



E2e Working Paper 020

Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program

Meredith Fowlie, Michael Greenstone, and Catherine Wolfram

June 2015

This paper is part of the E2e Project Working Paper Series.

E2e is a joint initiative of the Energy Institute at Haas at the University of California, Berkeley, the Center for Energy and Environmental Policy Research (CEEPR) at the Massachusetts Institute of Technology, and the Energy Policy Institute at Chicago, University of Chicago. E2e is supported by a generous grant from The Alfred P. Sloan Foundation.

The views expressed in E2e working papers are those of the authors and do not necessarily reflect the views of the E2e Project. Working papers are circulated for discussion and comment purposes. They have not been peer reviewed.



Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program

Meredith Fowle, Michael Greenstone, and Catherine Wolfram*

June 2015

Abstract

Conventional wisdom suggests that energy efficiency (EE) policies are beneficial because they induce investments that pay for themselves and lead to emissions reductions. However, this belief is primarily based on projections from engineering models. This paper reports on the results of an experimental evaluation of the nation's largest residential EE program conducted on a sample of more than 30,000 households. The findings suggest that the upfront investment costs are about twice the actual energy savings. Further, the model-projected savings are roughly 2.5 times the actual savings. While this might be attributed to the "rebound" effect – when demand for energy end uses increases as a result of greater efficiency – the paper fails to find evidence of significantly higher indoor temperatures at weatherized homes. Even when accounting for the broader societal benefits of energy efficiency investments, the costs still substantially outweigh the benefits; the average rate of return is approximately -9.5% annually.

*We received many helpful comments from seminar participants at Carnegie Mellon, Columbia University, ETH Zurich, National University of Singapore, NBER Summer Institute, Resources for the Future, the University of Basel, the University of Maryland, the University of Michigan, and the University of Wisconsin. The authors gratefully acknowledge the financial support of the Alfred P. Sloan Foundation, the MacArthur Foundation, the Rockefeller Foundation and the UC Berkeley Energy and Climate Institute, and institutional support from the Poverty Action Lab (JPAL) at MIT, the Center for Local, State, and Urban Policy (CLOSUP) at the University of Michigan, and the Energy Institute at Haas. We thank James Gillan, Brian Goggin, Walter Graf, Erica Myers, Daniel Stuart, and Matthew Woerman for excellent research assistance. We are indebted to Jesse Worker for outstanding management of a challenging project. Finally, we thank our contacts at both our partner utility and the community action agencies, without whom this project would not have been possible.

1 Introduction

Energy efficiency investments are widely believed to offer the rare win-win opportunity. Detailed engineering projections, such as those summarized by the well-known McKinsey curves (McKinsey & Company, 2009), routinely project that investments pay for themselves through the energy saved alone (win #1). Moreover, by reducing the energy necessary to achieve a given level of energy services (e.g., indoor heating), these investments promise to decrease the greenhouse gas emissions causing climate change and other pollutants that compromise human health (win #2).

Despite these apparent opportunities, there is a large and persistent difference between the levels of investment in energy efficiency that are projected to save consumers money and the investments that individuals actually pursue. This has become known as the “efficiency gap.” Over the last three decades, a wide variety of explanations have been offered for this apparent failure of consumers to avail themselves of profitable investment opportunities. The most popular explanations have emphasized the possibility of market failures, such as imperfect information, capital market failures, split incentive problems, and behavioral explanations, including myopia, inattentiveness, and prospect theory and reference-point phenomena (see, for example, Allcott and Greenstone, 2012; Gillingham and Palmer, 2014; Gerarden et al., 2015). In contrast, relatively little attention has been paid to the more pedestrian possibility that the real world returns on energy efficiency investments are lower than the engineering models indicate.¹

Mounting concern about climate change has increased the urgency of understanding this phenomenon. Governments around the world are pursuing a wide range of policies designed to narrow or close this gap. For example, the International Energy Administration has outlined a suite of policies that do not harm economic growth and limit warming to the 2 degrees C recommended by climate scientists; in this scenario, end-use energy efficiency improvements account for 49% of the greenhouse gas emissions abatement in 2020 (IEA, 2013).² Further, energy efficiency investments are a critical part of the U.S. government’s proposed Clean Power Plan (ICF, 2014), U.S. electric

¹Early work by Joskow and Marron (1992) raised concerns about overstated efficiency potential and underscored the importance of using ex post measures of consumer behavior to estimate energy savings. Subsequent work uses panel data on household energy expenditures to estimate realized returns on efficiency investments (see, for example, Metcalf and Hassett, 1999; Jacobsen and Kotchen, 2013; Graff Zivin and Novan, 2015).

²Indeed, energy efficiency is a central plank for virtually all serious climate mitigation plans (Loftus et al., 2015).

utilities are rapidly expanding their energy efficiency programs (Barbose et al., 2013), and federal and state regulators routinely tighten energy efficiency building codes, appliance standards, and fuel economy standards for automobiles and trucks.

This paper provides the first large-scale field evidence on the returns to energy efficiency investments from a randomized controlled trial. Specifically, we use experimental, as well as quasi-experimental, variation in participation in the federal Weatherization Assistance Program (WAP), to identify the returns to these investments. WAP is the nation’s largest residential energy efficiency program and has provided over 7 million low-income households with weatherization assistance since its inception in 1976. Recipient households in our study received approximately \$5,150 worth of home improvements on average, at zero out-of-pocket costs. The most common measures included furnace replacement, attic and wall insulation, and infiltration reduction. Importantly, WAP only pays for energy efficiency measures that pass a cost-benefit test, based on ex ante engineering estimates, with the aim of ensuring that only beneficial investments are undertaken.

The randomized encouragement design experiment was conducted on a sample of over 30,000 Michigan households that were presumptively eligible for participation in WAP. Approximately one quarter of these households were randomly assigned to a treatment group that was encouraged to apply for the program and received significant application assistance. The control households were free to apply for WAP but were not contacted or assisted in any way by our team.

There are three primary findings. First, an aggressive encouragement intervention increased WAP participation from less than 1% in the control group to about 6% in the encouraged group. The encouragement was implemented by a firm with extensive experience managing outreach campaigns and neighborhood canvassing operations, including among low-income populations. The field activities included almost 7,000 home visits, more than 32,000 phone calls, and 2,700 follow-up appointments. Ultimately, these extensive efforts only managed to increase the participation rate by 5 percentage points at a cost of more than \$1,000 per weatherized household (Fowlie et al., 2015), revealing low demand in the eligible population for a program with considerable potential benefits.

Second, the findings suggest that the benefits of these investments are substantially less than

the upfront costs. We estimate that the WAP energy efficiency investments reduce monthly energy consumption by 10-20% on average. Although this surely provides a substantial assist to participating low-income households in the form of reduced energy bills, the upfront investment costs are about twice the realized energy savings. Further, the model-projected savings are roughly 2.5 times the actual savings.

Third, while the modest energy savings might be attributed to the “rebound” effect, when demand for energy end uses increases as a result of greater efficiency, the paper fails to find evidence of significantly higher indoor temperatures at weatherized homes. This finding comes from a novel survey of measured indoor temperatures and thermostat set points that we conducted in the study population. Though the existence of the rebound effect has been the subject of much debate (Gillingham et al., 2013), our study is the first to provide a direct field test of this phenomenon for a broad spectrum of residential energy efficiency investments.³

Across a variety of metrics, the WAP energy efficiency investments appear to be poor performers on average. While these investments were free to participating households, we can nevertheless estimate the private returns if households had been responsible for the upfront costs, which is the case for households that do not qualify for WAP. Counting private returns - a household’s reductions in energy bills and willingness to pay for any change in indoor temperatures - the annual internal rate of return that would rationalize these efficiency investments is -2.2%. This finding of low, indeed negative, returns suggests that, at least for residential home retrofits, there may not be much of an efficiency gap to explain. Rather just like in all other sectors of the economy, investments with low returns are not taken-up frequently. Importantly, the engineering model used by WAP is similar, and in some cases identical, to those used to develop recommended residential efficiency investments in the broader population (i.e., not just low-income households).

These investments could still be beneficial socially, although this turns out not to be the case, on average, with the measures we evaluate. In contrast to the private calculation, the social one accounts for the benefits of reduced greenhouse gas and local pollutant emissions and the fact that part of households’ energy savings is a transfer from other energy consumers, rather than genuine

³Davis (2008) - a study of clothes washers - is a notable exception.

social savings. When the monetized reduction in environmental damages are added to the private benefits, the annual internal rate of return that would justify these investments is -0.8%, still far from what is considered an attractive rate of return. The annual social internal rate of return that also removes transfers from the calculation of benefits is -9.5%. Finally, we also calculate the average cost per ton of avoided CO₂ under a range of assumptions. The most plausible estimates are \$329/ton, which is about an order of magnitude larger than the U.S. government’s estimate of the social cost of carbon of roughly \$38 (Greenstone et al., 2013).

The paper proceeds as follows. Section 2 introduces a conceptual framework useful for valuing energy efficiency investments, including any behavioral adjustments they may cause. Section 3 outlines key details on the Weatherization Assistance Program and describes our study design. Section 4 describes the data sources and provides summary statistics. Section 5 reports the main results on actual savings and on observed rebound effects. Section 6 develops measures of the return to energy efficiency investments based on the previous section’s findings and Section 7 concludes.

2 Conceptual Framework

The standard definition of the ‘energy efficiency gap’ refers to the difference between actual and cost-effective energy use. In other words, a gap exists if individuals systematically overlook energy efficiency investments that confer benefits in excess of costs. Our primary objective is to estimate the value of ex post realized benefits derived from a set of policy-induced efficiency investments. This section introduces the conceptual framework underlying our valuation exercise.

Gains from an investment in energy efficiency are realized through two main channels: reduced energy consumption and increased consumption of energy services (e.g. lighting, space heating, air conditioning) due to reductions in the price of energy services. With respect to the first channel, any reduction in dollars spent on energy can be allocated to other forms of welfare-enhancing consumption. The second channel becomes important when an efficiency-induced reduction in energy end-use costs leads to an increase or “rebound” in the demand for the energy service. For utility-maximizing agents, any re-optimization of consumption that occurs in response to an efficiency improvement will be (weakly) welfare improving.

