schvaneveldt (1971). "Facilitation in Recognizing Pairs of Words: Evidence of a Dependence Between Retrieval Operations." *Journal of Experimental Psychology*, Vol. 90, pages 227-234.

[50] Mullainathan, Sendhil, Joshua Schwartzstein, and William Congdon (2012). "A Reduced-Form Approach to Behavioral Public Finance." *Annual Review of Economics*, Vol. 4, pages 17.1-17.30.

[51] Myers, Erica (2015). "Asymmetric Information in Residential Rental Markets: Implications for the Energy Efficiency Gap." Working Paper, UC Berkeley (September). Available from https://ei.haas.berkeley.edu/research/papers/WP246.pdf

[52] National Association of State Energy Officials (2015). "State 111(d) Resource Hub." Available from http://111d.naseo.org/. Accessed June 8, 2015.

[53] Newell, Richard, and Juha Siikamaki (2013). "Nudging Energy Efficiency Behavior: The Role of Information Labels." NBER Working Paper No. 19224 (July).

[54] NOAA (National Oceanic and Atmospheric Administration) (2015). "Climate Data Online." Available from http://www.ncdc.noaa.gov/cdo-web/search. Accessed June 14, 2015.

[55] Potomac Economics (2012-2015). "State of the Market Report for the MISO Electricity Markets." Available from https://www.misoenergy.org/MarketsOperations/IndependentMarketMonitor/Pages/IndependentMarketMon. Accessed August 18, 2015.

[56] PSD (Performance Systems Development) (2015a). "HERS BESTEST & DOE APPROVED." Available from http://psdconsulting.com/software/treat/hers-bestest-doe-approved/. Accessed June 8, 2015.

[57] PSD (Performance Systems Development) (2015b). "NYSERDA Home Performance with Energy Star Realization Rate Attribution Study." Prepared for New York State Energy Research and Development Authority. Ithaca, NY: Performance Systems Development.

[58] Sanstad, Alan, Michael Hanemann, and Maximilian Auffhammer (2006). "End-use Energy Efficiency in a "Post-Carbon" California Economy: Policy Issues and Research Frontiers." In Michael Hanemann and Alexander Farrell, Eds., Managing Greenhouse Gas Emissions in California. Berkeley, CA: The California Climate Change Center at UC Berkeley.

[59] Skumatz, Lisa (2008). "Non-Energy Benefits Publications / Bibliography for Skumatz / SERA." Available from http://www.serainc.com/NEBPublicationsforWebSite2008.pdf. Accessed July 9, 2015.

[60] Suri, Tavneet (2011). "Selection and Comparative Advantage in Technology Adoption." *Econometrica*, Vol. 79, No. 1, pages 159–209.

[61] Turrentine, Thomas, and Kenneth Kurani (2007). "Car Buyers and Fuel Economy?" *Energy Policy*, Vol. 35, pages 1213-1223.

[62] Van de Ven, Wynand, and Bernard Van Praag (1981). "The Demand for Deductibles in Private Health Insurance." *Journal of Econometrics*, Vol. 17, pages 229-252.

[63] Yergin, Daniel (1979). Energy Future: The Report of the Energy Project at the Harvard Business School. New York: Random House.

Tables and Figures

Table 1: **Overview of Related Energy Efficiency Programs**

Panel A: Policy Overview

	(1)	(2)	(3)	(4)
	Better Buildings Neighborhood Program (2010-2013)	Energy Efficiency and Conservation Block Grants (total)	U.S. and Canada Ratepayer-Funded Efficiency Programs (year 2013)	Weatherization Assistance Program (per year, pre-stimulus)
Number of programs	41	~2000	347	~400
Public expenditures	$508 million	$3.2 billion	$8.0 billion	$250 million
Buildings retrofitted	119,404	86,000		100,000
Value of predicted energy saved	$669 million		$2.95 billion	$453 million

Panel B: Common Program Evaluation Assumptions

	Share of programs
Include non-monetary benefits	30%
Use simulation predictions of savings	97%
of which: Use simulation predictions from other states	70%
Do not evaluate programs retrospectively	81%

Notes: "Public expenditures" includes costs of program operation, including administrative costs and subsidies, but not any investment costs paid by consumers. Better Buildings Neighborhood Program information in column 1 is from DOE (2015b,c). Energy Efficiency and Conservation Block Grant information in column 2 is from DOE (2014). U.S. and Canada ratepayer-funded program information in column 3 is from CEE (2015), and energy savings are those that accrue in 2013 only. Weatherization Assistance Program information in column 4 is from DOE (2011,2015) and Eisenberg (2010). Program evaluation assumptions are from Kushler, Nowak, and Witte (2012).

Table 2: **Summary Statistics for Wisconsin Experimental Sample**

Variable	Mean	Std. Dev.	Min.	Max.
Panel A: Household Data				
(N=101,881 households in experimental population)				
House age (years)	67.2	23.0	0	182
Property value ($000s)	156	90.7	0	2816
Building footprint (sq. feet/1000)	1.24	0.43	0	12.1
Madison	0.31	0.46	0	1
Census tract hybrid vehicle share	1.20	1.26	0	8.33
Audited	0.014	0.12	0	1
Invested	0.008	0.09	0	1
Panel B: Simulation Estimates of Monetary Costs and Benefits				
Recommended investments (N=6100 at the 1394 households that had audits)				
Cost ($)	1493	1472	17.1	43,600
Retail energy cost savings ($/year)	86.1	145	0.01	4359
Invested	0.51	0.50	0	1
Adopted investments (N=3834 at the 1394 households that had audits)				
Cost ($)	1206	1208	0.5	14,475
Retail energy cost savings ($/year)	79.9	134	-526	2501
Panel C: Electricity and Natural Gas Usage				
N=1212 households (gas); N=1217 households (electricity)				
Number of gas bills observed	51.9	14.2	9	88
Average gas use (therms/day)	2.50	0.98	0.20	9.59
Number of electricity bills observed	53.2	14.2	8	88
Average electricity use (kWh/day)	20.6	9.27	4.16	80.6

Notes: House age, property value, and building footprint are from county administrative data. Census tract hybrid vehicle share is the percent of registered vehicles in the Census tract that are hybrids, potentially ranging from 0 to 100. Recommended investments in Panel B include only those used for empirical estimates of investment takeup, and adopted investments in Panel B include only those used for the "accounting approach" to welfare analysis. Energy bills are observed only if a household had an audit. Energy prices are averages over 2011-2014. See Appendix B for more details.

Table 3: **Effects of Letter and Subsidy Treatments on Audit and Investment Takeup**

	(1)	(2)
Dependent Variable:	Audited	Invested
Received letter	0.158	0.041
	(0.093)*	(0.070)
Experiment audit subsidy ($00s)	0.525	0.152
	(0.146)***	(0.119)
Subsidy: $25 gift card	-0.006	-0.125
	(0.101)	(0.081)
House age (years)	0.014	0.009
	(0.001)***	(0.001)***
Property value ($millions)	0.750	0.204
	(0.439)*	(0.376)
Building footprint (sq. feet/1000)	0.326	0.130
	(0.076)***	(0.062)**
Madison	0.171	0.018
	(0.113)	(0.089)
Census tract hybrid vehicle share	0.198	0.122
	(0.039)***	(0.031)***
N	101,881	101,881
Dependent variable mean (percent)	1.4	.8

Notes: This table presents estimates of Equation (8), a probit model using the sample of all households in the Wisconsin experiment. We present marginal effects, with coefficients multiplied by 100 for readability. Audited is an indicator for whether the household had a home energy audit, and Invested is an indicator for whether the household made any energy efficiency investment. Robust standard errors in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table 4: **Joint Estimates of Audit and Investment Takeup**

	(1)	(2)	(3)	(4)
	Independent Probits		Joint Estimates	
Audit Parameters				
Received letter	0.0470	0.0482	0.0295	0.0318
	(0.0279)*	(0.0282)*	(0.0253)	(0.0254)
Experiment audit subsidy ($00s)	0.156	0.160	0.178	0.182
	(0.0438)***	(0.0443)***	(0.0436)***	(0.0447)***
Subsidy: $25 gift card	-0.00131	-0.00169	0.0376	0.0372
	(0.0304)	(0.0307)	(0.0280)	(0.0281)
Constant	-2.260	-2.813	-2.265	-2.833
	(0.0236)***	(0.0462)***	(0.0224)***	(0.0459)***
Household covariates	No	Yes	No	Yes
Investment Parameters				
Subsidy - Cost + Savings ($000s)	0.147	0.119	0.0619	0.0497
	(0.0150)***	(0.0173)***	(0.0213)***	(0.0182)***
Constant	0.0485		-2.363	
	(0.0308)		(0.141)***	
Investment category indicators	No	Yes	No	Yes
Household covariates	No	Yes	No	Yes
$\hat{\rho}$			0.941	0.944
			(0.0599)***	(0.0608)***

Notes: Columns 1 and 2 present the independent probit estimates of Equations (8) and (9), while columns 3 and 4 present estimates of Equation (10) using the maximum likelihood estimator of Van de Ven and Van Praag (1981). These are coefficients, not marginal effects. All estimates exclude outlier investments with net private monetary benefits larger than $5,000 in absolute value. $\hat{\rho}$ is the estimated correlation between the errors in the audit and investment takeup equations. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table 5: **Post-Audit Energy Use Changes**

	(1)	(2)	(3)	(4)
	Natural Gas (therms/day)		Electricity (kWh/day)	
Dependent Variable:	Simulation Prediction	Energy Use	Simulation Prediction	Energy Use
Post audit (<6 months)	-0.284	-0.093	-0.296	-0.874
	(0.026)***	(0.029)***	(0.026)***	(0.238)***
Post audit (≥6 months)	-0.444	-0.128	-0.408	-1.013
	(0.032)***	(0.036)***	(0.036)***	(0.319)***
N	61,845	61,845	63,655	63,654

Notes: This table presents estimates of Equation (12) with daily usage of natural gas and electricity, respectively, as the dependent variables. Columns 1 and 3 have the simulation predictions as the dependent variable, while columns 2 and 4 have energy use as the dependent variable. Mean pre-audit natural gas usage is 2.40 therms/day, and mean pre-audit electricity usage is 21.4 kWh/day. Average marginal natural gas price is $0.82 per therm, and average marginal electricity price is $0.136 per kWh. All columns control for heating and cooling degrees, household-by-calendar month fixed effects, and month-of-sample fixed effects. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table 6: **Energy Price and Externality Assumptions ($/million Btu)**

		(1)	(2)	(3)
Row		Electricity	Natural Gas	Heating Oil
1	Marginal retail price	$39.76	$8.19	$25.18
2	Marginal acquisition cost	$9.96	$5.44	$25.18
3	Retail markup (row 1 - row 2)	$29.80	$2.75	$0
4	Environmental externality	$27.26	$4.50	$12.01
5	Uninternalized externality (row 4 - row 3)	-$2.54	$1.75	$12.01

Notes: This table presents energy price and externality assumptions for the three main fuels in the data. All columns are reported in common units of dollars per million Btu. Energy prices are averages over 2011-2014, and externality savings are based on a $39 social cost of carbon and a $6 million value of a statistical life. See Appendix B.C for details.

Table 7: **Welfare Effects: Accounting Approach**

	(1)	(2)
Source of energy savings estimates:	Simulation Predictions	Empirically Adjusted
Cost ($millions)		
Audit costs (at $400 per audit)		0.56
Investment costs		4.52
Total cost		5.08
Energy Savings ($millions present value at 5% discount rate)		
Natural gas	1.92	0.56
Electricity	0.07	0.18
Heating oil	0.61	0.18
Total	2.61	0.92
Externality Reduction ($millions present value at 5% discount rate)		
Climate (at $39 per ton CO2)	1.37	0.57
SO2/NOx/PM	0.72	0.68
Total	2.09	1.26
Summary		
Benefits - Costs ($millions)	-0.39	-2.91
Benefit/Cost ratio	0.92	0.43
Internal rate of return (percent)	4.0	-4.1

Notes: Column 1 uses energy savings projected by simulation models, while column 2 adjusts for empirically observed savings. Column 2 multiplies electricity and gas/heating oil savings from Column 1 by 2.48 and 0.29, respectively, based on the estimates in Table 5. Energy savings are calculated at average wholesale prices over 2011-2014, and externality reductions are based on a $39 social cost of carbon and a $6 million value of a statistical life; see Appendix B.C for details.

Table 8: **Welfare Effects: Revealed Preference Approach, with Base Case Assumptions and Alternative Subsidies**

	(1)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
	No	Program		Externality		Audit	
Scenario:	Subsidy	Subsidy		Subsidy Only		Subsidy Only	
Panel A: Program Subsidy Counterfactuals							
Audit subsidy in Madison/Milwaukee	$0	$200/300	$600/900	$0	$0	$200/300	$400/600
Investment subsidy multiplier on share energy saved	–	5,000	15,000	–	–	–	–
Investment subsidy multiplier on externality reduction	–	–	–	1	2	–	–
Panel B: Program Impacts on Households and Externalities							
Audit probability	0.008%	1.17%	9.94%	0.27%	0.70%	0.17%	2.26%
Investment probability	0.005%	0.65%	1.03%	0.11%	0.40%	0.03%	0.16%
Subsidy paid per sample household	$0	$14.48	$142.18	$8.68	$17.73	$0.58	$13.27
Investment expenditures per sample household	$0.20	$40.34	$68.99	$11.82	$31.42	$1.39	$6.45
Uninternalized externality benefit per sample household	$0.02	$9.67	$42.93	$26.47	$30.29	$0.10	$0.55
Uninternalized externality benefit per dollar subsidy	--	0.67	0.30	3.05	1.71	0.17	0.04
Incremental uninternalized externality benefit per incremental dollar subsidy (relative to previous column)		0.67	0.26		0.42		0.04
Panel C: Program Impacts on Welfare (Relative to the No Subsidy Case in Column 1)							
ΔConsumer utility (net of tax) per sample household		-$12.25	-$82.29	-$4.53	-$10.45	-$0.40	-$10.31
ΔUninternalized externality per sample household		$9.65	$42.92	$26.46	$30.27	$0.08	$0.53
ΔWelfare per sample household		-$2.60	-$39.37	$21.93	$19.83	-$0.32	-$9.78
ΔWelfare per dollar subsidy		-0.18	-0.28	2.53	1.12	-0.54	-0.74
Incremental Δwelfare per incremental dollar subsidy (relative to previous column)		-0.18	-0.29		-0.23		-0.75

Notes: This table reports averages across 250 simulated draws of ϵ. "Program Subsidy" refers to an investment subsidy that scales in the share of household energy saved. "Externality Subsidy" refers to an investment subsidy that scales in uninternalized externality reduction. Both subsidies are restricted to be between $0 and $3,500 per investment. Investment probability divides the number of adopted investments by the total number of possible investments that could be made by the 101,881-household sample. "Per sample household" means per household in the 101,881-household Wisconsin sample. Uninternalized externality reductions are discounted at a five percent discount rate using the programs' assumed investment lifetimes and simulation predictions of savings, based on a $39 social cost of carbon and a $6 million value of a statistical life; see Appendix B.C for details. In Panel C, consumer utility is from Equation (13), and ΔWelfare is from Equation (14).