These basic ideas are illustrated in the upper quadrant of Figure 1, which plots consumption of a particular energy service, home heating, on the horizontal axis and consumption of the numeraire (i.e., all other goods), X , on the vertical axis. The figure focuses on home heating because it is a particularly important end-use in our empirical setting; over 93% of projected energy savings from the weatherization investments we analyze are heating related.

The two downward sloping lines in the upper quadrant of the figure reflect budget constraints, the lower before the efficiency improvement (i.e., the status quo) and the higher after weatherization. The budget constraint pivots post-weatherization because the price of heating services (e.g., the price of keeping the house at a certain indoor temperature in winter) has fallen; energy efficiency improvements reduce the cost of purchasing any given level of thermal comfort.

The figure also illustrates a family of indifference curves for a representative consumer, each of which trace out the bundles of the numeraire, X , and heating services, H , that deliver the same level of utility. The shape of the indifference curves reflects that households do not like to be too hot or too cold. In the status quo (i.e., absent an efficiency improvement), the representative agent will maximize utility through the choice of H_0 and X_0 . The weatherization-induced expansion of the budget constraint allows the agent to move to a higher level of utility associated with the bundle of H_1 and X_1 . In the figure, status quo consumption occurs below the satiation point for thermal comfort. Thus, when the price falls, demand for heating services increases by $H_1 - H_0$. The positive income effect also increases consumption of the numeraire by $X_1 - X_0$.

The paper's empirical challenge is to measure the welfare gains conferred by weatherization investments. Our empirical setting allows us to develop a measure of willingness to pay (WTP) for weatherization that accounts for both reductions in energy expenditures and the increased consumption of heating services. The effect of energy efficiency improvements on other consumption can be measured using data on monthly energy expenditures which vary one-for-one with the consumption of the numeraire. Put another way, a \$1 decrease in energy expenditures allows for a \$1 increase in consumption of all other goods. Measuring willingness to pay for the increase in heating services (i.e., the direct rebound effect) is more challenging because demand for energy services, such as heating, is not readily observable in household energy consumption or expenditure

data.

To obtain an estimate of the efficiency-induced increase in demand for heating ($H_1 - H_0$ in the figure), we conduct a survey of indoor temperatures in weatherized and non-weatherized homes. With this estimate of the treatment-induced change in indoor temperatures in hand, we can construct bounds for the welfare consequences of any observed increase in warmth by imposing some structure on the relationship between energy demand and heating services.⁴ The bottom quadrant of Figure 1 plots a representative building-specific relationship between heating services and the energy required to achieve that temperature, E , holding constant outdoor temperatures and building characteristics, Z .⁵ Efficiency improvements to the building envelope (e.g., insulation improvements, window sealing, a furnace upgrade) reduce the energy required to deliver any given level of heating services. This implies that the slope of the relationship between heating services and energy consumption becomes less steep following an efficiency improvement, depicted as the pivot in the function that determine energy consumption from $E(H; SQ, Z)$ to $E(H; W, Z)$ in the figure. Section 5.3 details our approach to empirically estimating the relationship between energy consumption and heating demand, which, in practice, is measured as indoor temperature.

We use the empirical relationship between indoor temperatures and energy consumption to construct bounds on the utility gains from any efficiency-induced increase in the demand for indoor temperature using revealed preference logic. Since the agent chooses to increase heating services from H_0 to H_1 following the efficiency improvement, it follows that a lower bound for the increased utility derived from this increase is the associated increase in heating costs incurred after weatherization. This value is represented by $P_E^*(E_{1,W} - E_{0,W})$ in the figure, where P_E is the exogenous price of energy. Note that the agent chose less heating than H_1 prior to the efficiency

⁴Our measure of the returns to weatherization investments accounts for a household's valuation of increased warmth, but we do not account for any increase in comfort conditional on indoor air temperature. Researchers have noted that improvements in insulation enhance comfort by reducing drafts and increasing humidity (Schwarz and Taylor, 1995). If a home is less drafty and more humid, a consumer may be able to achieve the same comfort at a lower indoor temperature. Consequently, what we measure here is the net effect of efficiency improvements on heating demand. We return to this point below.

⁵We assume a linear relationship between air temperature and energy demand over the relevant range of temperatures; we provide empirical support for this assumption in Section 5. Also, for ease of exposition, this figure depicts the limited range of indoor air temperatures over which energy consumption is increasing in indoor temperatures. This range is bounded from below by the outdoor air temperature (if the thermostat is set below the outdoor air temperature, no heating services are required). The lines stop where the capacity of the home heating system binds. Beyond this point, increased energy consumption ceases to generate heating services.

improvement. Thus, the cost of the change in heating services prior to weatherization, measured by $P_E^*(E_{1,SQ} - E_{0,SQ})$, provides an upper bound on the welfare gain. Preferences revealed prior to the efficiency improvement suggest that the agent values the increase in heating services less than this incremental cost.

With this conceptual framework as a guide, the paper will estimate the causal effect of WAP participation on annual energy consumption and willingness to pay for changes in heating services.

3 Background and Study Design

This section provides background information on WAP. It also details the quasi-experimental and experimental designs that we use to evaluate the consequences of the energy efficiency measures for WAP participants.

3.1 Weatherization Assistance Program

The Weatherization Assistance Program (WAP) is the nation's largest residential energy-efficiency program. WAP supports improvements in the energy efficiency of dwellings occupied by low-income families. Since its inception in 1976, over 7 million low-income households have received weatherization assistance through the program. Proponents credit the program with saving energy, creating jobs, reducing emissions, and assisting low-income households. The American Recovery and Reinvestment Act PL111-5 (ARRA) dramatically increased the scale and scope of WAP.⁶ Our analysis seeks to estimate the impacts of weatherization assistance over the ARRA-funded time period.

WAP funds are distributed to states based on a formula tied to a state's climate, the number of low-income residents, and their typical energy bills. The states distribute WAP money to over 1,000 local sub-grantees, which are typically community action agencies (CAAs) or similar nonprofit groups. These sub-grantees are then tasked with identifying and serving eligible households. Participating WAP households receive free energy audits and a home retrofit that typically includes some

⁶Funding increased from \$450 million annually in 2009 to almost \$5 billion for the 2011-2012 program years. Under the ARRA-funded program, all owner-occupied households at or below 200% of the poverty line were eligible to apply for assistance.

combination of insulation, window replacements, furnace replacement, and infiltration reduction. The average participating household in our data received an average of \$4,143 of energy efficiency investments and over \$1,000 worth of additional house improvements at zero out-of-pocket costs.⁷

Before implementing a weatherization retrofit, CAA program staff conduct an energy audit of the home. The purpose of the audit is to make recommendations regarding which efficiency improvements should be implemented at the home. During the visit, program auditors collect detailed information about the building structure and other construction details, heating and cooling systems, appliances, ventilation, etc. This information is combined with local climate conditions and retrofit measure costs, then fed into a computer-based audit tool: the National Energy Audit Tool (NEAT). This tool uses engineering algorithms to model the energy use of single-family and small multi-family residential units. NEAT is the most widely used tool for weatherization audits; it is used by state and local WAP sub-grantees, utility companies, and home energy auditors (EERE, 2010).

NEAT produces an estimate of the energy savings and costs associated with different combinations of efficiency measures. The present value of energy savings are calculated using a discount rate of 3% and an engineering estimate of the lifespan of the measures. The 3% discount rate is consistent with OMB guidance on how to evaluate benefits of federal spending but is substantially lower than the cost of borrowing for most households, especially low income ones. The WAP program requires that all recommended measures return a minimum of \$1.00 in incremental savings for every \$1.00 expended in labor and material costs. The process of applying for weatherization is highly onerous and time intensive, at least partially to prevent fraud. Applicants must submit extensive paperwork documenting their eligibility, including utility bills, earnings documentation, social security cards for all residents of the home and deeds to the home. The process that determines which applicants receive weatherization assistance is not purely random. Local agencies often identify potential applicants from the pool of households that are receiving other social services, although walk-in clients are routinely admitted. Local agencies screen potential applicants

⁷During the course of the retrofit, additional costs are incurred to ensure the safe and effective installation of the weatherization measures. For example, electric wiring updates or asbestos removal may be required to ensure a safe working environment. Once these safety measures are accounted for, the average cost per household is \$5,150.

for eligibility. Eligible applicants are then prioritized following guidelines that recommend CAAs assign the household a high rank if it has an elderly resident, a person with disabilities or children, or where the occupants typically face a high energy burden (energy as a share of income) or have high residential energy use (see 10 CFR 440.16(b) (1-5)).⁸ Given the nature of the process by which households end up in the program, any comparisons of energy consumption across weatherized and un-weatherized households risks confounding the effect of the program with pre-existing differences in determinants of energy consumption.

3.2 Research Design

The paper’s empirical challenge is to obtain causal estimates of the effect of participation in the WAP program on energy consumption and indoor heating. We use both quasi-experimental and experimental approaches to address the potentially confounding effects of non-random selection into the WAP program and compare the results across specifications.

Our analysis focuses on a sample of low-income Michigan households. Michigan is one of the largest recipients of WAP program funding on account of its cold winters and large low-income population. Further, we were able to secure collaborative agreements with a major Michigan utility and five Community Action Agencies (CAAs) working in this utility’s service territory. This allowed us access to detailed, household-level energy consumption and weatherization program data. A close collaboration with two of these agencies allowed us to implement a large-scale field experiment.

Michigan received over \$200 million in ARRA funding for weatherization assistance. A series of bureaucratic delays - for example, ensuring that contractors were paid prevailing wages - delayed spending until early 2011 (Radnofsky, 2010). Around March 2011, weatherization activities in Michigan increased markedly. All stimulus funds had to be spent by March 2012. After that point, the pace of weatherization activity dropped precipitously. We stopped collecting data in April 2014. A number of households that applied for WAP had not yet received services by the end of our study period.

⁸Given the high ARRA funding levels during our study period, the prioritization scheme was less binding as compared to lower funding periods.