Table 9: **Welfare Effects of Program Subsidies Under Alternative Market Failure Assumptions**

	(1)	(2)	(3)	(4)	(5)
		Empirically Adjust	Empirically Adjust	Empirically Adjust and	Empirically Adjust and
	Base	Externality	Savings and	Double	40% of Savings
Scenario:	Case	Only	Externality	the SCC	Not Capitalized
Panel A: Market Failure Assumptions					
Audit distortion γ_{Ai} in Madison/Milwaukee	$0	$0	$0	$0	$0
Mean investment distortion γ_{ij}	$0	$0	$489	$489	$489
Mean investment uninternalized externality ϕ_{ij}	$272	-$3	-$3	$159	$243
Panel B: Program Impacts on Welfare (Relative to the No Subsidy Case)					
ΔConsumer utility (net of tax) per sample household	-$12.25	-$12.25	-$27.04	-$27.04	-$27.04
ΔUninternalized externality per sample household	$9.65	-$0.80	-$0.80	$4.50	$7.82
ΔWelfare per sample household	-$2.60	-$13.05	-$27.83	-$22.53	-$19.22
ΔWelfare per dollar subsidy	-0.18	-0.90	-1.92	-1.56	-1.33

Notes: This table reports averages across 250 simulated draws of ϵ. Column 2 adjusts uninternalized externalities ϕ_{ij} by multiplying electricity and gas/heating oil savings by 2.48 and 0.29, respectively, based on the estimates in Table 5. Column 3 repeats column 2, additionally assuming that consumers misperceive energy savings because they do not anticipate the empirical shortfall. Column 4 repeats column 3, except using a $78 social cost of carbon. Column 5 repeats column 3, except assuming that energy efficiency investments impose an additional positive externality on future home buyers equal to 40 percent of present discounted retail energy cost savings. "Per sample household" means per household in the 101,881-household Wisconsin sample. In Panel B, consumer utility is from Equation (13), and ΔWelfare is from Equation (14).

Figure 1: **Audit Takeup Treatment Effects Scaled by Effect of a $1 Subsidy**
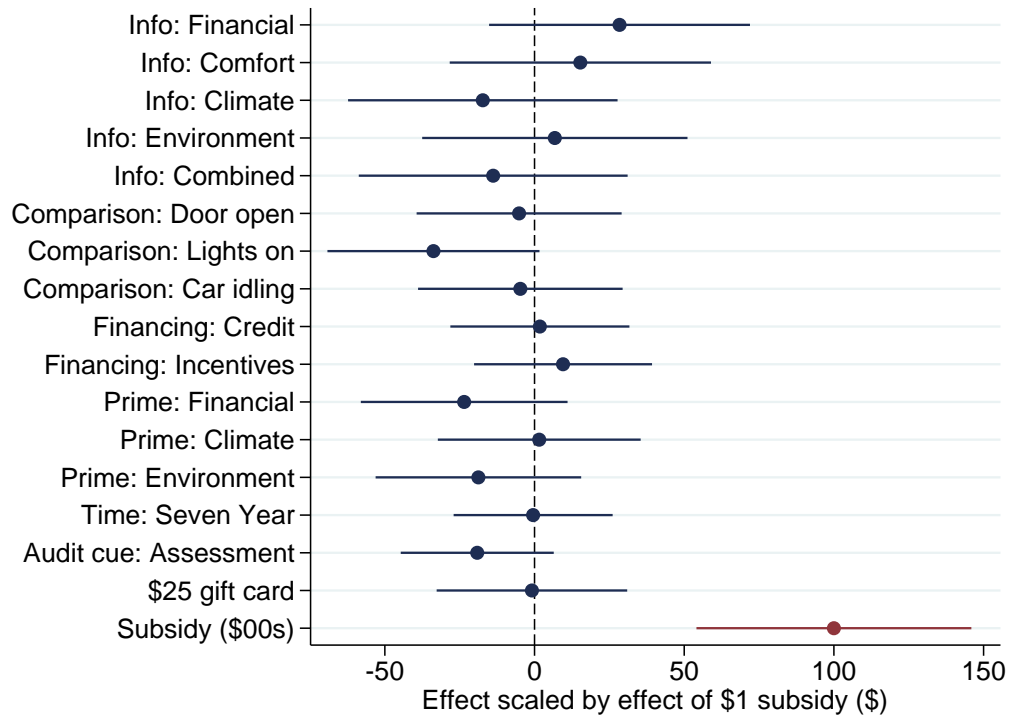
Notes: This figure presents the point estimates and 90 percent confidence intervals for the treatment effect estimates from the Wisconsin audit takeup experiment, scaled by the effect of a $1 subsidy. The estimating equation is Equation (8), and formal estimates are in Appendix Table A.7.

Figure 2: **Net Private Monetary Benefit and Investment Takeup**
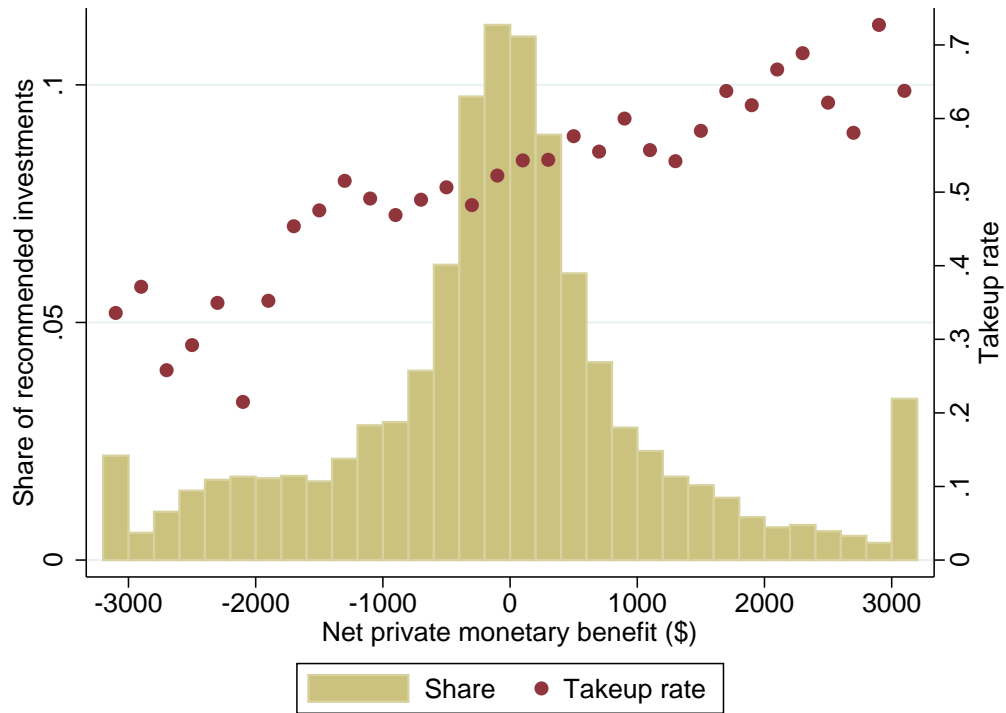
Notes: Net private monetary benefit is the present discounted value of retail energy cost savings net of subsidized cost, or $(S_{ij} - C_{ij} + E_{ij})$ in the notation of the model. Energy cost savings are computed at the 2011-2014 average retail natural gas and electricity prices in the household's city (Madison or Milwaukee). Subsidized costs are the upfront cost net of an imputed subsidy of \$500 per 10 percentage point projected decrease in the household's pre-audit energy use.

Figure 3: **Average and Marginal Investment Probabilities by Subsidy Level**
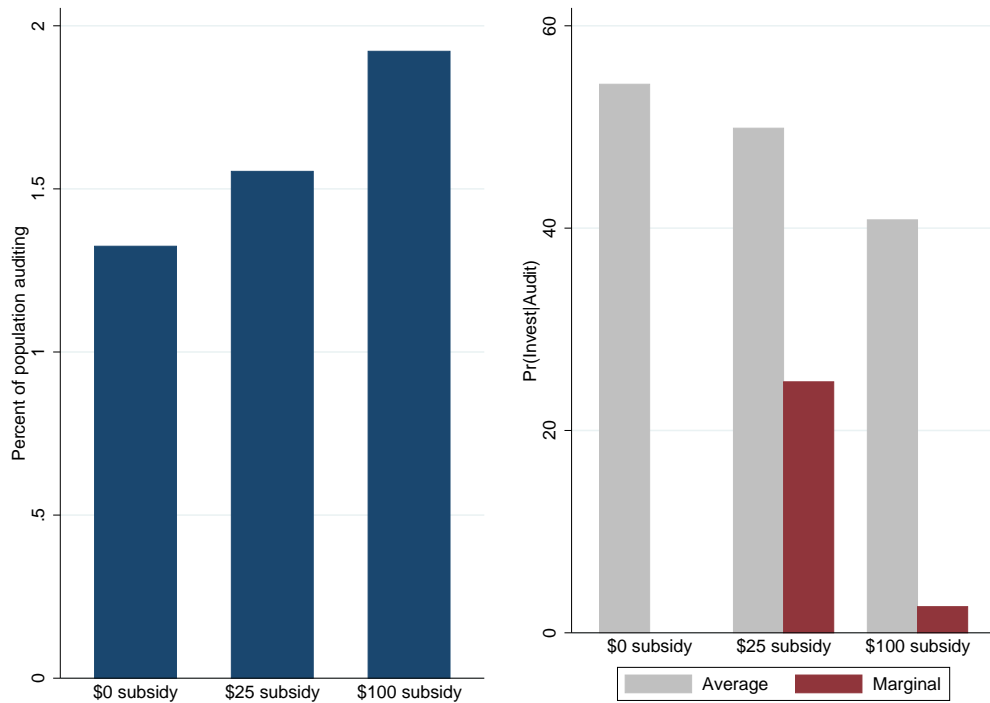
Notes: The left panel presents audit probability for each experimental audit subsidy group. The light (gray) bars on the right panel show the average investment probability by subsidy group, conditional on auditing. The dark (red) bars on the right panel show the investment probability for households marginal to each audit subsidy increase. Households that were in the letter control group or the $25 gift card group are excluded from this figure.

Figure 4: **Energy Use in Event Time**
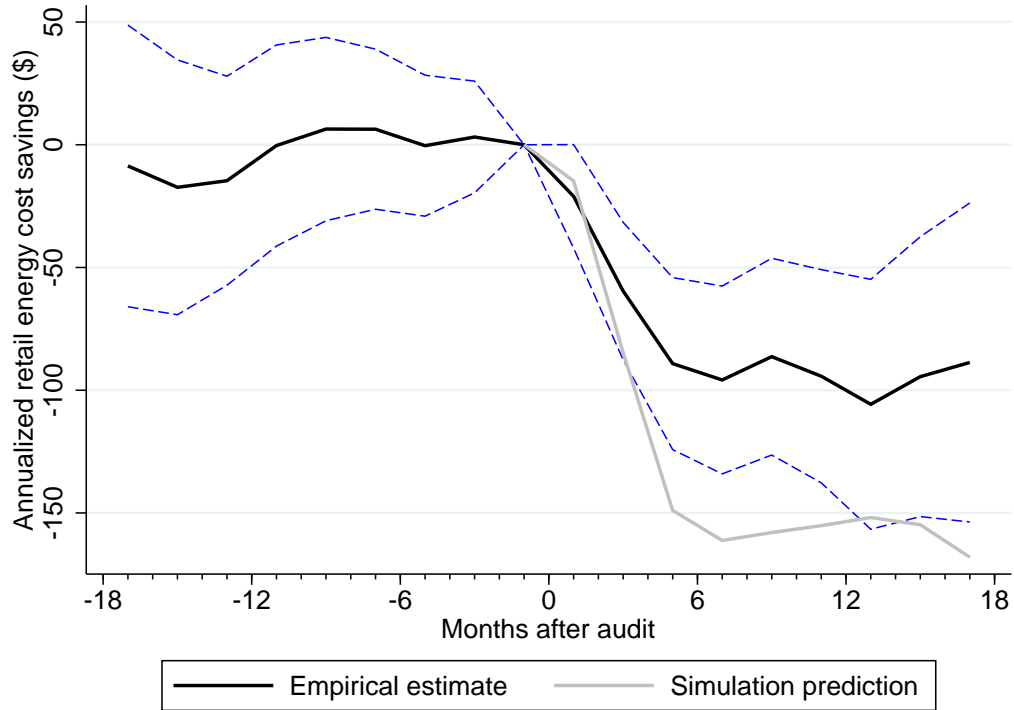
Notes: This figure presents energy use in event time relative to the household's audit. The figure is constructed by estimating Equation (12) separately for natural gas and electricity, replacing P_{it} with indicators for each two-month period within an event window extending 18 months before and after the audit. The excluded category is the month of the audit and the month before. After estimating separate equations for natural gas and electricity, we combine coefficients and standard errors for each two-month period, weighting fuels by their retail prices and multiplying by 365 to transform units to annualized retail cost savings. Dashed lines are 90 percent confidence intervals. Mean pre-audit retail energy costs are $1804 per year.

Figure 5: **Simulation Predictions vs. Empirical Estimates of Post-Audit Energy Savings**
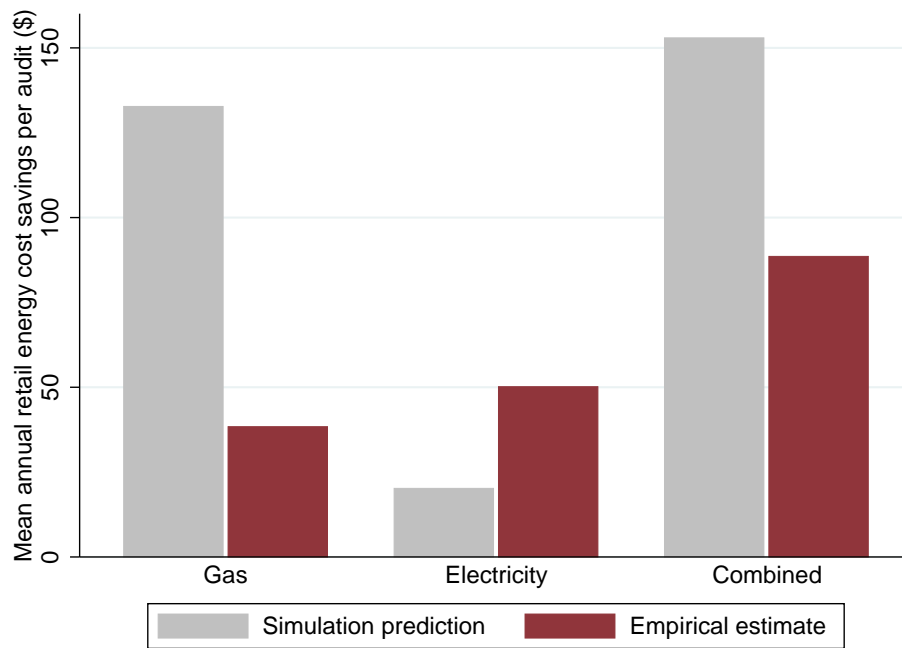
Notes: This figure takes the simulated and actual effects on energy use for households ≥ 6 months post-audit from Table 5. The regression coefficients are in therms per day (for gas) and kilowatt-hours per day (for electricity). They are multiplied by $365 \times$ sample average retail energy prices to give mean annual retail energy cost savings per audit.

MEASURING THE WELFARE EFFECTS OF RESIDENTIAL ENERGY EFFICIENCY
PROGRAMS

Hunt Allcott and Michael Greenstone

Online Appendix: Not for Publication

A   Example Audit Report

Figure A.1: **Example Audit Report (Page 1)**

Figure A.2: **Example Audit Report (Page 2)**

Reference Home

The below diagram is provided for reference as you read your personalized report. It illustrates common sources of home energy-efficiency issues and associated terminology.
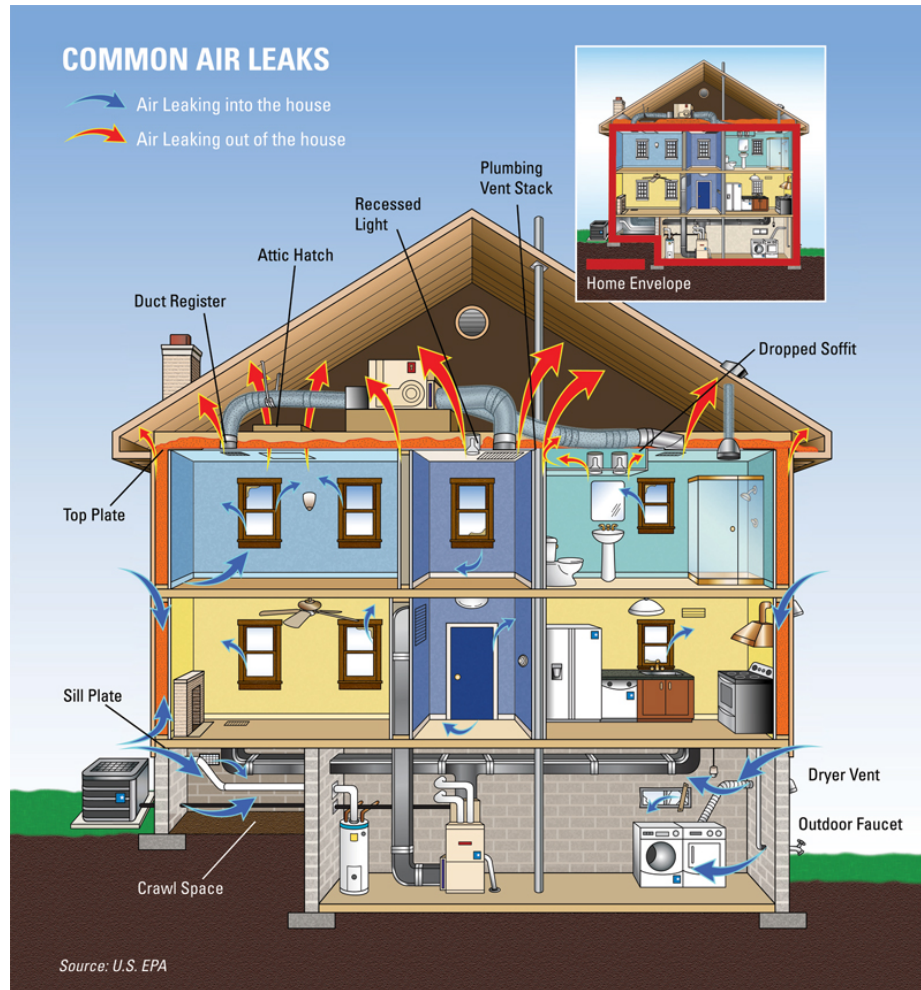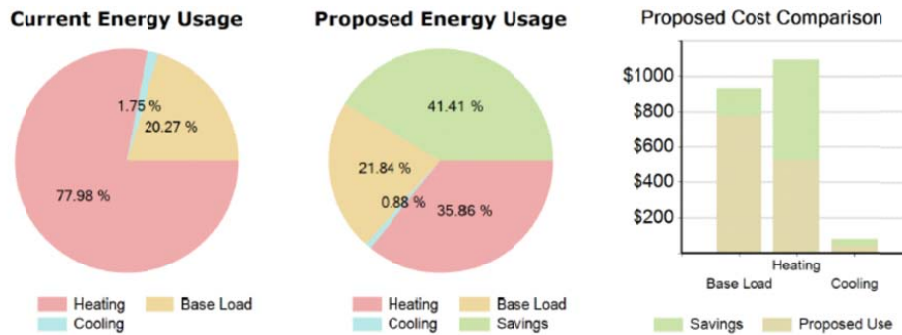
Figure A.3: **Example Audit Report (Page 3)**

Home Energy Snapshot - (Achieve 35+% Energy Savings)

How your home uses energy:

Heating usage includes all energy used to heat your home, similarly for cooling. They are both weather dependent. Base load is the energy use that is independent of the weather. This includes uses like appliances and lighting as well as hot water. This chart shows how your home is currently using energy among these different end uses.

Each improvement affects the energy profile of your house in different ways. Insulation will improve both heating and cooling while replacing a refrigerator will only improve the base-load. This chart indicates the proposed energy usage of your house with all the recommended improvements installed to indicate from where your savings will come.

Required energy-efficiency projects to achieve 35+% energy savings:

PROPOSED Improvement	Customer Cost	Annual Savings	Payback (years)	SIR	Savings Percent	Lifetime Savings
Attic Knee Wall Insulation	$225.00	$63.11	3.6	4.2	4.32%	$1,262.21
Vaulted Ceiling Insulation	$100.00	$12.23	8.2	1.8	0.79%	$244.53
Above Grade Wall Insulation	$3,400.00	$298.76	11.4	1.3	21.46%	$5,975.21
Air Sealing	$800.00	$70.05	11.4	1.3	5.21%	$1,400.91
Water Heater Improvement	$1,750.00	$160.95	10.9	1.0	-	$2,253.28
Rim Joist Insulation	$625.00	$31.18	20.0	0.8	2.25%	$623.55
Furnace Replacement*	$3,250.00	$68.05	47.8	0.3	5.02%	$1,361.00
Attic Insulation	$1,250.00	$21.74	57.5	0.3	1.43%	$434.83
Window Improvement	$11,500.00	$79.08	145.4	0.1	4.87%	$1,581.60

*Please note the software is only able to estimate gas savings for furnace replacements.

Note: Costs and savings are estimates only. Payback and SIR do not factor in cash-back rewards. Work must be completed by Me2-participating contractors to count towards potential bonuses.

The savings percentage of the individual items will not necessarily add up to the overall proposed savings.

Figure A.4: **Example Audit Report (Page 4)**

Energy project costs and financial incentives:

Depending on which improvements you choose to complete, you may be eligible for the following financial incentives:

Project Incentives	
Targeted Energy Savings	**Me^2 Incentives for Energy Savings**
35+% Energy Savings	$2,000*
25-34% Energy Savings	$1,500*
15-24% Energy Savings	$1,000*
10-14% Energy Savings	$750*
Additional Me^2 Incentives		

Complete recommended air sealing work or $1,000 in other energy-efficiency improvements with a Participating Contractor.	$100
Health & Safety Grant (Up to $1,000. Covers necessary electrical upgrades, removal of asbestos/vermiculite, or removal an oil tank. To receive this grant you must also complete the recommended energy-efficiency improvements.)	$1,000

*Incentive amount cannot exceed the energy efficiency improvement costs incurred to the customer. Total rewards from Me2 and Focus on Energy cannot exceed cost of installation.

Plus, you may qualify for special energy-efficiency financing through Summit Credit Union. Energy-efficiency financing between $1,000 and $15,000 is available to Me^2 clients, and you can finance up to 100 percent of the costs of recommended improvements.

Visit SummitCreditUnion.com/SummitMe2 or call 800.236.5560 to explore your options or to get pre-approved.

Focus on Energy cash-back rewards may also be available. Visit www.FocusonEnergy.com or call 800-767-7077 to determine your eligibility for these rewards. The Me^2 program does not guarantee Focus on Energy incentives for any project. Homeowners are responsible for working with their contractor(s) to see if they qualify for any other incentive programs.

B   Data Appendix

B.A   Wisconsin Experiment and Program Details

Table A.1 presents an overview of subsidies paid by the Wisconsin programs. Audit subsidies and the tiered investment subsidy based on percent energy savings are by far the most common subsidies, so these are the two that we model in the paper.

The numbers in the benefit information treatments in the promotional letters were based on our best *ex ante* estimates of the value of energy that the average participant would save. Based on the program's previous estimates, we assumed that a typical weatherization job would reduce energy use by 23 percent. We transformed this to private cost savings using average natural gas and electricity prices. We transformed this into reduced climate damages using emissions factors from the National Academy of Sciences and a $21 social cost of carbon, which was the current official estimate at the time of the experiment (Greenstone, Kopits, and Wolverton 2013). We included no quantitative information about the benefits through local air pollution reduction. Most of the energy saved is natural gas, and since natural gas generates little local air pollution, we calculated relatively small damages. Program staff hypothesized that revealing this would reduce takeup and asked us to remove the quantitative information.

The recipient households were randomly assigned to be mailed letters in one of eight initial waves in summer and fall 2012, which were June 11, June 25, July 9, July 23, September 10, September 24, October 8, and November 22. Households were randomly assigned to be mailed a second letter in one of four final waves, which were mailed on November 12, 2012, and January 7, January 28, and February 4, 2013. The 1,051 households that signed up for an audit after May 2012 but before their initial or final wave's mail date were not sent additional letters. These households are still included in the analysis.

We used a re-randomization algorithm to ensure balance. Specifically, we wrote an algorithm that carried out 500 randomizations for each city (Madison and Milwaukee) using 500 different seeds for the random number generator. For each seed, the algorithm regressed house age, property value, and building area on the set of all treatment group indicators and found the number of p-values less than 0.1 and the minimum p-value of each t-test. For each city, the algorithm then chose the randomization with the largest minimum p-value from the subset of randomizations with the fewest p-values less than 0.1.

Appendix Table A.2 presents tests of balance on observables. Unsurprisingly, F-tests fail to reject balance, with high p-values. The Census tract hybrid vehicle share, which was not part of the re-randomization algorithm, is also highly balanced across treatment groups; that F-test has a p-value of 0.993.

Appendix Table A.3 summarizes costs and benefits by investment category for both the Wisconsin and national programs.

Appendix Tables A.4 and A.5 detail the treatments in the Wisconsin field experiment, and Appendix Figures A.5 and A.6 present example letters from the experiment.

Table A.1: **Overview of Subsidies Paid by Wisconsin Programs**

Type of subsidy	Number of recipient households	Total value ($000s)
Audit subsidy	1338	352
Tiered investment subsidy based on energy savings	779	1080
Time-limited completion bonus	637	204
Blower door tests	635	92
Low-income Home Performance bonus	266	591
Randomly assigned experiment audit subsidy	115	6
Health and safety	36	29
Solar	9	18

Table A.2: **Test of Balance on Observables in Wisconsin Experiment**

Dependent Variable:	(1) House Age	(2) Property Value	(3) Building Area	(4) Madison	(5) Hybrid Share
Received letter	-0.238	-0.000	-0.001	0.002	-0.016
	(0.349)	(0.001)	(0.007)	(0.007)	(0.019)
Info: Financial	0.119	0.001	0.004	-0.002	0.003
	(0.296)	(0.001)	(0.006)	(0.006)	(0.016)
Info: Comfort	0.160	-0.000	0.001	0.001	0.005
	(0.295)	(0.001)	(0.006)	(0.006)	(0.016)
Info: Climate	-0.094	0.000	-0.006	0.001	0.018
	(0.293)	(0.001)	(0.006)	(0.006)	(0.016)
Info: Environment	0.399	-0.000	-0.001	0.001	0.013
	(0.296)	(0.001)	(0.006)	(0.006)	(0.016)
Info: Combined	0.299	-0.001	-0.008	-0.000	0.011
	(0.294)	(0.001)	(0.006)	(0.006)	(0.016)
Comparison: Door open	-0.032	-0.001	-0.000	-0.003	0.001
	(0.230)	(0.001)	(0.004)	(0.005)	(0.013)
Comparison: Lights on	-0.158	0.001	0.004	-0.002	-0.003
	(0.231)	(0.001)	(0.004)	(0.005)	(0.013)
Comparison: Car idling	-0.258	-0.000	-0.001	0.000	-0.000
	(0.228)	(0.001)	(0.004)	(0.005)	(0.013)
Financing: Credit	0.107	0.001	0.006	0.001	0.012
	(0.199)	(0.001)	(0.004)	(0.004)	(0.011)
Financing: Incentives	0.142	-0.000	-0.002	0.000	-0.001
	(0.199)	(0.001)	(0.004)	(0.004)	(0.011)
Prime: Financial	0.185	-0.001	-0.002	0.001	-0.014
	(0.230)	(0.001)	(0.004)	(0.005)	(0.013)
Prime: Climate	0.101	-0.000	0.004	0.002	-0.012
	(0.230)	(0.001)	(0.004)	(0.005)	(0.013)
Prime: Environment	0.187	-0.000	-0.001	-0.001	-0.002
	(0.230)	(0.001)	(0.004)	(0.005)	(0.013)
Time: Seven year	-0.066	0.001	0.003	-0.002	0.006
	(0.178)	(0.001)	(0.003)	(0.004)	(0.010)
Audit cue: Assessment	-0.065	0.000	0.002	-0.001	0.001
	(0.172)	(0.001)	(0.003)	(0.003)	(0.010)
Subsidy: $100	0.111	0.001	0.002	0.002	0.005
	(0.355)	(0.001)	(0.007)	(0.007)	(0.019)
Subsidy: $25	0.051	-0.001	-0.003	-0.001	-0.001
	(0.204)	(0.001)	(0.004)	(0.004)	(0.011)
Subsidy: $25 gift card	-0.035	0.001	-0.004	0.001	0.003
	(0.213)	(0.001)	(0.004)	(0.004)	(0.012)
N	101,881	101,881	101,881	101,881	101,881
F-test p-Value	.992	.878	.582	1	.993

Notes: Robust standard errors in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A.3: **Summary of Adopted Investments in Wisconsin and National Programs**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	National Pro-grams	*Wisconsin Programs*					
Category	Number of adoptions	Number on audit reports	Number of adoptions	Mean cost (\$)	Mean whole-sale energy savings (\$/year)	Median internal rate of return (%)	Mean assumed measure life
Insulation	50,622	3923	2059	1233	51	-3.4	20.0
Air sealing	27,454	1358	790	951	67	1.9	20.0
Exhaust fan	2,523	839	428	697	0	.	20.0
New heating/cooling system	25,567	707	279	3584	156	-7.7	18.5
New water heater	10,025	461	136	1711	13	-18.1	14.0
Replace lighting with CFLs	11,634	455	301	8.93	9	127	6.0
Aerators/showerheads	9,451	404	186	2.17	5	375.5	20.0
New windows	6,023	110	16	2867	9	-12.7	20.0
Programmable thermostat	5,104	49	61	305	21	-8.3	11.0
Pipe/duct sealing/insulation	11,629	35	6	481	41	1.9	18.3
Appliance replacement	1,959	33	0	.	.	.	.
Chimney liner	1,139	21	0	.	.	.	.
Unspecified weatherization	11,550	.	.	.	.	.	.