3.2.1 Quasi-Experimental Design

The quasi-experimental research design uses data on all of the households that applied for WAP after March 2011 with one of the five CAAs that shared data with us. In this sample of households, we compare patterns in energy consumption among weatherized households and households that applied for WAP but had not been weatherized by mid-2014, when our data ends. Forty percent of applicants were weatherized through WAP during this period.

A critical issue for the validity of the estimates from this design is how households in this sample were chosen for weatherization. The road from application to energy efficiency investments is long and there are many potential off-ramps. Applicant households may fail to complete the necessary – and involved – paperwork or may be deemed ineligible based on the information they provide. Once paperwork is completed successfully, households are put on a list where the waiting times can exceed one year. After rising to the top of the list, homeowners must accommodate scheduling of energy audits. Households may fail to receive weatherization if they miss an audit appointment, or if the auditors discover risks to WAP contractors (e.g., asbestos in the home). Because of significant delays in ramping up weatherization activities under ARRA, the agencies were unable to complete the weatherizations they anticipated prior to the March 2012 ARRA deadline, which helps to explain why fewer than half of the applicants in our sample were weatherized by mid-2014. Some of the explanations for variation in treatment status among applicants are orthogonal to household characteristics that determine energy consumption patterns, while others clearly are not.

We estimate the following equation designed to control for the effects of unobservable factors that determine energy consumption trajectories at households that apply for weatherization assistance:

$$\ln(y_{imt}) = \beta \mathbf{1}\{WAP\}_{imt} + \alpha_{im} + \alpha_{mt} + \epsilon_{imt}, \quad (1)$$

where $\ln(y_{imt})$ measures the natural log of energy consumption (natural gas, electricity, or a combined measure) at household i in month m and year t . The WAP indicator variable switches from zero to one in the month after a household’s weatherization retrofit is complete.⁹ The equation

⁹For approximately 70% of households we observe the date that the work is reported complete, the post-inspection is complete, and the job is finally “closed out.” For those households for whom we observe all three dates, these dates

includes household-by-month-of-year fixed effects, α_{im} , to account for permanent differences in a household’s energy consumption across months. It is possible to include such a rich set of fixed effects that account for household-specific seasonal variation in energy consumption because we observe households across multiple years. The model also includes month-by-year fixed effects, α_{mt} , to adjust for the average effects of time-varying factors (e.g., winter temperature) that generates variation in average consumption across all households.¹⁰

The parameter of interest is β , which measures the mean difference in energy consumption subsequent to the completion of WAP energy efficiency investments, after adjustment for the fixed effects. It is a difference-in-differences estimator that compares the change in energy consumption after weatherization to before, relative to consumption among households that have either not yet weatherized through WAP or never did during our sample period.

In equation (1), the primary threat to consistent estimation of β is the possibility that time-varying factors that affect household demand for energy also influence WAP participation. For instance, households may push forward their WAP application more aggressively when they anticipate an increase in their demand for energy as would be the case when the number of people in their household increases or they lose a job and expect to spend more time at home. While the quasi-experimental approach does not have a direct solution to this threat to identification, we take steps to balance observable characteristics and trends across weatherized and unweatherized households. To evaluate the robustness of our findings, we re-estimate equation (1), using alternative sets of controls and regression weights. Our preferred quasi-experimental specification re-weights control observations in order to achieve covariate balance across weatherized and un-weatherized controls, explained in more detail below.

typically fall within a month of each other, although we observe gaps between the complete date and the final close-out date as long as five months. For less than three percent of households, only a final close-out date is reported. For these households we assume that the weatherization work was completed the month before.

¹⁰We also estimate more highly saturated specifications that include month-by-year-by county and month-by-year-by-billing segment fixed effects.

3.2.2 Randomized Encouragement Design

The paper’s primary empirical estimates are derived from an experimental research design. The basis of the experiment is a randomly assigned encouragement intervention that aims to increase the probability of treatment households’ participation in WAP through recruitment and significant application assistance. The recruitment and assistance was conducted by FieldWorks LLC, a private company that specializes in running neighborhood canvassing operations and managing outreach campaigns. Importantly, we chose them because they had substantial experience working with low-income populations and their staff had generated millions of phone calls and knocked on millions of doors in previous engagements.

The experimental sample comprised 34,161 households that were both presumptively eligible for WAP and located within the counties served by the two CAAs that we established a close partnership with. Beyond sharing data, these agencies agreed to work with our field staff to process the applications we helped households complete. Approximately one quarter of these households were randomly assigned to our encouragement “treatment.” FieldWorks encouraged the treated households to apply for WAP and offered them extensive application assistance. For the remaining households assigned to the control group, we simply observe energy consumption and program participation decisions. The Data Appendix and Fowlie et al. (2015) provide more details on the encouragement and application assistance programs.

The encouragement campaign got underway as ARRA funds began flowing to the implementing agencies. Encouragement activities ran from March to May 2011. Table 1 summarizes our encouragement and enrollment activities in detail. During the encouragement phase, field staff made almost 7,000 initial, in-person house visits.¹¹ These ground operations were complemented with 23,500 targeted “robo-calls” to raise awareness of both the weatherization program and our encouragement campaign.

After an initial encouragement phase designed to generate interest in the program, we transitioned to an enrollment phase, which lasted through February 2012. The application process requires households to provide extensive documentation, including utility bills, social security cards

¹¹Most- but not all- houses assigned to the treated group were contacted. A small fraction were deemed inaccessible (e.g., because of a locked gate).

for all household members, and the deed for the house. Further, a minority of applicants were required to provide further documentation or to correct something on their initial application. Our staff made over 9,000 personal phone calls to provide assistance and to coordinate in-person meetings. Over the course of 2,720 home visits, our field staff helped individuals assemble documentation and complete paperwork. In some cases, our field staff provided transportation to and from the program agency offices.

The final row of Table 1 reports that we spent around \$475,000 on the encouragement or a little more than \$55 per household in the treatment group. It is noteworthy that we did not initially intend to devote such extensive efforts and resources to the encourage and enrollment phases. However, the early results suggested that we were failing to have a substantial impact on applications and concerns about the ultimate precision of our estimated treatment effects motivated us to raise additional funds to expand the share of treated households.¹²

We use the random assignment to encouragement as an instrumental variable for weatherization status. In a first stage, we use OLS to estimate:

$$\mathbf{1}\{WAP\}_{imt} = \theta \mathbf{1}\{Encouraged\}_{imt} + \delta_{im} + \delta_{mt} + \eta_{imt}, \quad (2)$$

where the dependent variable is as described in the previous subsection. The indicator variable, $\mathbf{1}\{Encouraged\}_{imt}$ is set to zero for all households prior to the encouragement intervention. After March 2011, this indicator switches to 1 for the 25% of the households randomly assigned to the treatment group.

We substitute $\mathbf{1}\{W\hat{A}P\}_{imt}$ from the estimation of equation (2) to fit equation (1) and obtain $\hat{\beta}_{IV}$. In this instrumental variables (IV) framework, $\hat{\beta}_{IV}$ is identified using the exogenous variation in program participation that is generated via the random assignment of encouragement.

If there is heterogeneity in how weatherization retrofits impact residential energy consumption, the expectation of unbiased quasi-experimental and experimental estimates of β need not be

¹²On the one hand, our encouragement costs may have been higher than necessary. To our knowledge, ours was the first encouragement program for WAP, and we learned by doing. On the other hand, the costs in Table 1 do not reflect the time that the research team devoted to overseeing the encouragement effort.

equivalent. Our quasi-experimental approach is designed to provide an estimate of the average treatment effect on all treated households (ATET), or an average across the full distribution of treatment effects. In contrast, the randomized encouragement design estimates the so-called local average treatment effect (LATE), or the average effect for the subset of the population who must be encouraged to participate in the program (i.e., the compliers) (Angrist, Imbens, and Rubin, 1996). In other words, significant differences in the quasi-experimental and experimental estimates could be due to bias in the former or differences in the LATE and ATET.

4 Data Sources and Summary Statistics

4.1 Data Sources

The data collected to support this analysis correspond to two overlapping groups of households. The first group comprises the 34,160 households in our experimental sample drawn from the counties served by two agencies. The second group of households corresponds to our quasi-experimental research design. As this design did not require the same degree of coordination with our agency partners, we were able to expand the scope of this sample by collecting detailed data from three additional implementing agencies. The quasi-experimental sample includes the 7,304 households that applied for weatherization assistance at these five agencies. The quasi-experimental sample is smaller overall but has a larger number of applicants and weatherized households, relative to the experimental sample. The two groups overlap as 1,773 applicant households are also part of our experimental sample.

4.1.1 Energy Consumption Data

We obtained monthly natural gas and electricity consumption data over the period June 2008 to May 2014. This period includes at least two years of pre-retrofit data for all weatherized households in our sample.¹³ The utility data track monthly kilowatt-hours (kWh) of electricity and thousand cubic feet (Mcf) of gas used at the dwelling. We convert both of these variables to million British

¹³We identify a household as a particular account number at an address. For a small number of addresses, we have data on multiple accounts if one household moved out and another moved in.

thermal units (MMBtu) using the standard conversion factors employed by the WAP program.¹⁴

Energy consumption records obtained from the utility are merged with households in our experimental sample and the applicant data obtained from the five implementing agencies. Data are merged using detailed name and address information. Not all households find an exact match in consumption records. Match rates are 85% and 69% in our experimental and quasi-experimental samples, respectively. The higher match rate in our experimental sample is to be expected; when selecting this sample we focused exclusively on zip codes within the territory of our partner utility.

4.1.2 Energy Efficiency Measures

The primary output of the household-level energy audits, which are part of the WAP implementation as described above, is the list of energy efficiency measures for which projected energy savings exceed costs. These are the measures that are eligible for implementation through WAP. In addition to the household-specific list of recommended measures, we also acquired the work summaries that are filed after the work at each house is completed. These data allow us to confirm that the recommended measures were installed and to compare realized costs with projections.

4.1.3 Indoor Temperature Data

Two years after the encouragement effort was initiated, we randomly selected a subset of weatherized and unweatherized households for a field survey. These households were selected from the quasi-experimental sample. The primary purpose of the survey was to collect measurements of thermostat set points and indoor temperatures that can be used to test for a direct “rebound effect.”

Michigan field staff attempted to contact 6,400 households on cold days (projected maximum temperature below 45 degrees Fahrenheit) in March and early April 2013. Survey questions were

¹⁴We also reconfigure the data to account for the fact that consumption records correspond to billing cycles or segments versus calendar months. The utility assigns households to one of 21 billing portions. In a given month, each portion maps to a different set of calendar days. For example, one household’s June bill may reflect consumption in all of June, whereas another household’s June bill captures the last half of May and the first half of June. The utility provided us with a detailed mapping of billing segments to calendar days over the duration of our study period. For each customer and for each billing cycle, we divide meter reads evenly across days in the cycle. These “meter-day” measures are then aggregated by calendar month to construct estimates of monthly consumption at the household. We also estimate regression equations that include month-by-year-by-billing segment fixed effects using unadjusted data. Coefficient estimates are unaffected.

designed to collect information about thermostat set points and, where applicable, household's experience with weatherization. With the homeowner's permission, surveyors entered the home, closed the door, moved to the center of the room, and recorded multiple indoor air temperature measurements using two different thermometers. Of our initially targeted sample, surveyors spoke with 1,658 homeowners. Of these, 899 allowed us to enter their homes and record their thermostat set point and 688 allowed us to close the door, and collect two or more indoor thermometer readings. We obtained thermometer readings from a slightly larger share of the weatherized households in the survey compared to households that did not participate in WAP.

4.2 Summary Statistics

Table 2 summarizes pre-treatment information on the households in both the RED and quasi-experimental samples. The top panel summarizes monthly energy consumption during the two years immediately preceding the treatment period. The first two columns summarize means for the randomized encouragement and experimental control groups, respectively. The third column reports differences between the treatment and control groups as well as the standard error of the difference (in parentheses).

There are seasonal patterns in energy consumption that differ for natural gas and electricity. Across all groups, winter natural gas consumption (which is dominated by space heating) is significantly higher than summer gas use (comprised primarily of hot water heating and cooking). Electricity usage, also measured in MMBtu, is fairly consistent across seasons.¹⁵

Because households in the experimental sample were randomly assigned to the encouraged and control groups, it is unsurprising that differences in pre-March 2011 natural gas and electricity consumption across these two groups, reported in column 3, are all small and statistically indistinguishable from zero. The table also provides an opportunity to judge the credibility of the comparisons that underlie the quasi-experimental estimates that complement the paper's experimental estimates. The fourth column reports on all weatherized households in the territory covered

¹⁵Overall, the lower income households in our sample consume less energy on average than the typical household served by our partner utility. In our sample, average monthly natural gas consumption is approximately 7.33 MMBtu. Across the entire service territory, the annual residential natural gas consumption averages 7.75 MMBtu per month.

by the five community action agencies that provided us with data. The fifth column summarizes households in these territories that applied for weatherization but had not received weatherization assistance as of April 2014. In practice, the variation in the weatherization dates means that the identification of the quasi-experimental estimator is not just based on comparisons between the samples summarized in columns (4) and (5), but also relies on within household comparisons. It is nevertheless informative to compare these two sets of households. The mean differences in column (6) show that weatherized households have historically consumed significantly less natural gas than the unweatherized applicants during both winter and summer months

Panel B of Table 2 summarizes the detailed demographic information and dwelling characteristics that are collected for most clients as part of the application process and also documents important differences between the weatherized and unweatherized applicant subsamples. Our summary of these variables focuses exclusively on program applicants because these data are not available for the majority of households in the experimental sample that did not apply to the program. Weatherized households have higher incomes in an absolute and relative sense as compared to unsuccessful applicants.¹⁶ Weatherized households also report having more children, are more likely to report an elderly resident, and are more likely to use natural gas as their primary heat source. They are less likely to report a disabled resident.

The significant differences between the weatherized and unweatherized applicants motivate us to re-weight observations in the control group so that observable factors are distributed similarly in the weatherized and unweatherized applicant groups. Although our preferred specifications includes household-by-month-of sample fixed effects that control for all time invariant differences between households, we are concerned that observable fixed differences might be correlated with time-varying differences. We use an estimated propensity score to balance covariates that presumably play a role in determining program take-up. Section 10.2 provides exact details on the participation equation but the set of controls includes income, demographics, and levels and trends in energy

¹⁶Program eligibility is based on the household's income, relative to the poverty line. The Census Bureau definition of the poverty line varies by family size and composition. To qualify for the program, a household's income cannot exceed 200% of poverty. Applicants fall well below this threshold on average. In addition to having higher incomes, applicants who ultimately receive weatherization are farther above the poverty line than the unsuccessful applicants.

consumption.¹⁷

Appendix Table 1 summarizes some of the detailed information collected during the household energy efficiency audits. The entries are based on 1,638 households or roughly 75% of the households that participated in WAP in the quasi-experimental sample.¹⁸ A typical weatherization retrofit involved several measures such as furnace replacement (34% of retrofits), attic insulation (85%), wall insulation (44%), and infiltration reduction (76%). The NEAT model predicts that on average these investments will reduce natural gas consumption by 46% and electricity consumption by 15%. The average project involved over \$5,100 in total expenditures. This includes materials, labor, and construction costs, but does not include any program overhead. Using a 3% discount rate, the projected net present value of energy bill savings average \$10,689. The average projected savings:investment ratio (across measures) exceeds 2:1.

5 Results

5.1 Quasi-Experimental Estimates of Energy Savings

Table 3 presents the quasi-experimental estimates based on the estimation of equation (1). The dependent variable in all regressions is the log of monthly energy consumption (i.e., the sum of electricity and natural gas both measured in MMBtu). The first two columns use data from all weatherization applicants, with the second specification allowing time period effects to vary across counties. Columns (3) and (4) trim the sample to obtain estimates that are more directly comparable to the experimental estimates. Specifically, this sample is limited to the implementing agencies that participated in the experiment and to applicants that applied after the encouragement intervention was initiated. Columns (5) and (6) report estimates comparable to columns (1) and (2) reweighted by the propensity score.¹⁹ All specifications include household-by-month fixed effects

¹⁷Reassuringly, none of the differences in average covariate values across weatherized households and propensity-score weighted controls are statistically significant.

¹⁸Only 1,638 of the NEAT data files could be confidently matched with weatherized households (leaving 436 weatherized households in our data with no match in the NEAT audit data). We applied fairly strict matching criteria so as not to mismatch weatherized households with audit information. Occupant names were often not included in the audit files. Addresses appear to have been miscoded in several instances.

¹⁹Because some of the covariates included in the propensity score estimating equation (e.g., reported disability and number of children) are not reported by all households, this sample is somewhat smaller.

and standard errors are clustered by household in all specifications.

The first row in Table 3 reports the estimated average treatment effect. Across the columns, the estimates suggest that WAP participation reduces energy consumption by roughly 8-10%. This is the first indication that the realized savings from the WAP-induced energy efficiency investments are substantially smaller than the projections from the engineering model. These estimated energy savings are approximately 25% of projected savings (measured in mmbtu).

Figure 2 provides another perspective on the results in Table 3. Specifically, this figure reports on the estimation of a version of the column (1) specification, except the weatherization indicator is interacted with indicator variables for each of the potential quarters before and after weatherization was completed (the time zero effect captures energy consumption in the month of weatherization). We also interact the quarter indicators with heating degree days (HDD) and HDD squared, to account for the fact that weatherization retrofits may lead to higher energy savings during colder quarters.²⁰ The figure then plots the coefficients on the quarter dummy variables plus the coefficients on the HDD and HDD squared variables multiplied by the average HDD and HDD squared in our data. We also plot 95% confidence intervals associated with these coefficient sums. The decline in energy consumption is apparent and seems roughly constant throughout the period of our study.

Panel B of Table 3 computes the present value of the estimated energy savings under alternative assumptions about investment time horizons and discount rates. To express the estimates of monthly energy savings in dollar terms, we first construct estimates of fuel-specific savings (see Appendix). We use these estimates to construct fuel-specific savings in percentage terms. Estimates of average monthly natural gas savings are multiplied by the product of monthly gas consumption in the control group and the residential retail price of natural gas in Michigan in 2013 (EIA, 2015).²¹ Similarly, average 2013 electricity savings (in percentage terms) are multiplied by the

²⁰Heating degree days reflect the difference between the outdoor temperature and a base temperature of 65° F. A day's HDD equal zero for days when the hourly temperature exceeds 65° F for all hours of the day. It is equal to a weighted difference between 65° F and hourly temperatures when the temperature dips below the base. Weights are determined by the share of hours at each temperature. Quarterly HDDs sum daily HDDs over all days in the quarter.

²¹The average retail price of natural gas in Michigan in 2013 was \$10.46/MMBtu (expressed in \$2013). This is higher than the average price charged by this utility over the entire treatment period (\$7.98/MMBtu). Natural gas prices were at historic lows over this period, so using the observed average price would likely underestimate the real

product of average monthly electricity consumption in the control group and retail electricity price (\$0.11/kWh).²² Taken together, our estimate of average energy savings is approximately \$155 per year.

To compute the net present value of energy savings over the useful life of the improvements, we invoke some additional assumptions. First, we rely on the NEAT simulation program's assumptions about measure-specific lifespans. These projected lifespans range from 3 years (for a furnace tune-up) to 20 years (for attic insulation). The energy savings-weighted average lifespan for installed measures in our dataset is 16 years. In the table, we report discounted benefits with assumed lifespans of 10, 16, and 20 years. We also assume that the effect of weatherization on energy consumption - and real energy prices - do not vary over the life of the measure. Discounted benefits are calculated at discount rates of 3%, 6%, and 10% (see columns).

The energy savings are small relative to the upfront costs. The nine estimates of the present value of the savings range from just below \$1,000 (high discount rate and short time horizon) to about \$2,300 (low discount rate and long time horizon). These estimates are between 20% and 45% of the average upfront cost of the energy efficiency retrofits.