Notes: Adoption counts for national programs reflect the number of households that made an investment in that category within the 58,418 households that have valid data. Count on audit reports and adoption counts for Wisconsin programs reflect the number of investments within the category from the 1394 households that had audits, possibly including more than one per household. Cost is unsubsidized upfront cost, and energy savings are based on simulation estimates and 2011-2014 wholesale energy price averages. Internal rate of return is at unsubsidized costs and energy acquisition costs.

Table A.4: **Details of "Informational" Treatments in Wisconsin Experiment**

Treatment	Share	Text
Benefit Info		
Financial	1/6	**More savings.** Using less energy saves you money. Act on recommended energy-efficiency improvements and you could save around $[340 / 2,400] on your energy bills over the next [one year / seven years].
Comfort	1/6	**Feel better at home.** Using less energy can lead to greater comfort and a healthier home. When done right, energy-efficiency projects can improve indoor air quality while reducing humidity, drafts, and mold. Plus, you could save around $[340 / 2,400] on your energy bills over the next [one year / seven years].
Climate	1/6	**Reduce your carbon footprint.** Using less energy reduces greenhouse gas emissions that can contribute to climate change. Climate experts calculate that your reduced greenhouse gas emissions over the next [one year / seven years] would benefit people across the world by the equivalent of $[60 / 420]. Plus, you could save around $[340 / 2,400] on your energy bills over that same time.
Environment	1/6	**Help the planet.** Using less energy reduces local air pollution. Reductions in air pollution can lead to longer lives and reduced allergies and asthma, creating a healthier environment for you and your community. Plus, you could save around $[340 / 2,400] on your energy bills over the next seven years.
Combined	1/6	**Savings for you, your community, and our world.** Using less energy saves you money while reducing air pollution and greenhouse gases. Reductions in air pollution can lead to longer lives and reduced allergies and asthma, creating a healthier environment for you and your community. Climate experts calculate that your reduced greenhouse gas emissions over the next seven years would benefit people across the world by the equivalent of $[60 / 420]. Plus, you could save around $[340 / 2,400] on your energy bills over that same time.
Control	1/6	**Reduce energy use.** Reduce your energy use at home and enjoy the many benefits of energy efficiency. You'll soon see that conserving energy is great for you, your family, and the greater community.
Financing		
Credit	1/3	Low-interest, flexible term financing is available. Energy savings can help pay off your loan!
Incentives	1/3	Collect cash incentives!
Control	1/3	Remember, saving energy is smart for everyone!
Comparison		Giving up this opportunity ...
Door	1/6	wastes as much energy as leaving your front door wide open for three hours each day in the winter.
Light	1/6	wastes more energy than leaving all the lights in your house on all day, every day.
Car	1/6	wastes about as much energy as leaving your car idling outside for an hour every day.
Control	1/2	is like missing out on the chance to improve your home.

Table A.5: **Details of "Behavioral" and Subsidy Treatments in Wisconsin Experiment**

Treatment	Share	Text
Graphical Prime		
Financial	1/4	**make BANK.** Shrink your energy bills. Get paid.
		CASH in. Save energy. Save money.
Environmental	1/4	**earth LOVER.** Save energy at home. Save the planet.
		breathe EASY. Energy-efficient home. Cleaner air.
Climate	1/4	**take CONTROL.** Save energy. Reduce your carbon footprint.
		PEACE of MIND. Energy-efficient home. Brighter future
Comfort	1/4	**COMFY.** Save energy. Live comfortably.
		COZY. Feel better in your home.
Time Frame		
Seven years	1/2	Example: "you could save around $2,400 on your energy bills over the next seven years."
One year	1/2	Example: "you could save around $340 on your energy bills over the next year."
Audit Cue		
Assessment	2/3	"home energy assessment"
Audit	1/3	"home energy audit"
Subsidy**		Call to schedule a home energy [assessment/audit]. Usual cost: $400 ...
$100 Subsidy	0.06	You pay [nothing / only $200]!
$25 Subsidy	0.23	You pay only $[75 / 175]!
$25 Gift Card	0.20	You pay only $[100 / 200]! Bonus! Receive a $25 Visa Reward Card just for completing an [assessment/audit]!
Control	0.51	You pay only $[100 / 200]!

**: These prices differed by city: the program subsidized the base price of audits (without the additional experimental subsidy) to $100 in Milwaukee and $200 in Madison.

Figure A.5: **Wisconsin Experiment Sample Letter 1**

Figure A.6: **Wisconsin Experiment Sample Letter 2**

B.B   Better Buildings Neighborhood Program Data Preparation

The Better Buildings Neighborhood Program (BBNP) website provides household-level microdata for 75,110 retrofits of single-family homes at 37 different sites.[29] For each home, the data include total unsubsidized cost, simulation estimates of total annual energy savings in dollars (top-coded at $2500) and in physical quantities of natural gas, electricity, heating oil, propane, kerosene, and wood, and counts of investments made by category. We combine the physical quantities with state-specific average energy prices for 2011-2014 (described below) to construct annual cost savings, then construct present discounted values using five percent discount rates and investment life assumptions from Heaney and Polly (2015).

Column 1 of Appendix Table A.3 presents counts of adopted investments. The distribution of investment categories in Wisconsin is similar to in the nationwide program: primarily insulation, air sealing, and new heating and cooling systems.

Because the data were reported by 37 different agencies, there are some inconsistencies and implausible observations. We drop observations with:

- missing or negative costs,

- missing or negative constructed cost savings,

- reported energy cost savings that are not top-coded and differ from our constructed cost savings by more than a factor of two, if either the reported or constructed annual energy cost savings exceed $1000,

- payback periods faster than one year and costs greater than $2500 (these are primarily air sealing, insulation, and windows, not CFL replacement, so a one-year payback strongly suggests misreported data), or

- constructed cost savings less than the 1st percentile or larger than the 99th percentile, similar to the Heaney and Polly (2015) approach.

This leaves 58,418 valid retrofits in our sample.

B.C   Energy Price and Externality Assumptions

All fuel prices are means for 2011-2014. We observe the state for each home retrofitted in the national Better Buildings Neighborhood Program data, so we match in state-specific average fuel prices for 2011-2014.[30] We use natural gas "citygate" prices as wholesale prices.

---

[29]The data are available from http://energy.gov/eere/better-buildings-neighborhood-program/downloads/better-buildings-neighborhood-program-data-1, with documentation at http://energy.gov/eere/better-buildings-neighborhood-program/downloads/better-buildings-neighborhood-program-data.

[30]For retail electricity, these are from the EIA "Electricity Data Browser," available from http://www.eia.gov/electricity/data.cfm. Quantity-weighted "all-in" wholesale electricity prices for 2011-2014 in the ERCOT, New York, New England, PJM, Midcontinent, and California power markets are from data reported by market operators to Potomac Economics and presented in Figure 4 of Potomac Economics (2015b). These prices include energy prices, ancillary services and uplift charges, and capacity costs where applicable, so they reflect the long-run marginal cost of electricity. We map each US state to one of these markets, and we impute the national average price for states that are not in one of the markets. Citygate natural gas prices are from http://www.eia.gov/dnav/ng/ng_pri_sum_a_epg0_pg1_dmcf_a.htm. Residential retail natural gas prices are from http://www.eia.gov/dnav/ng/ng_pri_sum_a_epg0_prs_dmcf_a.htm. Residential heating oil prices are from http://www.eia.gov/dnav/pet/pet_pri_wfr_a_epd2f_prs_dpgal_w.htm, and residential propane prices are from http://www.eia.gov/dnav/pet/pet_pri_wfr_a_epllpa_prs_dpgal_w.htm.

For electricity and retail natural gas in the Wisconsin programs, we gathered marginal retail prices from staff at the Madison and Milwaukee utilities. Both utilities use time-invariant residential electricity prices, except that Madison charges higher prices during the June-September summer peak. We construct a consumption-weighted average price using 2011-2014 monthly average electricity consumption data for Madison Gas and Electric reported on the EIA Form 826. For natural gas, both utilities have prices that vary by month. We construct consumption-weighted average retail prices using Wisconsin 2011-2014 monthly average retail natural gas consumption from EIA.[31] For citygate natural gas, wholesale electricity, and residential heating oil prices, we use the Wisconsin data from the same EIA sources as for the national program.

Externality damage assumptions build primarily on Holland, Mansur, Muller, and Yates (2015), who graciously shared their data with us. Their key assumptions are a \$6 million value of a statistical life and a fine particulate dose response function from Pope *et al.* (2002). For electricity, Holland *et al.* (2015) provide estimates of average marginal damages by pollutant, NERC region, and hour. We weight the hours by residential load shapes from the Integral Analytics DSMore model and sum across pollutants to get average marginal damages by state per kilowatt-hour conserved. Holland *et al.* also provided county-by-pollutant marginal damages for non-point sources, which are relevant for natural gas and heating oil combustion from sources such as homes that do not have smokestacks. We inflate these to 2013 dollars, collapse these to the state level, weighting by 2011 population, and multiply the state average marginal damages by pollutant-specific emission factors for natural gas, heating oil, propane, kerosene, and wood from the AP-42 database (EPA 1995). We use a \$39 social cost of carbon, based on estimates from the Interagency Working Group on the Social Cost of Carbon (2013) for 2013 and inflated to real 2013 dollars.

Drawing on Howarth *et al.* (2012) and Abrahams *et al.* (2015), we assume that three percent of natural gas leaks as methane during drilling and transportation before arriving in homes. We translate this to carbon dioxide equivalents using a methane global warming potential of 34 from the Intergovernmental Panel on Climate Change (Myhre *et al.* 2013). This increases the natural gas climate change externality by 38 percent, or \$0.10/therm for residential natural gas.[32]

For heating oil sulfur content, we assume 2500 parts per million, except in Connecticut, Delaware, Maine, Massachusetts, New Hampshire, New Jersey, New York, and Pennsylvania, which are phasing in maximum fuel oil sulfur content regulations by 2018. In these states, we construct the present value of average sulfur content over 2013-2032, using a five percent discount rate and maximum allowable fuel sulfur content reported by the New England Fuel Institute (2014). We assume that maximum fuel oil sulfur content restrictions larger than 2500 ppm are non-binding.

Appendix Table A.6 summarizes the energy price and externality assumptions. Climate externality refers to carbon dioxide emissions, plus methane leakage for residential natural gas. Column 1 presents averages for the Wisconsin experiment population. Retail gas and electricity prices differ only slightly between Madison and Milwaukee, and the table presents the average weighted by the share of the audited population in each city. Column 2 presents averages across all households in the national sample.

The table highlights three key issues. First, retail prices substantially exceed wholesale prices for both electricity and natural gas. Indeed, consistent with Davis and Muehlegger (2010), the retail natural gas markup more than exceeds the externality damages for the national programs, meaning that it may be optimal to *tax* investments that conserve natural gas. This is not the case

---

[31]This is available from http://tonto.eia.gov/dnav/ng/hist/n3010wi2m.htm.

[32]Although this leakage also affects climate damages from electricity consumption, we do not include this because our data from Holland, Mansur, Muller, and Yates do not include whether natural gas was on the margin. If natural gas were on the margin 50 percent of the time, the electric power climate externality in Appendix Table A.6 would be $38\% \times 50\% \times 2.3$ cents/kWh ≈ 0.45 cents/kWh larger.

16

in Wisconsin, as the retail markup is lower. Second, the SO2/NOx/PM externality for electricity is more than twice the climate externality. This is largely driven by the SO2 component. Of course, reducing sulfur content in coal and fuel oil and/or using post-combustion controls would be more direct approaches to sulfur emission abatement than reducing energy use. Third, the retail price distortion per unit of energy use (in the bottom row of each panel) differs markedly between the three fuels. As discussed in the body of the paper, this means that subsidizing physical units of energy reduction is very different than subsidizing to correct uninternalized externalities.

Table A.6: **Energy Price and Externality Assumptions**

	(1)	(2)
	Wisconsin Programs	National Programs
Electricity ($/kWh)		
Retail price	0.136	0.139
Wholesale price	0.034	0.049
Climate externality	0.024	0.023
SO2/NOx/PM externality	0.069	0.048
Retail price-social cost	0.009	0.019
Retail price-social cost ($/mmBtu)	2.54	5.57
Natural gas ($/therm)		
Retail price	0.82	1.10
Citygate price	0.54	0.53
Climate externality	0.35	0.35
SO2/NOx/PM externality	0.10	0.09
Retail price-social cost	-0.17	0.13
Retail price-social cost ($/mmBtu)	-1.75	1.30
Heating oil ($/gallon)		
Price	3.50	3.82
Climate externality	0.48	0.48
SO2/NOx/PM externality	1.19	1.14
Retail price-social cost	-1.67	-1.62
Retail price-social cost ($/mmBtu)	-12.01	-11.65

Notes: This table presents household-weighted averages of prices and externality damages per unit of energy. Energy prices are averages over 2011-2014, and externality savings are based on a $39 social cost of carbon and a $6 million value of a statistical life. The bottom row of each panel uses the fact that there are 293 kWh, 10 therms, or 7.19 gallons of electricity, natural gas, or heating oil per million British thermal units (mmBtu), respectively.

C   Appendix to Audit and Investment Takeup Estimates

C.A   Additional Experimental Results

Table A.7: **Effects of All Treatment Variations on Audit and Investment Takeup**

	(1)	(2)
Dependent Variable:	Audited	Invested
Info: Financial	0.148	0.149
	(0.138)	(0.109)
Info: Comfort	0.080	0.103
	(0.138)	(0.108)
Info: Climate	-0.090	-0.086
	(0.142)	(0.114)
Info: Environment	0.035	0.148
	(0.140)	(0.108)
Info: Combined	-0.072	-0.034
	(0.142)	(0.113)
Comparison: Door open	-0.027	0.070
	(0.108)	(0.083)
Comparison: Lights on	-0.176	-0.076
	(0.112)	(0.087)
Comparison: Car idling	-0.025	-0.007
	(0.108)	(0.085)
Financing: Credit	0.009	0.047
	(0.094)	(0.074)
Financing: Incentives	0.050	0.067
	(0.094)	(0.074)
Prime: Financial	-0.122	-0.108
	(0.109)	(0.086)
Prime: Climate	0.008	0.055
	(0.107)	(0.083)
Prime: Environment	-0.098	-0.099
	(0.108)	(0.086)
Time: Seven year	-0.002	-0.009
	(0.084)	(0.066)
Audit cue: Assessment	-0.100	-0.018
	(0.081)	(0.064)
Subsidy: $25 gift card	-0.005	-0.124
	(0.100)	(0.081)
Experiment audit subsidy ($00s)	0.521	0.153
	(0.145)***	(0.118)
N	101,881	101,881

Notes: This table presents estimates of Equation (8). We present marginal effects, with coefficients multiplied by 100 for readability. Audited is an indicator for whether the household had a home energy audit, and Invested is an indicator for whether the household made any energy efficiency investment. All regressions include controls for house age, property value, building footprint, city, and hybrid share. Robust standard errors in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A.8: **Tests of Whether Groups of Treatments Jointly Affect Takeup**

| | (1) | (2) |
	Audited	Invested
Dependent Variable:		
Benefit info treatments	0.45	0.11
Comparison treatments	0.47	0.58
Graphical prime treatments	0.52	0.16
All "informational" treatments	0.63	0.24
All "behavioral" treatments	0.58	0.38
All non-subsidy treatments	0.67	0.24
Subsidy treatments	0.001	0.07

Notes: This table presents p-values of Wald tests of whether groups of coefficients in Appendix Table A.7 jointly differ from zero. Audited is an indicator for whether the household had a home energy audit, and Invested is an indicator for whether the household made any energy efficiency investment.

C.B   Investment Takeup Estimates

Appendix Figure A.7 re-creates Figure 2 with internal rate of return instead of net benefit on the x-axis. Because IRR and net benefit are closely connected, the figure looks very similar.

Appendix Table A.9 presents marginal effects from probit estimates of Equation (9). Except for in column 4, we limit the sample to observations with $\|(S_{ij} - C_{ij} + E_{ij})\| \leq \$5000$ to reduce the influence of outliers. Column 1 presents estimates excluding household covariates $\mathbf{X}_i$. Consistent with Figure 2, takeup increases by 5.9 percentage points for every \$1000 increase in projected monetary benefit.

As discussed in the body of the paper, because $(S_{ij} - C_{ij} + E_{ij})$ is not randomly assigned, we can only cautiously interpret $\hat{\eta}_I$ as a demand slope. The unobservable ϵ_{ij} would be positively correlated with E_{ij} if energy savings bring warm glow utility or are associated with more in-home comfort. Furthermore, ϵ_{ij} would be negatively correlated with C_{ij} if higher-cost projects also require more non-monetary effort to implement, e.g. if larger home construction jobs are both more costly and more of a hassle for the homeowner. Both of these possible correlations would bias $\hat{\eta}_I$ upward, meaning that investment takeup might be even more inelastic than we estimate.

Columns 2 and 3 explore these concerns. Column 2 adds the $\mathbf{X}_i$ covariates to column 1; the marginal effect of $\hat{\eta}_I$ does not change. Column 3 allows separate $\hat{\eta}_I$ coefficients on future savings E_{ij} vs. net upfront costs $C_{ij} - S_{ij}$. The coefficients are economically and statistically similar to each other and to the $\hat{\eta}_I$ coefficients in columns 1 and 2. Thus, either the η_I coefficient does not have large bias, or ϵ_{ij} happens to have very similar correlations with E_{ij} vs. $C_{ij} - S_{ij}$.

Column 4 includes the additional 110 recommended investments with $\|(S_{ij} - C_{ij} + E_{ij})\| > \$5000$. These estimates could be more heavily driven by outliers – for example, there are four recommendations with $\|(S_{ij} - C_{ij} + E_{ij})\| \geq \$25,000$. The $\hat{\eta}_I$ coefficient is smaller, further reinforcing the finding of inelastic demand.

The theoretical framework and welfare analysis allow for various market failures that might affect demand for audit and investments. One potential example is asymmetric information problems that could prevent investments' full benefits from being capitalized into home resale prices. If people have some foresight into their possible future moves, such a market failure would cause households that move sooner to be less likely to invest. However, column 5 shows that the 37 audited households that close their utility accounts more than six months post-audit but during our sample are no less likely to invest. This provides no evidence that this potential market failure is relevant for investment decisions, although it could in principle affect whether people have audits.

Column 6 includes an additional control for the experimental audit subsidy offered to household i, in units of \$100s. This subsidy does not affect investment incentives conditional on auditing, but it does cause selection into auditing. Households that audited under a \$100 larger audit subsidy are a remarkable 12.9 percentage points less likely to invest, conditional on $\mathbf{X}_i$ and $S_{ij} - C_{ij} + E_{ij}$. This reinforces our discussion of the positive correlation between ϵ_{Ai} and ϵ_{ij}.

Different types of home energy efficiency investments entail substantially different unobserved net benefits. Appendix Figure A.8 presents the estimated $\hat{\xi}_j$ from a regression that is identical to the investment takeup probit in column 1 of Appendix Table A.9, except that it allows the constant term to vary by investment category. The close correspondence between takeup rates (against the right axis) and $\hat{\xi}_j$ (against the left axis) emphasizes that takeup rates identify $\hat{\xi}_j$: if many consumers adopt an investment after being informed about its costs and benefits, we infer that they must derive more utility from it. Air sealing, insulation, and programmable thermostats have the most positive $\hat{\xi}_j$, while windows and pipe and duct sealing/insulation have the most negative.

Figure A.7: **Private Internal Rate of Return and Investment Takeup in Wisconsin**
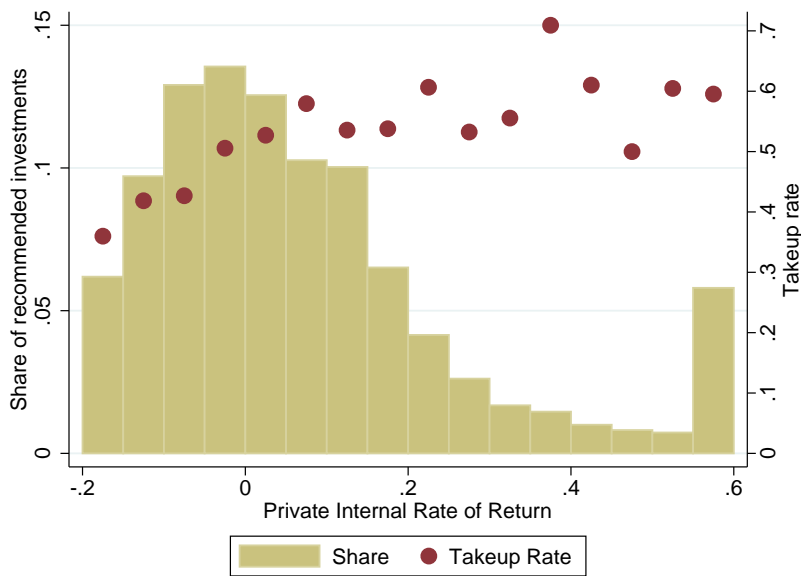
Notes: Internal rates of return are computed at the 2011-2014 average retail natural gas and electricity prices in the household's city (Madison or Milwaukee). Subsidized costs are the upfront cost net of an imputed subsidy of $500 per 10 percentage point projected decrease in the household's pre-audit energy use. These are annualized IRRs, assuming that energy savings accrue in equal monthly installments until the end of the investment's projected life.

Table A.9: **Post-Audit Investment Takeup Estimates**

	(1)	(2)	(3)	(4)	(5)	(6)
Subsidy - Cost + Savings ($000s)	0.059	0.059		0.028	0.059	0.059
	(0.006)***	(0.006)***		(0.007)***	(0.006)***	(0.006)***
Constant	0.019	0.132	0.134	0.111	0.133	0.151
	(0.012)	(0.056)**	(0.056)**	(0.056)**	(0.056)**	(0.056)***
Energy Savings PDV ($000s)			0.056			
			(0.007)***			
Subsidy - Cost ($000s)			0.062			
			(0.007)***			
Closed account					0.026	
					(0.083)	
Experiment audit subsidy ($00s)						-0.129
						(0.046)***
N	5,990	5,990	5,990	6,100	5,990	5,990
Household covariates	No	Yes	Yes	Yes	Yes	Yes
Exclude outliers	Yes	Yes	Yes	No	Yes	Yes

Notes: This table presents marginal effects estimates of Equation (9), a probit model using the sample of recommended investments in the Wisconsin experiment. All columns except for column 4 exclude outliers with net private monetary benefits larger than $5,000 in absolute value. The additional variable in column 5, Closed account, is an indicator for the 124 investments at 37 households that closed their gas and electric accounts more than six months after the audit but before our data end. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Figure A.8: **Mean Unobserved Attribute of Recommended Investments**
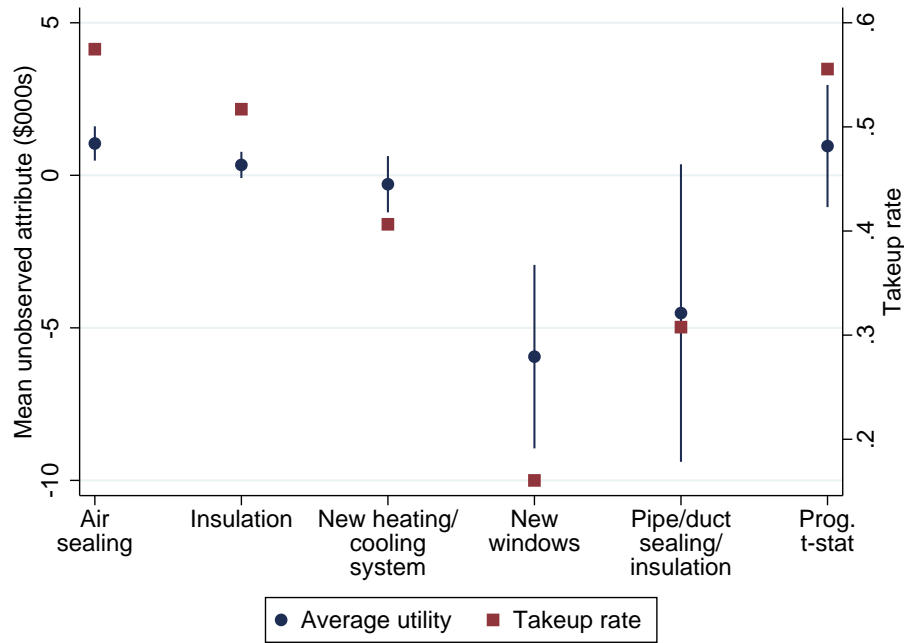
Notes: Point estimates and 90 percent confidence intervals for the dollar value of the mean unobserved attribute by investment category, or ξ_j in the notation of the model. This uses estimates in column 1 of Appendix Table A.9. Standard errors are calculated using the Delta method.

D   Appendix to Estimates of Effects on Energy Use

D.A   Additional Figures

Figure 4 combines event time estimates for natural gas and electricity. In this appendix, we present the separate estimates for each fuel. To parallel Appendix Table A.11 below, we also consider using the first investment date as the event, instead of the audit. Note that even in the post-investment figures, simulation model savings continue increasing (in absolute value) after month 0 because that month marks the *first* investment, which for many households is a "direct install" measure put in during the audit, and many of these same households later make larger investments through a contractor.

Appendix Figures A.9 and A.10 present the natural gas and electricity estimates underlying Figure 4. In each figure, Panel (a) presents the estimates relative to the audit date, while Panel (b) presents the estimates relative to the first investment date. There are no significant pre-audit or pre-investment trends in either gas or electricity use. Immediately after the audit or investment, energy use decreases. By the end of the first year post-audit, natural gas usage averages 0.09 therms/day lower, and electricity usage averages about 1.6 kWh/day lower. Relative to pre-audit averages of 2.40 therms/day and 21.4 kWh/day, these represent approximately four and seven percent reductions.

Relative to the empirical estimates, the simulations predict much larger natural gas savings, both post-audit and post-investment. By contrast, the empirical estimates for post-audit electricity savings are much larger than predicted, and the post-investment estimates are noisy and indistinguishable from the prediction. Note that although $E_{it} = 0$ for all t before the audit, the coefficient estimates are not exactly zero before the audit and investment due to the inclusion of month-of-sample effects μ_m.

Appendix Figures A.11 and A.12 present analogues to Appendix Figures A.9 and A.10, except limiting the sample to a "balanced" set of households whose energy bills are observed over the entire event window of 18 months before to 18 months after. (For the investment figures, households that made no investments are also included as controls.) This eliminates the possibility of compositional effects, i.e. that systematically different sets of households would identify treatment effects for different months. The "balanced" group of households turns out to have made larger investments: on average, they were projected to save 17 percent more gas and 36 percent more electricity than the non-balanced group by the end of the first year post-audit. Thus, the energy savings in Figures A.11 and A.12 are somewhat larger than for the full sample. Otherwise, the figures look very similar.

Appendix Figures A.13 and A.14 present analogues to Figures A.11 and A.12, except over a shorter time window. Because these figures limit to the "balanced" set of households observed over the entire window in order to eliminate compositional effects, limiting to a smaller window allows us to include additional households that are observed only over a shorter period. This and other time windows all give essentially the same pictures as Figures A.11 and A.12.

Appendix Figure A.15 presents the audit event studies for households with no recorded investments. Again, there are no discernible pre-trends. The point estimates suggest a slight but imprecisely estimated post-audit increase in natural gas use.