5.2 Experimental Estimates of Energy Savings

5.2.1 First-Stage: Program Take-Up

It may seem straightforward to encourage households to participate in a program that provides free efficiency retrofits worth an average of approximately \$5,000 that are designed to significantly reduce energy expenditures. In our experience, that was hardly the case. The impact of reducing barriers to participation (e.g., information and process costs) on program uptake is of independent interest both to policymakers and researchers. This section evaluates the impact of our intervention on a multi-stage participation process.

prices that will prevail over the life of these investments. Prices in 2013 are somewhat lower than the average real prices over the period 2000-2013. The shale gas boom has arguably ushered in a new domestic price regime, such that a longer average real price will overestimate future prices.

²²The NEAT program audits assume an electricity price of \$0.11/kWh and a natural gas price of \$11.46/MMBtu. The higher gas price is presumably based on 2006 prices which averaged around \$11.50/MMBtu in this service territory in 2006.

Table 4 summarizes program take-up at three separate stages. In the first stage, our goal was simply to increase the share of households filing applications. The coefficient estimate in column (1) indicates that the encouragement intervention increased the rate of application to the program by 13 percentage points from the control group mean of 2%.

Once a household's application is approved, the energy audit is the critical next step. Several months can elapse between application approval and energy audit scheduling. Audits are only scheduled when the implementing agency is in a position to hire the construction crews and allocate any other resources needed to implement a weatherization retrofit at the home. Column (2) reveals that the fraction of households who received an energy audit was 5 percentage points higher in the encouraged group (off a base of about 1%). As we discussed above, several factors can explain why so many households that submit an application are not audited, including if the household fails to follow through on requests for more information or if the submitted information indicates that the household does not meet the program's eligibility requirements.

Column (3) of Table 4 documents that the treatment increased the fraction of households that were successfully weatherized by about 5 percentage points, against a 1% rate in the control group.²³ The encouragement treatment is a statistically significant predictor of weatherization, which we will use to instrument for program participation.

The low take-up rates in the encouraged group are quite striking. Program participants receive substantive home improvements, yet incur no out-of-pocket expenses. All households in the encouraged group received some information about the program via a phone call or door hanger. A majority of households (i.e., those who spoke with our canvassers in person or by phone) received further information about our offer of application assistance. Given that households had detailed, specific information about the program, it seems reasonable to surmise that some combination of high perceived costs of applying for the program, low expectation of an application leading to a weatherization, high unmeasured process costs, and low expected benefits of participating in the program are impediments to WAP participation. In the end, the average cost of encouragement per completed weatherization was about \$1,050, which is more than 20% of the average costs of

²³A small fraction of households get audited but not weatherized, primarily because the auditors deem the home a possible danger to weatherization contractors (e.g., due to the presence of asbestos).

measures installed. See Fowlie et al. (2015) for further details.

5.2.2 Instrumental Variables Estimates of Energy Savings

This section presents the experimental estimates of the impact of weatherization on energy consumption, using random assignment to the encouragement group as an instrument for program participation. Figure 3 provides a graphical overview of this IV design. The broken line shows the cumulative effect of the randomly assigned encouragement intervention on the monthly rate of weatherization, relative to the control group which received no encouragement. This effect increases over time as the treatment households submit applications and receive weatherization assistance. The figure also plots month-by-month estimates of intent to treat (ITT) effects on energy consumption. Conceptually, monthly estimates of the local average effect of weatherization on energy consumption can be constructed as a ratio of the monthly ITT estimates and the corresponding effect of encouragement on program participation.

Panel A of Table 5 summarizes results that relate measures of energy consumption to the WAP participation indicator. In the first two columns, the dependent variable is the log of total energy consumption (MMBtu/month). In the third and fourth columns, the dependent variables are the log of natural gas and electricity consumption, respectively.

The first column reports on the application of one conventional associational approach. Specifically, it reports on the coefficient associated with the WAP indicator from an OLS regression with month-of-sample and household-month fixed effects that is fit to the experimental sample. This specification indicates that WAP participation is associated with a 9.5% decline in energy consumption, which is similar to the quasi-experimental estimate in Table 3.

The IV estimates based on our randomized encouragement design are reported in columns (2) - (4). The estimate in column (2) indicates that WAP participation causes a reduction in monthly energy consumption of approximately 19% among households that were encouraged into the WAP program. The experimental point estimate of energy savings is twice as large the non-experimental estimate. Columns (3) and (4) report local average treatment effects for natural gas and electricity, respectively. Natural gas accounts for 94% of projected savings (measured in

MMBtu), so it is not surprising that natural gas consumption is more significantly impacted by weatherization.²⁴ Our IV strategy is predicated on an exclusion restriction: we assume that our encouragement activities affected energy consumption only through its effect on participation in the WAP program. To informally test whether the treatment’s encouragement activities had a direct effect on energy consumption, we test for an effect of our encouragement activities on the households in the encouraged group that did not receive weatherization assistance. We fail to reject the null of no effect on energy consumption among these households.²⁵

Panel B of Table 5 computes the present value of the estimated energy savings under alternative assumptions about investment time horizons and discount rates. The approach we take is similar to that described in section 5.1. One difference is that we cannot directly observe energy consumption in the post-assignment period among compliers assigned to the control group. Instead, we impute the average consumption among unweatherized compliers using average consumption among unweatherized households in the encouraged and control groups, respectively, and the estimated proportion of compliers in the population (reported in Table 4).²⁶

As with the quasi-experimental results, the experimentally estimated energy savings are small relative to the projected savings and the upfront costs. Our estimates imply energy savings of 17.2 MMBtu per year, whereas the average projected savings among compliers is 43.7 MMBtu. Our central estimate of the realized average savings per household is roughly \$2,400. This is

²⁴The coefficient in column (2) is not a weighted average of the gas and electricity coefficients in columns (3) and (4) for two reasons. First, the samples differ slightly across the columns, and second the coefficients on the fixed effects and other covariates are not constrained to be equal in columns (3) and (4). It is also worth noting that the effect of weatherization on gas consumption is more precisely estimated. This is not surprising, because natural gas consumption is driven primarily by the end uses targeted by weatherization (space and water heating), whereas electricity consumption is derived from many end uses that are unaffected by weatherization (e.g., lighting and appliances).

²⁵To conduct this test, we drop all households receiving weatherization assistance from the experimental sample. Using the remaining households, we regress the log of monthly energy consumption on an indicator that equals zero before encouragement activities were initiated and one after they began in March 2011 for all households assigned to the encouraged group. We estimate a precise zero effect which suggests that our encouragement intervention had no effect on energy use in these households and supports the validity of the exclusion restriction. Of course, our exclusion restriction also implies that the encouragement intervention does not directly effect energy consumption among households taking up the weatherization treatment but we cannot test this assumption directly.

²⁶In Table 4 we estimate the share of always takers and compliers in the population to be 1% and 5%, respectively. We directly observe the average energy consumption among unweatherized households in the encouraged group (i.e., the never takers) during the post-assignment period. We also observe the average consumption among unweatherized households during the post-assignment period in the control group. As this represents a weighted average of consumption across never takers and compliers, we can impute the average post-assignment consumption among unweatherized compliers. This is the imputed baseline consumption reported in Table 5.

significantly lower than the average net present value of savings predicted by the ex ante engineering analysis which is \$9,206 among compliers (and \$10,689 averaged across all weatherized households in our data).²⁷ The nine estimates of the present value of the savings range from approximately \$1,450 (high discount rate and short time horizon) to about \$3,500 (low discount rate and long time horizon). These estimates are just 32% to 77% of the upfront cost of the energy efficiency measures.²⁸

The estimated local average treatment effect of weatherization on energy consumption is approximately twice as large as the quasi-experimental estimate of the average effect among all households receiving weatherization assistance. The Appendix investigates several potential explanations for this difference. We note that there are significant observable differences between the group of complier households in the experimental sample and the weatherized households in the quasi-experimental sample. Adjusting parametrically for these observable differences explains approximately 20% of the discrepancy between the experimental and quasi-experimental treatment effect estimates. The remaining difference in estimated causal effects could be due to unobservable differences between complier households and households that need no encouragement to pursue weatherization assistance. Additional explanations could be that omitted time-varying factors bias the quasi-experimental estimates towards zero or that measurement error in the WAP indicator lead to attenuation bias in the quasi-experimental estimator. Note that the last explanation is consistent with the fact that the estimate in column (1) of Table 5 is comparable to the quasi-experimental estimator.²⁹ Finally, given that we are using two distinct data sets, the differences in the estimated parameters could be due to sampling variation. We have no way of distinguishing between these explanations.

²⁷The average present discounted savings among weatherized households in the control group was projected to be \$11,215. The average projected discounted savings among weatherized households in the encouraged group is \$9,490. We estimate that compliers account for 85% of weatherized households in the encouraged group. We thus impute that the average projected savings among complier households is \$9,206.

²⁸To compute these percentages, we use an average cost of \$4,550 among compliers. To construct this average, we use our estimate that 85 percent of weatherized households in the encouraged group are compliers. We observe the average cost among always takers in the control group to be \$5,123. We interpret the average cost among weatherized households in the encouraged group, \$4,636, to be a weighted average across compliers and always takers. Note that the range of the percentages would likely be different if we had household-level estimates of savings and treatment costs.

²⁹Results are robust to requiring more and less documentation from the CAAs to define a household as “weatherized.” Measurement error could exist, for instance, if there are typos in the CAAs’ paperwork.

5.3 Household Reoptimizing Behavior, Building Thermal Properties, and the Welfare Implications of Rebound

In Section 2 we introduced a utility-maximization framework that uses a revealed preference approach to bound the average willingness to pay for any efficiency-induced increase in households' consumption of energy services. In this section, we implement this bounding exercise in three steps.