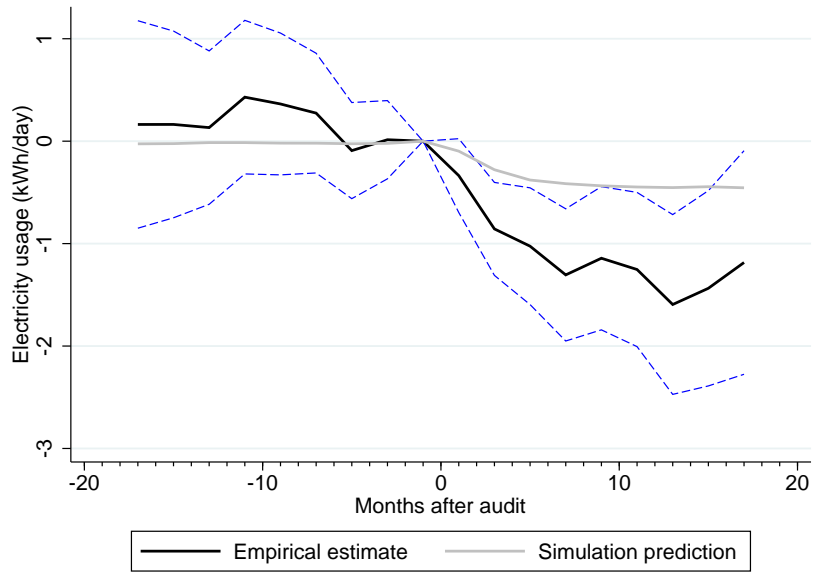Figure A.9: **Natural Gas Usage in Event Time**
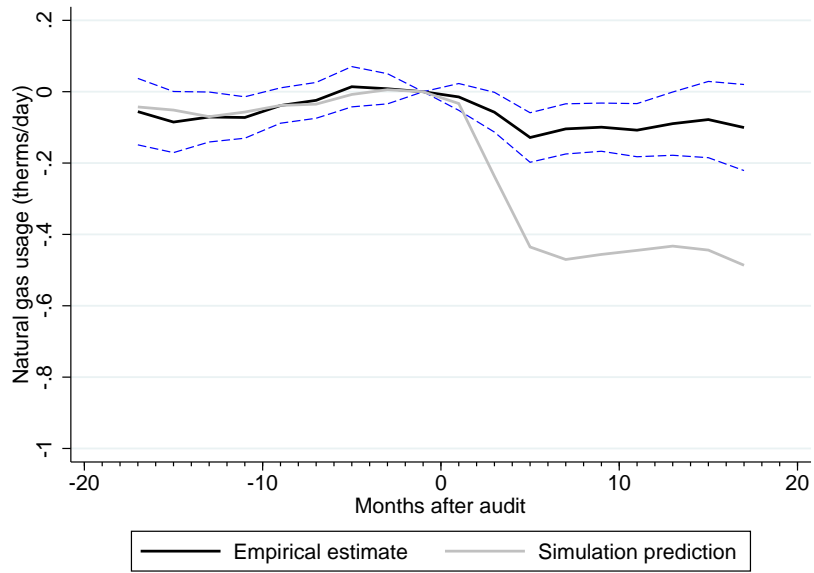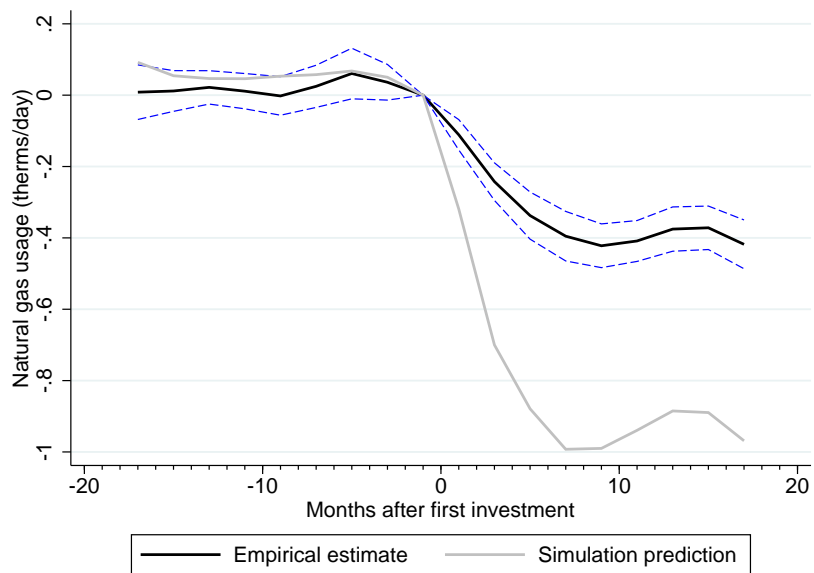
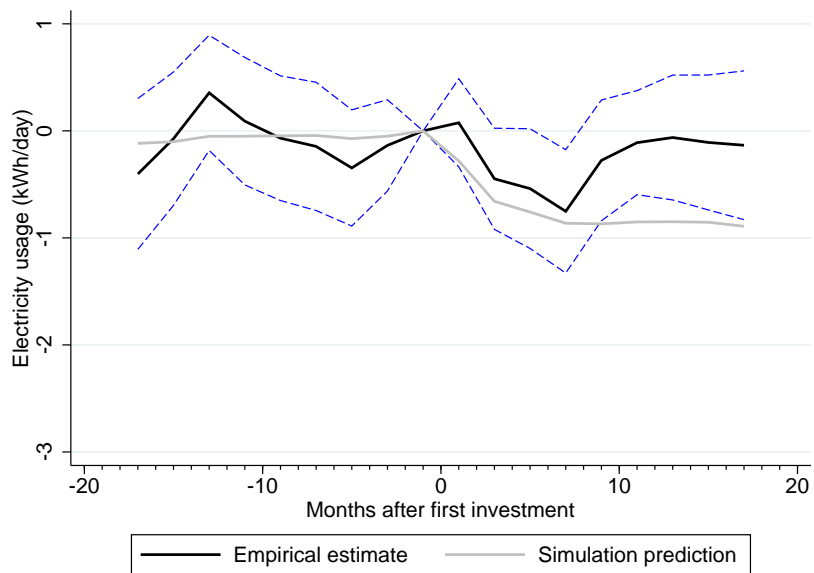(a) **Post-Audit**

(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit natural gas usage is 2.40 therms/day.
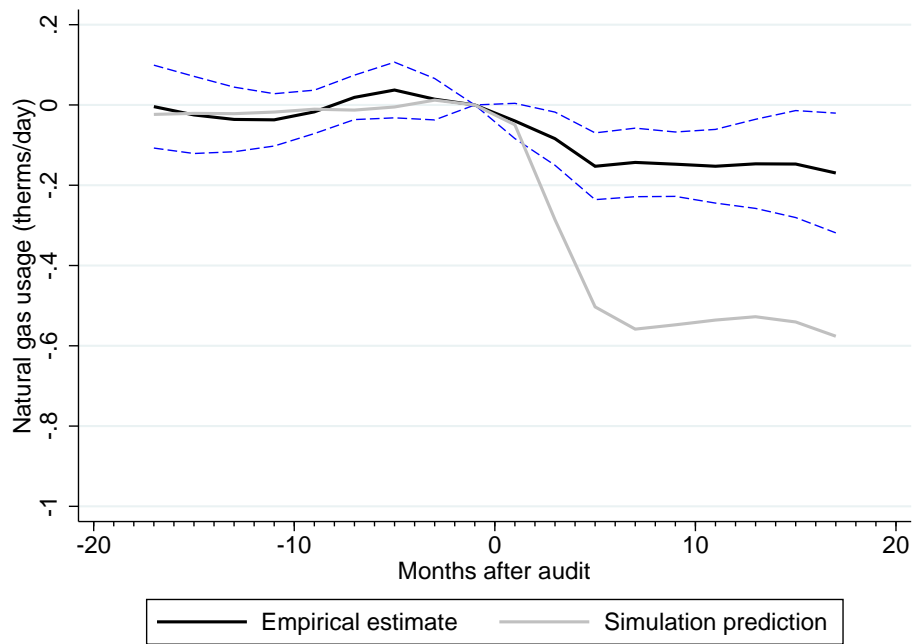
Figure A.10: **Electricity Usage in Event Time**

(a) **Post-Audit**

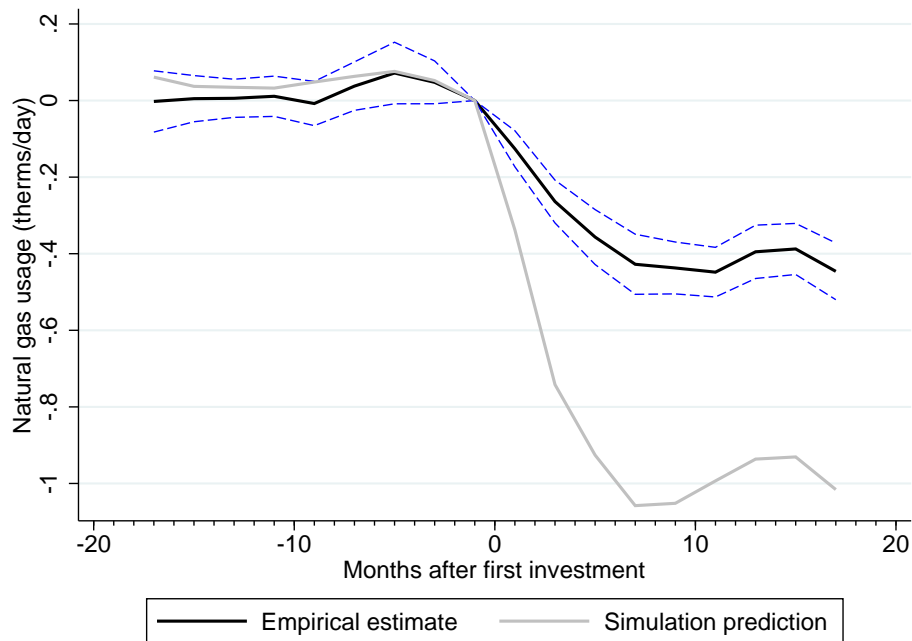(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit electricity usage is 21.4 kWh/day.

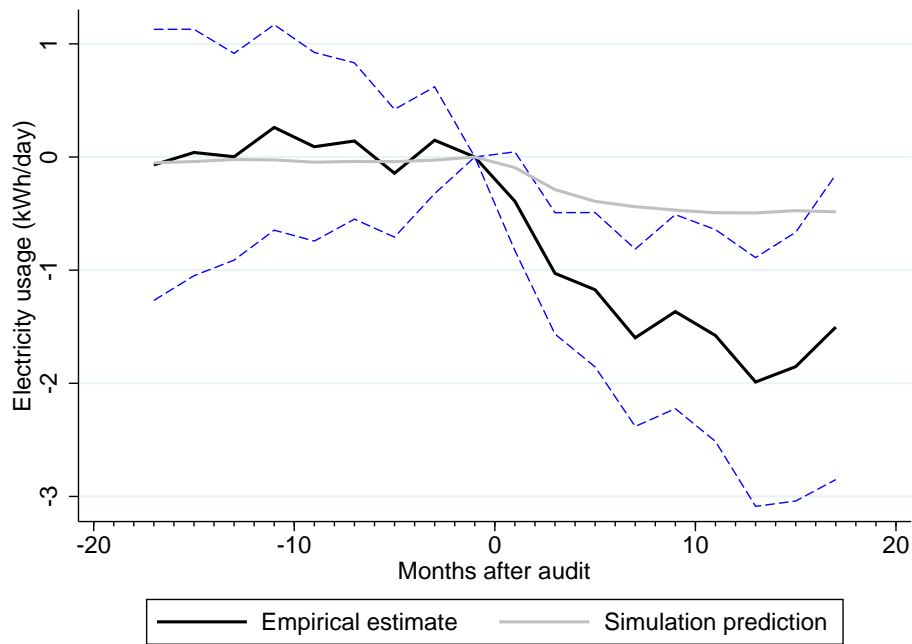Figure A.11: **Natural Gas Usage in Event Time in Balanced Sample**

(a) **Post-Audit**

(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit natural gas usage is 2.40 therms/day. This parallels Figure A.9 but includes on the "balanced" sample of households.

28

Figure A.12: **Electricity Usage in Event Time in Balanced Sample**
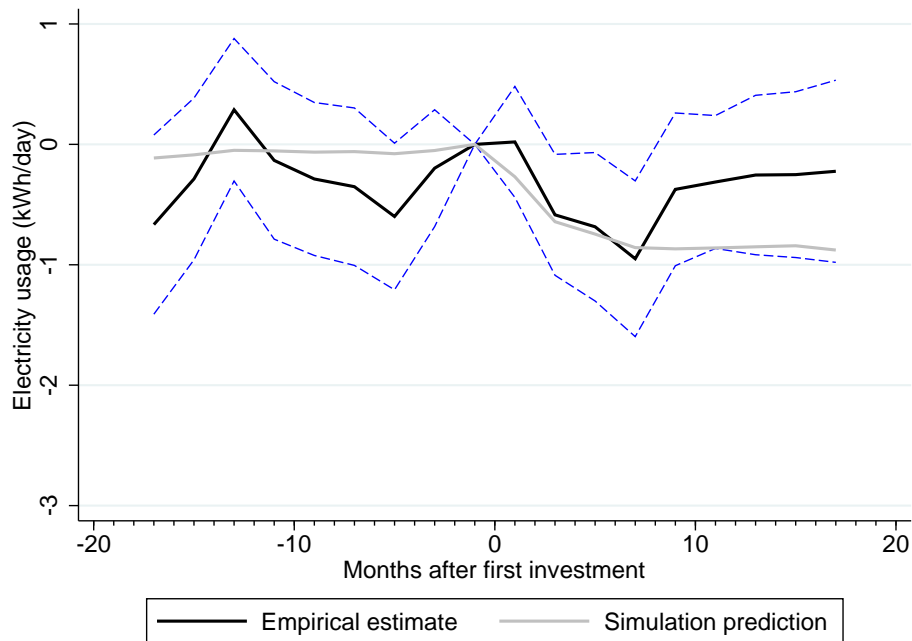

(a) **Post-Audit**

(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit electricity usage is 21.4 kWh/day. This parallels Figure A.10 but includes on the "balanced" sample of households.

Figure A.13: **Natural Gas Usage in Event Time in Balanced Sample over Shorter Window**
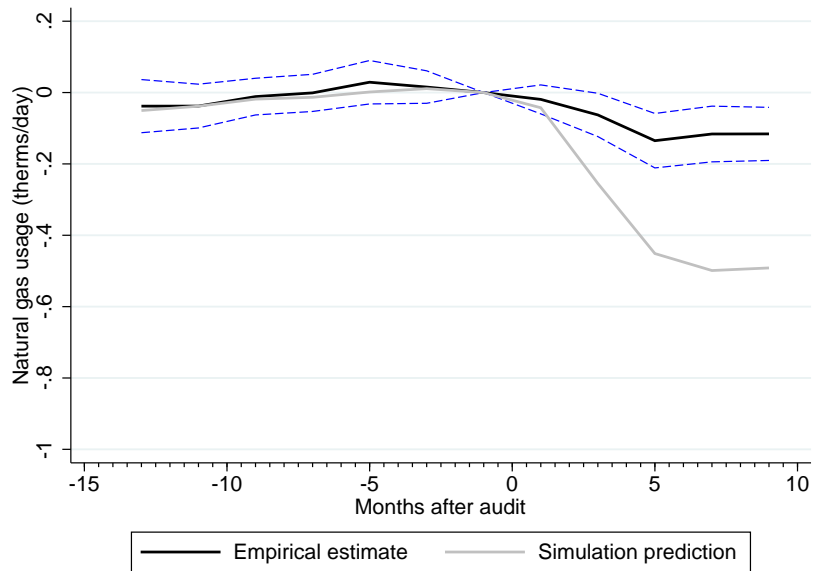

(a) **Post-Audit**

(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit natural gas usage is 2.40 therms/day. This parallels Figure A.11 but covers a shorter time window.