5.3.1 Does Weatherization Lead to Temperature “Take Back”?

The first step tests for an effect of weatherization on household demand for space heating. Table 6 summarizes the results from our survey of indoor air temperatures and thermostat set points collected during the winter of 2013 at a subset of our quasi-experimental households. All weatherized households surveyed had received efficiency improvements at least one year before the survey was administered, allowing plenty of time for residents to observe the extent to which the weatherization retrofit affected winter heating costs. A total of 899 households allowed our survey team to enter their home to record the thermostat set point, and 688 allowed our surveyors to linger long enough to record the actual indoor temperature. Approximately half of the households for which we have thermostat or temperature data had received weatherization assistance (453/899 and 349/688). Anticipating some measurement error, we used two different devices to measure indoor temperatures at each home.

Table 6 reports results from regressing thermometer readings and thermostat set points on a binary variable indicating whether the household had been weatherized.³⁰ Columns (1) and (3) report results from our base specification. Survey respondents comprise a non-random subset of our sample as only a fraction of the targeted household were home and/or willing to open the door to receive our surveyors. Further only 41% of households who opened the door to our surveyors allowed us to come in and collect temperature measurements. Because we observe differences between survey respondents and the larger sample, we also estimate specifications where survey household observations are weighted to match the covariate means in the larger quasi-experimental sample

³⁰If surveyors recorded temperatures from both devices, the temperature specifications include two measurements per household. Standard errors are clustered at the household level.

of weatherized households.³¹ These results are reported in columns (2) and (4). All specifications control for the outdoor temperature on the day of the survey, measured by heating degree days (HDD).

Columns (1) and (2) show that, after controlling for outdoor temperatures, indoor temperatures may be slightly warmer at weatherized households. The point estimate suggests an increase of 0.65 degrees. This effect is imprecisely estimated; we fail to reject the null of zero. The coefficient associated with the HDD variable is statistically significantly negative in the temperature specifications, suggesting that indoor temperatures measure lower on colder days. Although we asked surveyors to wait several minutes before recording temperatures, this finding suggests that cold air brought with the surveyor could be affecting the measurements.³² Because the standard errors of our estimates do not allow us to reject a small increase in indoor temperatures, we will estimate an upper bound on possible welfare gains using an estimated increase in temperature of 0.65 degrees based on the point estimate in Column (2).³³

In contrast, Columns (3) and (4) suggest that weatherized households set their thermostats lower, on average, by approximately 0.6 degrees F. This is inconsistent with a rebound in demand for indoor heat. One possible explanation is that retrofits, by reducing cold air infiltration, allow households to maintain the same (or slightly higher) levels of indoor comfort at lower thermostat set points.

³¹For both the thermostat set point and indoor temperature samples, we developed pairwise comparisons of observable household and dwelling characteristics across weatherized and unweatherized surveyed households and the larger quasi-experimental sample. The surveyed sub-sample is observationally similar among most - but not all - dimensions. For example, for both dependent variables, survey respondents are significantly more likely to report having children or elderly family members, less likely to be unemployed, and are more likely to use gas as their primary source of heat, as compared to the larger quasi-experimental sample. Among unweatherized households, surveyed households tend to be larger.

³²Notably, our estimated constants in columns (1) and (2) minus approximately 5° F, recognized by engineers as a typical internal heat gain (e.g., from body heat), is equal to the base temperature of 65° F used in standard HDD calculations.

³³The positive coefficient on the weatherization dummy in columns (1) and (2) appears sensitive to outliers in our data, which may reflect coding errors by our survey team. For example, when we trim the sample by excluding the top and bottom 5% of the observations, the coefficient on a specification equivalent to column (4) is 0.01 (standard error = 0.27). The results on the thermostat set point are not sensitive to data trimming.

5.3.2 Building Energy Performance

In order to estimate how a given increase in demand for indoor space heating translates into an increase in monthly gas consumption, we need to estimate the thermal properties of a representative building. We use the so-called “degree day method” to model the energy required to increase temperatures in an average home in our data (Thorpe, 2013) . Heating degree days (HDDs) measure the difference between outdoor temperatures and a temperature that people generally find comfortable indoors. In the literature that analyzes energy use in buildings, it is common to assume a linear relationship between energy consumption and this difference (e.g., Friedman, 1987; Dyson et al., 2014). We find empirical support for this assumption in our setting. We use this estimated relationship between HDD and monthly energy consumption to estimate the effect of increased demand for indoor temperatures on energy consumption. This approach assumes that a household’s choice of the indoor temperature is independent of outdoor temperatures, thus outdoor temperatures are a valid proxy for the desired level of heating services.³⁴

In residential buildings where building envelope losses are the major determinant of heating energy requirements, it is standard to summarize the technical relationship between energy consumption and heating demand by regressing energy consumption on HDDs. Specifically, we estimate the following equation:

$$C_{imt} = \alpha_i + \beta_1 \mathbf{1}\{WAP\}_{imt} + \beta_2 HDD_{mt} + \beta_3 HDD_{mt} * \mathbf{1}\{WAP\}_{imt} \quad (3)$$

$$+ \beta_4 HDD_{mt}^2 + \beta_5 HDD_{mt}^2 * \mathbf{1}\{WAP\}_{imt},$$

where C_{imt} measures the natural gas consumption at household i in month m and year t . We estimate this equation using data collected from program applicants during winter months (September-March) over the entire sample period. Panel data allow us to include household-level fixed effects in this regression. We include both a linear and a quadratic HDD term, allowing each coefficient to vary with weatherization status, as well as a separate intercept for WAP participating

³⁴Note that this assumption is consistent with the results in Table 6 which fail to reject that there is no relationship between HDD and thermostat set point.

households.

All of the estimated coefficients in equation (3) are highly statistically significant and very precisely estimated. The R-squared value is 0.73. Figure 4 summarizes the relationship between energy consumption (during winter months) and HDD separately for weatherized and unweatherized observations. This figure is analogous to the bottom panel of Figure 1 (flipped into the top quadrant).

The estimated slope of the relationship is less steep among weatherized homes. That is, weatherizations effectively reduce the marginal cost of indoor space heating during the winter. Moreover, it is noteworthy that this relationship is approximately linear though we estimated quadratic terms.

5.3.3 Bounding the Average Valuation of Increased Indoor Heat

This subsection implements an approach to bounding the welfare gain from higher indoor temperatures. The product of the slope of the relationship between natural gas consumption and HDD (depicted in Figure 4) and the average natural gas price in the post-encouragement period provides a measure of the the marginal cost of gas heating. Among unweatherized households, this product is equal to approximately \$0.072 per heating degree day (or \$2.17 per heating degree month). The analogous calculation for weatherized households lead to an estimated marginal cost of \$0.056 per heating degree day (or \$1.67 per heating degree month). This implies that weatherization led to a reduction in the marginal cost of approximately 20%.³⁵

These estimates of the marginal costs of heating among weatherized and non-weatherized households can be combined with the estimated 0.65 degree increase in indoor temperatures (see Table 6, column (2)) to bound the average welfare gain from weatherized households' reoptimization.³⁶ The lower bound of households' valuation of the higher temperatures is given by $0.65^{\circ}\text{F} * \$1.67/\text{degree-month}$ or \$1.09 per winter month. At this lower bound, the utility gains from increased warmth

³⁵These estimates of incremental heating costs are comparable to a "rule of thumb" popularized by the American Council for an Energy-Efficient Economy. This rule states that a household will pay approximately 3% on their gas bill for a degree increase in winter thermostat settings (see, for example, <http://www.improvenet.com/a/5-easy-ways-to-lower-your-gas-bill-during-the-winter>). Average natural gas bills during winter months are \$85.95 and \$57.96 at unweatherized and weatherized homes, respectively.

³⁶Because the relationship between heating services and energy consumption is approximately linear, the average temperature increase yields the average willingness to pay for heating services.

are exactly offset by the increase in the energy expenditures incurred to achieve the temperature increase, implying a zero gain in welfare. To define the upper bound, we note that by revealed preference, unweatherized households chose not to pay $0.65^{\circ}\text{F} * \$2.17/\text{degree-month}$ or \$1.41 per winter month to achieve this incremental increase in temperature. It follows that average marginal benefits from this temperature increase cannot exceed \$1.41 per winter month. Since increasing indoor temperatures by 0.65°F costs the average weatherized household \$1.09 per winter month, the average net gain from the weatherization-induced increase in warmth does not exceed \$0.32 per winter month. Assuming 6 full winter months in Michigan, this implies an upper bound on the annual welfare gain of roughly \$1.92 from higher indoor temperature. In sum, this bounding exercise suggests that the welfare gains from any efficiency-induced rebound in heating demand are very small, likely less than 0.1% of the energy expenditure savings. Put another way, the efficiency induced rebound in demand for indoor heating appears to be inconsequential in this setting. Our methodology also allows us to calculate the potential welfare gains from much larger increases in indoor temperatures. Note that even a 10 degree F increase in indoor temperatures would lead to a welfare gain of less than \$30 per year.

6 Interpretation

This section evaluates the returns to energy efficiency investments in residential retrofits from private and social perspectives. We also provide an alternative summary measure of cost effectiveness, which is the cost per ton of carbon abated. To conduct this part of the analysis, we use the experimental estimates of the effect of weatherization on energy consumption to estimate average energy savings.

6.1 Returns on Residential Energy Efficiency Investments

Panel A of Table 7 evaluates the internal rate of return (IRR) on investment from a private perspective. More precisely, we report the discount rate at which the discounted value of average avoided energy expenditures exactly equals the average upfront investment. While these investments were free to participating households, it is nevertheless informative to estimate the private returns if

households had been responsible for the upfront costs because most households that consider the exact same investments do not qualify for WAP. That is in the broader population, the IRR is a critical factor in determining take-up of energy efficiency investments.

Column (1) computes the internal rate of return using the average upfront investment costs³⁷ and the average reductions in annual energy expenditures projected by the WAP program audit. These savings are monetized or valued by using the average retail residential natural gas and electricity prices in 2013 as reported by the EIA.³⁸ Over a range of time horizons, the estimated IRR is quite high, as they should be given the investment rule that required projected savings exceed costs. The rate of return associated with the savings-weighted average lifespan (i.e., 16 years) is approximately 12%. By this measure, efficiency investments supported under WAP appear to be very attractive investments that greatly exceed typical returns available in equity, real estate, and bond markets.

The second column of Table 7 replaces the projected savings with an estimate of realized welfare benefits. This is calculated as the sum of the actual energy savings (in monetary terms), derived from the randomized encouragement design, and the estimated upper bound of the monetized value of the net welfare gain from higher indoor temperatures.³⁹ The estimate of realized savings is about \$235, which is only 36% ($=\$235/\651) of the average projected annual monetary savings.⁴⁰ When the upper bound of the monetized value of the higher indoor temperatures is added to the energy savings, the annual benefits are roughly \$237. Using this measure of annual benefits, the IRR is -10.5% for the 10 year horizon, -2.2% for 16 years and 0.3% for 20 years. This finding of low, indeed

³⁷The upfront investment costs are calculated as the imputed average installation, construction, and materials costs among compliers.

³⁸The average monthly residential gas price in Michigan is \$10.46/MMBtu. The average residential electricity price is \$0.11/kWh.

³⁹By adding the upper bound, we use a generous estimate of the welfare gain associated with heating demand rebound as the lower bound is zero. Because the estimated rebound effect is so small, the decision to use the upper versus lower bound estimate has little impact on these calculations. It is important to note that there may be additional benefits and costs to households that received weatherization assistance that are not reflected in increased indoor air temperatures. For example, consumers may realize health benefits associated with reduced draftiness. At the same time, there are speculations in the public health literature that the program may increase health risks as less drafty homes may trap indoor air pollutants such as radon.

⁴⁰The experimental estimate of average energy savings (measured in MMBtu) is 39% (17.2 MMBtu/43.7 MMBtu) of projected energy savings. This is modestly larger than our estimate of the ratio of realized annual monetary savings to projected annual monetary savings, because 2013 energy prices are modestly lower than the prices assumed by the audits.

negative, returns suggests that, at least for residential home retrofits, there may not be much of an efficiency gap to explain as investments with these returns are infrequently taken-up in the broader economy.

Panel B conducts a similar exercise but adds the value of avoided emissions to the benefit side. Avoided emissions of CO₂ are valued at \$38 per ton of (Greenstone et al., 2013). Nitrogen oxide and sulfur dioxide emissions from residential gas consumption are valued at \$250 per ton and \$970 per ton, respectively (Muller and Mendelsohn, 2009).⁴¹ The IRRs for the 10, 16, and 20 year horizons are -8.8%, -0.8%, and 1.5%, respectively.

The societal perspective is an especially important one to judge these investments, because a wide set of policies encourage residential energy efficiency investments. Panel C reports estimates of the social internal rate of return that are calculated by using the social value of avoided fuel consumption, rather than the retail prices, to monetize the energy savings and adding the monetized value of the avoided emissions (as in Panel B). The retail electricity and natural gas prices used in Panels A and B include a fixed cost component that covers connection costs; these fixed costs are not avoided when efficiency investments reduce residential natural gas and electricity consumption (Davis and Muehlegger, 2010). Since most electricity and natural gas distribution utilities are subject to cost-plus regulation, the fixed costs recovery will be shifted to other customers; thus, households that reduce energy consumption receive a transfer from other households. We use the average of the 2013 spot prices (\$3.73/MMBtu) set at the Henry Hub distribution point, a standard reference price for natural gas in North America, to reflect the true marginal cost of natural gas over the lifetime of the measures.⁴² To value reductions in electricity consumption, we use the average wholesale electricity price in the midwest electricity market.

These adjustments lead to a meaningful decrease in the IRR, relative to the IRRs in Panels A and B. For example, the social IRR for the 16 year time horizon is -9.5%; it is -20.0% and -6.1% for

⁴¹We assume that burning natural gas emits 116.39 lbs CO₂ per mmbtu, 0.092 lbs NO_x per mmbtu, and 0.000584 lbs SO₂ per mmbtu. We assume a marginal operating emissions rate of 0.916 lbs CO₂ per kWh in the Midwest power sector (Callaway et al., 2015). Sulfur dioxide and nitrogen oxide emissions from residential natural gas consumption are monetized using the median marginal damage estimates in (Muller and Mendelsohn, 2009). NO_x and SO₂ emissions from electricity generation are subject to a (barely) binding cap.

⁴²This is lower than the average gas recovery charge reported by the utility in regulatory proceedings over the post-weatherization period (\$5.54), but the utility's recovery charge rolls in contract positions.

the 10 and 20 year horizons, respectively. Overall, these residential energy efficiency investments have a negative rate of return across all reasonable time horizons.⁴³

An alternative method to summarize the return on WAP energy efficiency investments is to estimate the cost per ton of CO₂ avoided. This is calculated as the ratio of the net cost of the investments (i.e., the annual rental cost of the upfront investment less the value of annual energy savings) and the tons of CO₂ emissions reduced per year. Panel D uses a 3% discount rate to calculate the annual rental cost of capital, while Panel E uses 7%. Both panels report results based on assumptions of 10, 16, and 20 year lifespans for the investments and use wholesale energy prices as above.

As in Panels A through C, the conclusions differ starkly depending on whether one uses *projected* or *realized* energy savings. Using projected energy savings values generated by the NEAT program audits and the 16 year lifespan, the cost per ton of CO₂ avoided is -\$19 with a 3% discount rate and \$14 with a 7% discount rate. With these estimates, the energy efficiency investments would be judged beneficial because they are less than the United State Government's official value of the social cost of carbon (i.e., the monetized value of the avoided damages from a ton of abated CO₂ emissions) of \$38. Indeed, the Panel D cost of -\$19 supports the claims that energy efficiency investments are win-win because the engineering estimates suggest that it is possibly to simultaneously reduce households' energy costs and carbon emissions.

However, the estimates that are based on actual energy savings again tell a different story. When the experimental estimates of actual natural gas and electricity savings are used in column (2), the analogous costs per ton of CO₂ avoided are \$329 (3% discount rate) and \$484 (7% discount rate). These costs exceed the United States Government's social cost of carbon by roughly an order or magnitude, suggesting that, at least in this study's context, residential energy efficiency investments are an inefficient approach to mitigating climate change.

⁴³An alternative approach would be to calculate the social internal rate of return from the government's perspective. This approach would require accounting for the social cost of public funds, administrative costs of the program, and costs of the encouragement design.

6.2 What Explains the Low Rate of Return on These Efficiency Investments?

It is natural to ask why the returns to residential energy efficiency investments are so low. After all, WAP is designed so that the only measures implemented are ones with projected savings to costs ratios greater than one.

An important factor leading to negative returns on investment is the incomplete realization of projected energy savings. The projected savings are about 2.5 times the preferred experimental savings estimate. Further, the projected savings are roughly 4 times the quasi-experimental estimates of energy savings. There are relatively few ex-post estimates in the academic peer-reviewed literature, with Davis et al. (2014) and Dubin et al. (1986) serving as notable exceptions.⁴⁴ Both of those papers similarly find low realization rates, although they largely attribute it to behavioral responses (i.e., the rebound effect) which we have shown plays at most a minor role in this paper's setting. Moreover because energy efficiency programs are implemented by regulated utilities, there are a number of regulatory filings that estimate ex post program savings and it is hardly unusual for them to find that the programs deliver estimated savings considerably lower than projected.⁴⁵

Having ruled out the rebound effect as the primary explanation for the gap between projected and realized energy savings, we conclude that the efficiency audit tool must systematically overstate the real returns to these investments. Along these lines, we explore some alternative sources of this bias. First, we compare the distribution of temperatures observed during our study period against the typical weather patterns on which engineering calculations are based. Although we do observe some moderate spells in our timeframe, on average we observe colder than average temperatures and higher than average degree day measures in our sample; these colder temperatures should lead to greater than average savings.

A second potential source of bias concerns the over-statement of baseline energy use. Several studies and utility reports have documented how software-based energy analysis of existing homes

⁴⁴A recent working paper with a nonexperimental design finds that, for residential energy efficiency investments similar to those in the WAP program, projected energy savings are about 2 times larger than actual energy savings (Allcott and Greenstone, 2015).

⁴⁵See, for example, CPUC (2015), which summarizes findings across California programs implemented from 2010-2012 and reports realization rates from below 30% to above 100%, though the estimated savings are mainly derived from observational approaches that produce associational evidence. Findings appear similar in other jurisdictions. An early analysis of several thousand home retrofits in Hood River, Oregon similarly found realization rates below 40% (Hirst, 1987).

tends to over-predict pre-retrofit energy use and retrofit energy savings.⁴⁶ Indeed, we found in our data that the NEAT program predicts baseline natural gas consumption that exceed actual consumption by more than 25% prior to weatherization. This suggests that the auditing tool could be under-estimating the efficiency properties of the average home prior to weatherization, which may partly explain the over-statement of the benefits of upgrading to a given efficiency standard.

Overall, our findings suggest that the NEAT audit tool over-estimates returns by a significant margin. Further, this overestimation of savings does not appear to be due to behavioral responses. This is an important finding in its own right; NEAT is widely used by state and local WAP sub grantees, utility companies, and home energy audit firms.⁴⁷

7 Conclusion

We conducted a large-scale randomized encouragement design experiment on a sample of over 30,000 households presumptively eligible for participation in WAP in the state of Michigan. Approximately one quarter of these households were randomly assigned to a treatment group that was encouraged to apply for the program and received significant application assistance. The control households were free to apply for WAP but were not contacted or assisted in any way by our team. We also analyze corroborating evidence from a quasi-experimental analysis covering over twice as many weatherizations as well as a survey of indoor conditions at weatherized and unweatherized homes.

We document three primary findings. First, the aggressive encouragement efforts were disappointing. This encouragement increased take-up rates from less than 1% in the control group to about 6% at a cost of over \$1,000 per weatherized household. Second, we find that WAP participation reduced energy consumption by 10-20% among participating households. However, the upfront cost of the energy efficiency investments are about twice the cost of the realized energy savings. Further, the projected savings are about 2.5 times the actual savings. Third, while the

⁴⁶For example, a recent report found that modeling software consistently overestimated energy consumption; mean modeled total annual use was 40% greater than billed use (SBW, 2012).