Figure A.14: **Electricity Usage in Event Time in Balanced Sample over Shorter Window**
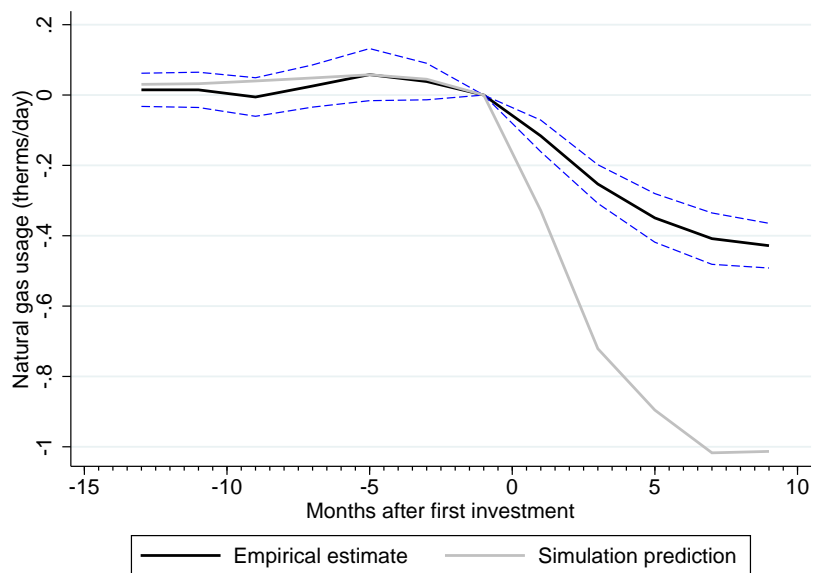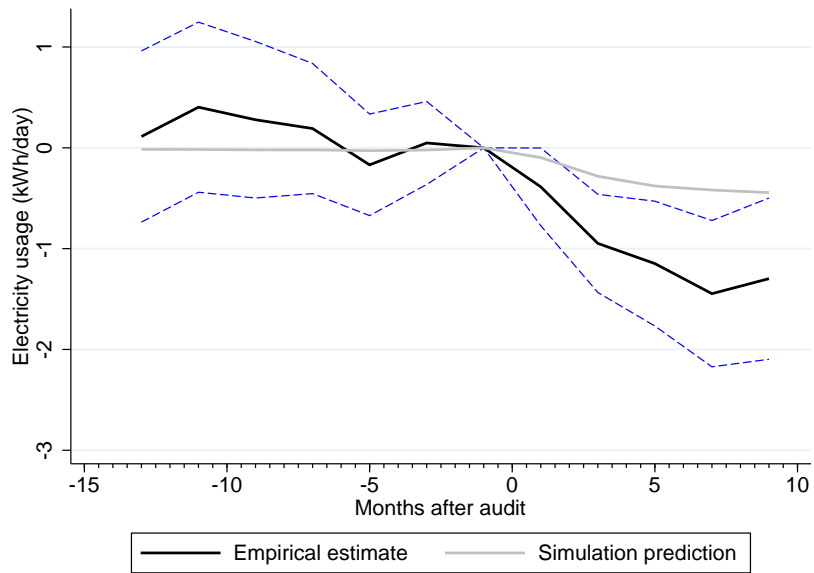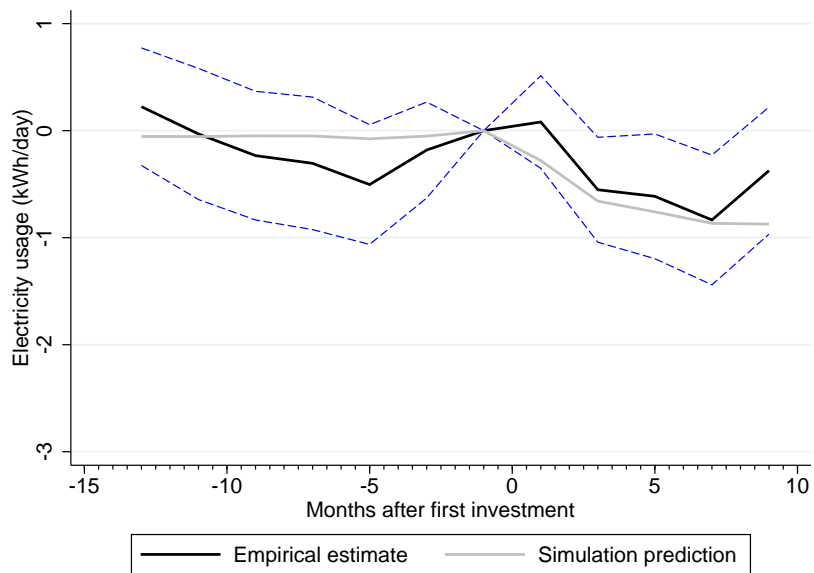

(a) **Post-Audit**

(b) **Post-Investment**

Notes: This figure presents energy use in event time relative to the household's audit or investment. Dashed lines are 90 percent confidence intervals. Mean pre-audit electricity usage is 21.4 kWh/day. This parallels Figure A.12 but covers a shorter time window.

Figure A.15: **Energy Use in Event Time for Households with No Recorded Investments**
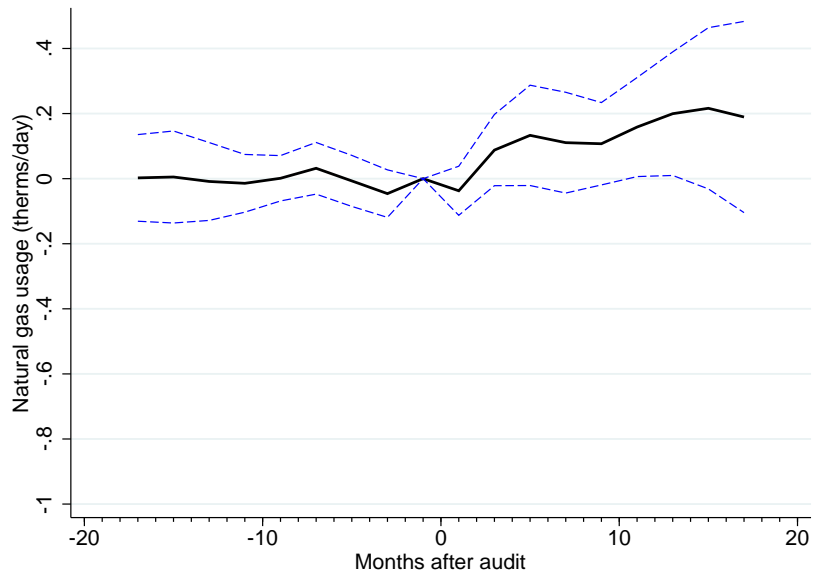
(a) **Natural Gas**
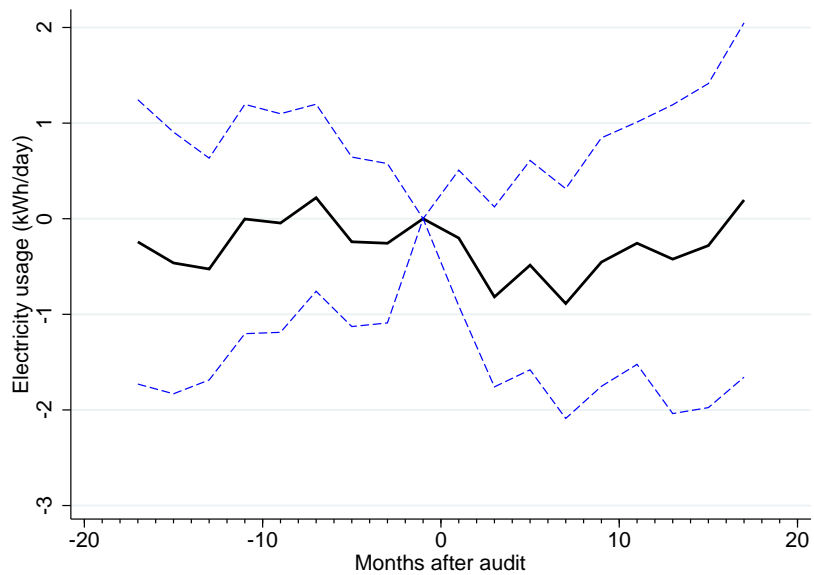
(b) **Electricity**

Notes: This figure presents energy use in event time relative to the household's audit, for the subsample of households that made no recorded investments. Dashed lines are 90 percent confidence intervals. Mean pre-audit natural gas usage is 2.40 therms/day, and mean pre-audit electricity usage is 21.4 kWh/day.

D.B   Extensions to Table 5

Appendix Table A.10 presents robustness checks and extensions to Table 5. As with Table 5, columns 1 and 2 present natural gas estimates, while columns 3 and 4 present electricity estimates. Columns 1 and 3 limit the sample to households that made no observed investments, paralleling Appendix Figure A.15. Three of the four coefficients on the post audit indicators are statistically insignificant, which would suggest that these households do not make significant unobserved investments. For natural gas, however, the coefficient on Post audit ($\geq$6 months) is positive and significant with 90 percent confidence.[33]

Columns 2 and 4 repeat columns 1 and 3 of Table 5, except with weather-unadjusted simulation estimates as the dependent variable. The estimates are similar to but slightly smaller than the estimates using weather-adjusted predictions. This implies that the post-audit and post-investment samples include slightly more extreme temperatures, in which more savings would be expected relative to prediction. The ≥ 6 month gas realization rate using the weather-unadjusted results in column 3 is -0.128/-0.409$\approx$31 percent. For comparison, the weather-adjusted figure reported in Figure 5 was 29 percent.

In the body of the paper, we defined $\mathbf{P}_{it}$ as a pair of post-audit indicators, estimating the effects of audits plus subsequent investments. Appendix Table A.11 parallels Table 5, except adding post-investment indicators to $\mathbf{P}_{it}$. Recall that E_{it} takes value 0 before investments are made and equals the predicted savings thereafter. Many households audited but did not invest, so the post-audit indicators from Table 5 represent a diluted effect of investments in the larger population of households that audited. In Appendix Table A.11, the predicted post-investment savings now load more heavily onto the post-investment indicators. There are small decreases associated with the post-audit indicators in column 3 due to the combination of minor compositional effects with month of sample dummies μ_m.

These results can be used to construct an alternative realization rate based on the post-investment coefficients. The natural gas realization rate is 0.458/1.036$\approx$44%, while the electricity realization rate is 0.199/0.836$\approx$24%. When combined at retail prices, this gives an over realization rate of 42 percent.

The main interesting feature of this table is that the electricity savings in column 4 are much more strongly associated with the post-audit indicators than the post-investment indicators. This suggests that the excess electricity savings may not be attributable to the observed investments, which in any event are largely projected to save natural gas. Instead, the excess savings could result from unobserved actions taken after the audit, such as additional CFL installation and appliance replacement, that would not be recorded in the programs' administrative data. When adjusting the accounting welfare analysis in Section VIII for empirically realized savings, we assume that the excess electricity savings accrue because of the *observed* investments, which may cause us to understate total investment costs for the empirically estimated energy savings. This biases against our finding that program costs outweighed benefits.

To better understand the differences between the simulation and empirical estimates, we construct a "shortfall" variable $Y_{it} - E_{it}$, measuring the difference between actual usage and predicted savings. Columns 1 and 3 of Appendix Table A.12 regress this shortfall on the interaction of post-audit indicators with indicators for whether the household made made one of the three most

---

[33]Additional regressions show that this apparent post-audit gas use increase is unlikely to be driven by unobserved energy efficiency investments: the increase is no larger at households that received recommendations of new heating and cooling systems or water heaters. The increase is unlikely to be driven by substitution from electricity to gas: post-audit gas use changes are not associated with electricity use changes. We do find that the apparent usage increase is limited to only above-median natural gas users. Our best guess is that the apparent increase is an idiosyncratic result for a subset of heavy users where the difference-in-differences controls are imperfect.

common investments: insulation, heating/cooling, and "other," which includes air sealing, new water heaters, new windows, pipe and duct sealing and insulation, and programmable thermostats. Positive coefficients mean that the investment is associated with a savings shortfall, while negative coefficients mean that the investment is associated with excess savings. Column 1 shows that insulation and heating/cooling are not statistically significantly associated with the shortfall variable, although the t-statistic on insulation is around 1.3. By contrast, natural gas shortfalls are statistically significantly associated with investments in the "other" category. (We do not have sufficient power to further disaggregate the "other" category.) Consistent with our results, a recent report by the TREAT model developers (PSD 2015b) also finds that air sealing and insulation predictions are important contributors to low realization rates. The TREAT model is "in close alignment with the predictions from best-in-class modeling tools" (PSD 2015b), so this finding may be more general than just TREAT.

Columns 2 and 4 add an indicator for whether a new appliance appears in the household's list of recommended measures. The program's takeup data do not include appliances because they were not subsidized, but the fact that this last interaction is significant suggests that these households may have reduced both natural gas and electricity use through unobserved appliance purchases. However, only 25 households in the energy use data were recommended appliances, so this is not enough to fully explain the excess post-audit electricity savings relative to the simulation predictions.

See Blasnik (2010), Nadel and Keating (1991), and PSD (2015b) for more in-depth discussions of why simulations can systematically overestimate empirically realized savings. One broad class of explanations has to do with the assumptions and parameters used by the program and its auditors. For example, a home's baseline energy use could be overstated, giving excess simulated savings. This explains most of the empirical shortfall in the PSD (2015b) New York study. In the Wisconsin programs, however, only 12 percent of pre-audit observations were calibrated with estimates instead of the audited household's true baseline energy use, so this particular factor seems unlikely to explain much of the our estimated shortfall.[34] Another factor that could cause excess simulated savings in Wisconsin and elsewhere is that simulations often assume best-case installation scenarios that could not be realized except through great effort by the most expert contractors. This, or any other source of systematic bias, can be addressed (on average) by scaling back all predictions by some factor less than one.

---

[34]The default calibrations were 23 percent lower and 12 percent higher than the average audited household's pre-audit usage of gas and electricity, respectively.

Table A.10: **Post-Audit Energy Use Changes: Alternative Estimates**

	(1)	(2)	(3)	(4)
	Natural Gas (therms/day)		Electricity (kWh/day)	
	Non-Investors' Energy Use	Weather-Unadjusted Simulation Prediction	Non-Investors' Energy Use	Weather Unadjusted Simulation Prediction
Dependent Variable:				
Post audit (<6 months)	0.068	-0.225	-0.416	-0.292
	(0.043)	(0.021)***	(0.395)	(0.025)***
Post audit (≥6 months)	0.112	-0.409	-0.145	-0.402
	(0.057)*	(0.029)***	(0.524)	(0.034)***
N	22,455	61,845	23,098	63,655

Notes: This table presents alternative estimates of Equation (12). Columns 1 and 3 present estimates for households with no recorded investments. Columns 2 and 4 present estimates of Equation (12) with weather-unadjusted simulation predictions as the dependent variable. Mean pre-audit natural gas usage is 2.40 therms/day, and mean pre-audit electricity usage is 21.4 kWh/day. Average marginal natural gas price is \$0.82 per therm, and average marginal electricity price is \$0.136 per kWh. All columns control for heating and cooling degrees, household-by-calendar month fixed effects, and month-of-sample fixed effects. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A.11: **Post-Investment Energy Use Changes**

Dependent Variable:	(1) Natural Gas (therms/day)	(2)	(3) Electricity (kWh/day)	(4)
	Simulation Prediction	Energy Use	Simulation Prediction	Energy Use
Post audit (<6 months)	0.010	0.016	-0.082	-0.796
	(0.021)	(0.030)	(0.019)***	(0.259)***
Post audit (≥6 months)	0.013	0.067	-0.045	-0.918
	(0.022)	(0.037)*	(0.023)**	(0.340)***
Post investment (<6 months)	-0.849	-0.323	-0.626	-0.217
	(0.030)***	(0.030)***	(0.037)***	(0.293)
Post investment (≥6 months)	-1.036	-0.458	-0.836	-0.199
	(0.031)***	(0.037)***	(0.052)***	(0.345)
N	61,845	61,845	63,655	63,654

Notes: This table presents estimates of Equation (12) with daily usage of natural gas and electricity, respectively, as the dependent variables, using data from the Wisconsin experiment. This parallels Table 5, except adds post-investment indicators. Columns 1 and 3 have the simulation predictions as the dependent variable, while columns 2 and 4 have energy use as the dependent variable. Mean pre-audit natural gas usage is 2.40 therms/day, and mean pre-audit electricity usage is 21.4 kWh/day. Average marginal natural gas price is \$0.82 per therm, and average marginal electricity price is \$0.136 per kWh. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A.12: **Associations of Energy Savings Shortfalls with Specific Investments**

	(1) Natural Gas (therms/day)	(2)	(3) Electricity (kWh/day)	(4)
Post audit × Insulation	0.179	0.158	-0.997	-1.265
	(0.140)	(0.138)	(1.454)	(1.356)
Post audit × Heating/cooling	-0.003	0.004	-0.518	-0.432
	(0.064)	(0.064)	(0.539)	(0.534)
Post audit × Other	0.348	0.367	1.373	1.607
	(0.142)**	(0.140)***	(1.489)	(1.393)
Post audit × Recommended appliance		-0.219		-3.255
		(0.069)***		(0.879)***
N	61,845	61,845	63,654	63,654

Notes: This table presents alternative estimates of Equation (12), interacting P_{it} with indicators for installed and recommended investment categories. The dependent variable is $Y_{it} - E_{it}$, the difference between actual usage and predicted savings. Mean pre-audit natural gas usage is 2.40 therms/day, and mean pre-audit electricity usage is 21.4 kWh/day. Average marginal natural gas price is \$0.82 per therm, and average marginal electricity price is \$0.136 per kWh. All columns control for heating and cooling degrees, household-by-calendar month fixed effects, and month-of-sample fixed effects. Robust standard errors in parentheses, clustered by household. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

D.C   Can the "Rebound Effect" Explain the Empirical Shortfall?

We present two calculations to test whether the "rebound effect," i.e. increased utilization in response to a decreased cost of energy services, can explain the low natural gas realization rate. For this section, we use the realization rate implied by the post-investment coefficients in Table A.11, which is $0.458/1.036 \approx 44\%$. Since this is slightly higher than the realization rate from the post-audit coefficients, our calculations are "conservative": it would be even more difficult for the rebound effect to explain the lower post-audit realization rate.

First, we calculate the price elasticity of demand for energy services that would be implied if all of the empirical shortfall is due to increased utilization. The change in quantity of energy services would be

$$\%\Delta q_{energy\ services} = \frac{\text{Actual Post-Investment Usage}}{\text{Simulated Post-Investment Usage}} - 1 \approx 0.419. \tag{15}$$

The change in price of energy services is

$$\%\Delta p_{energy\ services} = \frac{\text{Simulated Usage Reduction}}{\text{Baseline Usage}} \approx -0.429. \tag{16}$$

The implied elasticity is thus

$$\frac{\%\Delta q_{energy\ services}}{\%\Delta p_{energy\ services}} \approx -0.98. \tag{17}$$

A utilization elasticity of -0.98 appears to be unrealistically large in this population. For comparison, the most closely related estimates of energy utilization elasticity are much smaller: -0.06, -0.3, and -0.22 for washing machines (Davis 2008), home electricity (Dubin and McFadden 1984), and autos (Gillingham 2014), respectively.

The second calculation is to determine the change in indoor temperature that would be required to generate the entire empirical shortfall. To do this, we make two assumptions. First, we assume that natural gas consumption is a linear function of the difference between indoor and outdoor temperatures. Second, we assume that households keep indoor temperatures constant as outdoor temperatures fluctuate, or more weakly that indoor temperatures are not conditionally correlated with outdoor temperatures. Under these two assumptions, we can estimate the natural gas required to increase *indoor* temperatures using variation in *outdoor* temperatures. We estimate the following analogue to Equation (12):

$$Y_{it} = \iota P_{it}\mathbf{W}_{it} + \omega(1 - P_{it})\mathbf{W}_{it} + \nu_i + \mu_m + \varepsilon_{it}. \tag{18}$$

where Y_{it} is natural gas consumption, P_{it} is a post-investment indicator, $\mathbf{W}_{it}$ is heating and cooling degrees, ν_i is a household-by-calendar month fixed effect, and μ_m are month-of-sample indicators. The coefficient ι measures natural gas consumption as a function of the difference between indoor and outdoor temperatures.

In the post-investment sample, we estimate $\hat\iota \approx 0.035$ therms/cooling degree and $\hat\iota \approx 0.099$ therms/heating degree. If 1/3 of the year in Wisconsin requires cooling and 2/3 of the year requires heating, then the average therms per degree of indoor temperature per day is $\frac{1}{3} \cdot 0.035 + \frac{2}{3} \cdot 0.099 \approx 0.078$. Comparing the ≥ 6 month post-investment simulation vs. empirical estimates in columns 2 vs. 4 of Table 5 gives an empirical shortfall of $1.036 - 0.458 \approx 0.578$ therms/day. Thus, to explain the full empirical shortfall, indoor temperatures would have to change (increase in winter and decrease in summer) by $0.589/0.078 \approx 7.5$ degrees. This would be unrealistically large.

For comparison, Fowlie, Greenstone, and Wolfram (2015) find no statistically significant post-weatherization temperature change in a low-income (and thus likely more price elastic) population in Michigan. They can reject a post-weatherization temperature change of more than 1.4 degrees with 90 percent confidence in a two-sided test.

   In summary, the "rebound effect" is likely to explain only a small fraction of the empirical shortfall for natural gas.

E   Accounting Welfare Analysis of National Better Buildings Program

It is important to know whether our results in Table 7 are specific to the Wisconsin programs, or whether they generalize more broadly. To test whether the Wisconsin experimental sample was somehow unusual, we augment the analysis with microdata from all Better Buildings programs nationwide, representing $435 million in upfront costs across 58,418 households at 37 sites. We downloaded the data from the DOE website and prepared the data according to the process described in Appendix B.B.

Table A.13 describes data from the 37 Better Buildings Neighborhood Program sites that reported to the Department of Energy. These data are more limited than for the Wisconsin programs: for each household retrofitted through the program, we observe the categories of investments made (as in Wisconsin, typically insulation, air sealing, and heating and cooling), total (unsubsidized) cost, and simulation estimates of total annual energy savings in physical units. The primary two fuels saved are electricity and natural gas, but the data also include predicted savings of heating oil, propane, kerosene, and wood. We group these latter four fuels as "other fuels" for Table A.14. We translate physical units to dollars and externality reductions using the same process described in Section IV, with data from the state and electricity market where the site is located. Of the 75,110 households in the original data, 58,418 survive the data cleaning process described in Appendix B.

Table 7 presents welfare effects using the "accounting approach." Columns 1 and 2 replicate Table 7, presenting results for the Wisconsin programs, while columns 3 and 4 present the nationwide BBNP results.[35] Columns 1 and 3 use the simulation predictions to calculate energy and externality reductions. Column 2 multiplies the Wisconsin simulation estimates by realization rates of 2.48 for electricity and 0.29 for natural gas and other fuels, on the basis of Table 5 and Figure 5. Column 4 multiplies the national program simulation estimates by realization rates of 0.59 for electricity and 0.47 for natural gas and other fuels, on the basis of the DOE's (2015c) national BBNP evaluation.

Under the assumptions of the accounting approach, the programs reduce welfare. Even using the simulation predictions, the Wisconsin and national programs have benefit/cost ratios of 0.92 and 0.75 at five percent discount rates, with internal rates of return of 4.0 and 1.5 percent, respectively. After adjusting for the empirical shortfall, the benefit/cost ratios are 0.43 and 0.38, and the IRRs are negative 4.1 and negative 5.2 percent. The projects from the national program with valid data involved total cost of $435 million, compared to empirically adjusted social benefits of $166 million.

Of course, the results of the engineering approach depend on energy price and externality assumptions. For empirically adjusted benefits to exceed costs, energy acquisition costs would need to be 4.2 and 3.9 times larger, all externality damages would need to be 3.3 and 4.7 times larger, or the social cost of carbon would need to be $239 and $301 per ton for the Wisconsin and national samples, respectively.

The U.S. Department of Energy (DOE) has conducted an official evaluation of the nationwide Better Buildings Neighborhood Program (DOE 2015c). The DOE report evaluates BBNP as an economic stimulus program, which is relevant given that it was funded through the American Recovery and Reinvestment Act, but does not present a benefit-cost analysis relevant for typical macroeconomic conditions. By contrast, we do not consider the economic stimulus benefit, so while our analysis is appropriate under normal economic conditions, we are not evaluating whether the

---

[35]We assume that the unsubsidized audit cost is $c_A = \$400$, based on typical market prices. The national Better Buildings program dashboard reports a grand total of 138,323 single-family home audits and 74,493 retrofits. Applying that 54 percent follow-through rate implies that 108,474 audits were required to generate the 58,418 valid projects in our data.

programs increased welfare during the Great Recession. The DOE's headline benefit/cost ratio of 3.0 is the ratio of economic activity created (i.e. private sector expenditures plus tax revenues) divided by federal government outlays. This calculation is not comparable to ours. Indeed, it is mechanically opposite: energy efficiency investment costs count as a cost in our framework, while they count as a benefit in the DOE framework because they represent additional economic stimulus. The details of the DOE report suggest that if they had done an analysis like ours, they would have arrived at similar conclusions: total program costs exceed retail energy savings plus carbon emission benefits, and their empirical analysis finds a 52 percent realization rate for residential retrofits after combining electricity and gas.[36]

Table A.13: **Summary Statistics for National Better Buildings Program Data**

Variable	Mean	Std. Dev.	Min.	Max.
Total cost ($)	6705	5711	100	30,000
Retail energy cost savings ($/year)	472	430	12.7	2734

Notes: Sample size is the 58,418 households that made investments and survive the data cleaning process described in Appendix B. Energy prices are averages over 2011-2014; see Appendix B for details.

---

[36]The DOE evaluation reports $669 million in lifetime bill savings and $972 million in total costs, plus 7.2 million metric tons of CO2 abated, which sums to $274 million at a $39 social cost of carbon. In this sense, costs exceed benefits by $29.4 million. However, this is not directly comparable to our benefit-cost analysis, because it considers energy savings at retail prices instead of avoided costs, uses un-discounted future energy and environmental benefits, and does not include the local air pollution benefits.

Table A.14: **Welfare Effects: Accounting Approach, Including National Programs**

	(1)	(2)	(3)	(4)
	Wisconsin Programs		*National Programs*	
	Simulation	Empirically	Simulation	Empirically
Source of energy savings estimates:	Predictions	Adjusted	Predictions	Adjusted
Cost ($millions)				
Audit costs (at $400 per audit)	0.56		43	
Investment costs	4.52		392	
Total cost	5.08		435	
Energy Savings ($millions present value at 5% discount rate)				
Natural gas	1.92	0.56	77	36
Electricity	0.07	0.18	43	25
Other fuels	0.61	0.18	68	32
Total	2.61	0.92	187	93
Externality Reduction ($millions present value at 5% discount rate)				
Climate (at $39 per ton CO2)	1.37	0.57	81	40
SO2/NOx/PM	0.72	0.68	59	32
Total	2.09	1.26	140	73
Summary				
Benefits - Costs ($millions)	-0.39	-2.91	-108	-270
Benefit/Cost ratio	0.92	0.43	0.75	0.38
Internal rate of return (percent)	4.0	-4.1	1.5	-5.2

Notes: Columns 1 and 2 present estimates for the Wisconsin sample, and Columns 3 and 4 present estimates for the 58,418 households that invested under the national Better Buildings program and have valid data. Columns 1 and 3 use energy savings projected by simulation models, while columns 2 and 4 adjust for empirically observed savings. Column 2 multiplies electricity and gas/other fuels savings from Column 1 by 2.48 and 0.29, respectively, based on the estimates in Table 5. Column 4 multiplies electricity and gas/other fuels savings from Column 3 by 0.59 and 0.47, respectively, based on estimates from DOE (2015c). Energy savings are calculated at average wholesale prices over 2011-2014, and externality reductions are based on a $39 social cost of carbon and a $6 million value of a statistical life; see Appendix B for details.

F   Appendix to Revealed Preference Welfare Analysis

Table A.15: **Association Between Recommended Investment Characteristics and Audit Subsidy**

	(1)	(2)	(3)	(4)
	Subsidy - Cost + Savings ($000s)	Energy Savings PDV ($000s)	Subsidy - Cost ($000s)	Number of Recommendations
Experiment audit subsidy ($00s)	-0.083	-0.014	-0.069	0.136
	(0.152)	(0.114)	(0.109)	(0.184)
N	6,100	6,100	6,100	1,394
Household covariates	Yes	Yes	Yes	Yes

Notes: This table presents regressions of characteristics of recommended investments on the household's experimental audit subsidy. The sample for columns 1-3 is all recommended investments; the sample for column 4 is all households that audited. Robust standard errors in parentheses. Standard errors are clustered by household in columns 1-3. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

References

[1] Abrahams, Leslie S., Constantine Samaras, W. Michael Griffin, and H. Scott Matthews. 2015. "Life Cycle Greenhouse Gas Emissions From U.S. Liquefied Natural Gas Exports: Implications for End Uses." *Environmental Science and Technology* 49: 3237-3245.

[2] Blasnik, Michael (2010). "Energy Models vs. Reality." Presentation, Southface Lecture Series (January).

[3] EPA (U.S. Environmental Protection Agency) (1995). "AP 42, Ffth Edition: Compilation of Air Pollutant Emission Factors, Volume 1: Stationary Point and Area Sources. Chapter 1: External Combustion Sources." Available from http://www.epa.gov/ttn/chief/ap42/ch01/index.html. Accessed June 12, 2015.

[4] Heaney, Michael, and Ben Polly (2015). "Analysis of Installed Measures and Energy Savings for Single-Family Residential Better Buildings Projects." National Renewable Energy Laboratory Technical Report 5500-64091 (April).

[5] Holland, Stephen, Erin Mansur, Nicholas Muehller, and Andrew Yates (2015). "Measuring the Spatial Heterogeneity in Environmental Externalities from Driving: A Comparison of Gasoline and Electric Vehicles." NBER Working Paper No. 21291 (June).

[6] Howarth, Robert, Drew Shindell, Renee Santoro, Anthony Ingraffea, Nathan Phillips, and Amy Townsend-Small. (2012). "Methane Emissions from Natural Gas Systems." Available from http://www.eeb.cornell.edu/howarth/publications/Howarth_et_al_2012_National_Climate_Assessment.pdf. Accessed August 15, 2015.

[7] Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura and H. Zhang. 2013. "Anthropogenic and Natural Radiative Forcing." In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[8] Nadel, Steven, and Kenneth Keating (1991). "Engineering Estimates vs. Impact Evaluation Results: How Do They Compare and Why?" Research Report U915, American Council for an Energy-Efficient Economy. Available from http://www.aceee.org/research-report/u915

[9] New England Fuel Institute (2014). "State Sulfur & Bioheat Requirements for No. 2 Heating Oil in the Northeast & Mid-Atlantic States." Available from http://www.nefi.com/news/docs/heating-oil-standards-chart.pdf. Accessed June 12, 2015.

[10] Pope, C. Arden, Richard T. Burnett, Michael J. Thun, Eugenia, E. Calle, Daniel Krewski, Kazuhiko Ito, and George D. Thurston (2002). "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association*, Vol. 287, No. 9 (March 6), pages 1132-1141.

[11] Potomac Economics (2015b). "2014 State of the Market Report for the ERCOT Wholesale Electricity Markets." Available from https://www.potomaceconomics.com/uploads/ercot_documents/2014_ERCOT_State_of_the_Market_Report.pdf. Accessed August 18, 2015.