⁴⁷While more sophisticated building simulation models exist, an appeal of NEAT is that it can be inexpensively used by the thousand of implementers who have a wide range of skills and technical training. Indeed, the DOE cites one of the primary benefits of NEAT as its accessibility to non-technical users (EERE, 2010). While more complex models exist, they are very likely to be more expensive to use.

modest energy savings might be attributed to the rebound effect, when demand for energy end uses increases as a result of greater efficiency, the paper fails to find evidence of economically or statistically significant increases in indoor temperature at weatherized homes.

Overall, the energy efficiency investments we evaluate are poor performers on average across a variety of metrics. From a household's perspective, the annual internal rate of return that would rationalize these efficiency investments is -2.2%. The household's perspective differs from society's because it fails to recognize the benefits of greenhouse gas and local pollutant emissions reductions and because the retail prices for natural gas and electricity exceed their marginal costs of delivery. Accounting for these two factors, the annual social internal rate of return that would justify these investments is -9.5%, which is even less favorable. Finally, we also calculate the average cost per ton of avoided CO_2 under a range of assumptions. The most plausible estimates are approximately \$329/ton, which is about an order of magnitude larger than the U.S. government's estimate of the monetized benefits of avoided emissions (i.e., the social cost of carbon) of roughly \$38.

This study demonstrates that the returns to common residential energy efficiency investments are negative both privately and socially among low-income households in Michigan. The results are striking because Michigan's cold winters and the likelihood that the weatherized homes were not in perfect condition suggests that it may have been reasonable to expect high returns in this setting. Regardless of one's priors, this paper underscores that it is critical to develop a body of credible evidence on the true, rather than projected, returns to energy efficiency investments in the residential and other sectors. The findings also suggest that the last several decades may have seen too much investigation into the why of the energy efficiency gap and not enough into whether there really was one.

From a policy perspective, WAP does not appear to pass a conventional cost-benefit test, although its full-set of goals may not be reflected in such tests. On the broader question of optimal climate change policy, this paper's findings indicate that residential energy efficiency retrofits are unlikely to provide the least expensive carbon reductions. Future research should examine whether the real world returns to energy efficiency investments differ so starkly from engineering projections in other settings.

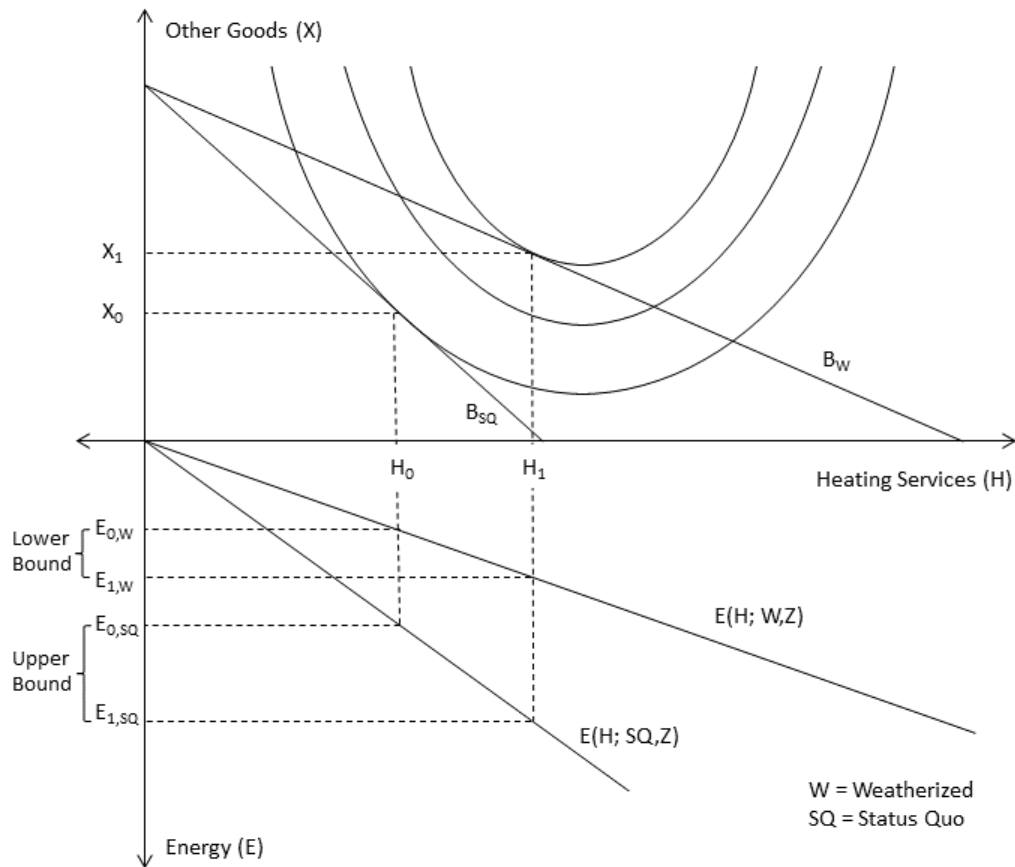
References

- Allcott, H. and Greenstone, M. (2012). Is there an energy efficiency gap? *The Journal of Economic Perspectives*, 6(1):3–28.
- Allcott, H. and Greenstone, M. (2015). Maximizing money vs. utility: Measuring the welfare effects of energy efficiency programs. *Mimeo graph*.
- Barbose, G. L., Goldman, C. A., Hoffman, I. M., and Billingsley, M. A. (2013). The future of utility customer-funded energy efficiency programs in the United States: projected spending and savings to 2025. *Energy Efficiency Journal*, 6(3):475–493.
- Callaway, D., Fowlie, M., and McCormick, G. (2015). Location, location, location: the variable value of renewable energy and demand-side efficiency resources. *Working paper*.
- CPUC (2015). 2010 to 2012 energy efficiency annual progress evaluation report. Technical report, California Public Utilities Commission.
- Davis, L. (2008). Durable goods and residential demand for energy and water: evidence from a field trial. *RAND Journal of Economics*, 39(2):530–546.
- Davis, L., Fuchs, A., and Gertler, P. (2014). Cash for coolers: evaluating a large-scale appliance replacement program in Mexico. *American Economic Journal: Economic Policy*, 6(4):207–238.
- Davis, L. and Muehlegger, E. (2010). Do Americans consume too little natural gas? An empirical test of marginal cost pricing. *The RAND Journal of Economics*, 41(4):791–810.
- Dubin, J. A., Miedema, A. K., and Chandran, R. V. (1986). Price effects of energy-efficient technologies: a study of residential demand for heating and cooling. *The RAND Journal of Economics*, 17(3):310–325.
- Dyson, M., Borgeson, S., Michaelangelo, T., and Callaway, D. (2014). Using smart meter data to estimate demand response potential, with application to solar energy integration. *Energy Policy*, 73:607–619.
- EERE (2010). Review of selected home energy auditing tools: in support of the development of a national building performance assessment and rating program. Technical report, U.S. Department of Energy, Energy Efficiency & Renewable Energy.
- EIA (2015). Natural gas monthly. Technical report, Energy Information Administration.
- Fowlie, M., Greenstone, M., and Wolfram, C. (2015). 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*.
- Friedman, D. (1987). Cold houses in warm climates and vice versa: A paradox of rational heating. *Journal of Political Economy*, 95(5):1089–97.
- Gerarden, T. D., Newell, R. G., and Stavins, R. N. (2015). Assessing the energy-efficiency gap. Technical report, Harvard Environmental Economics Program.

- Gillingham, K., Kotchen, M. J., Rapson, D. S., and Wagner, G. (2013). Energy policy: The rebound effect is overplayed. *Nature*, 493(7433):475–476.
- Gillingham, K. and Palmer, K. (2014). Bridging the energy efficiency gap: policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy*, 8(1):18–38.
- Graff Zivin, J. and Novan, K. (2015). Upgrading efficiency and behavior: electricity savings from residential weatherization programs. *Working paper*. Paper can found at arefiles.ucdavis.edu.
- Greenstone, M., Kopits, E., and Wolverton, A. (2013). Developing a social cost of carbon for US regulatory analysis: a methodology and interpretation. *Review of Environmental Economics and Policy*, 7(1):23–46.
- Hirst, E. (1987). Cooperation and community conservation comprehensive report. Technical report, Oak Ridge National Laboratory.
- ICF (2014). EPA’s 111(d) clean power plan could increase energy efficiency impacts, net benefits, and total value. White paper and webinar. Technical report, ICF International.
- IEA (2013). Redrawing the energy-climate map. Technical report, International Energy Agency.
- Jacobsen, G. D. and Kotchen, M. J. (2013). Are building codes effective at saving energy? Evidence from residential billing data in Florida. *Review of Economics and Statistics*, 95(1):34–49.
- Joskow, P. L. and Marron, D. B. (1992). What does a negawatt really cost? Evidence from utility conservation programs. *The Energy Journal*, 13(4):41–74.
- Loftus, P. J., Cohen, A. M., Long, J. C. S., and Jenkins, J. D. (2015). A critical review of global decarbonization scenarios: what do they tell us about feasibility? *Wiley Interdisciplinary Reviews: Climate Change*, 6(1):93–112.
- McKinsey & Company (2009). Unlocking energy efficiency in the U.S. economy. Technical report, McKinsey Global Energy and Materials.
- Metcalf, G. E. and Hassett, K. A. (1999). Measuring the energy savings from home improvement investments: evidence from monthly billing data. *The Review of Economics and Statistics*, 81(3):516–528.
- Muller, N. and Mendelsohn, R. (2009). Efficient pollution regulation: getting the prices right. *The American Economic Review*, 99(5):1714–1739.
- Radnofsky, L. (2010). A stimulus project gets all caulked up. *Wall Street Journal*.
- SBW (2012). 2010-2012 pge and sce whole house retrofit program process evaluation study. Technical report.
- Schwarz, P. M. and Taylor, T. N. (1995). Cold hands, warm hearth? Climate, net takeback, and household comfort. *The Energy Journal*, 16(1):pp. 41–54.
- Thorpe, D. (2013). *Energy Management in Buildings: The Earthscan Expert Guide*. Routledge.

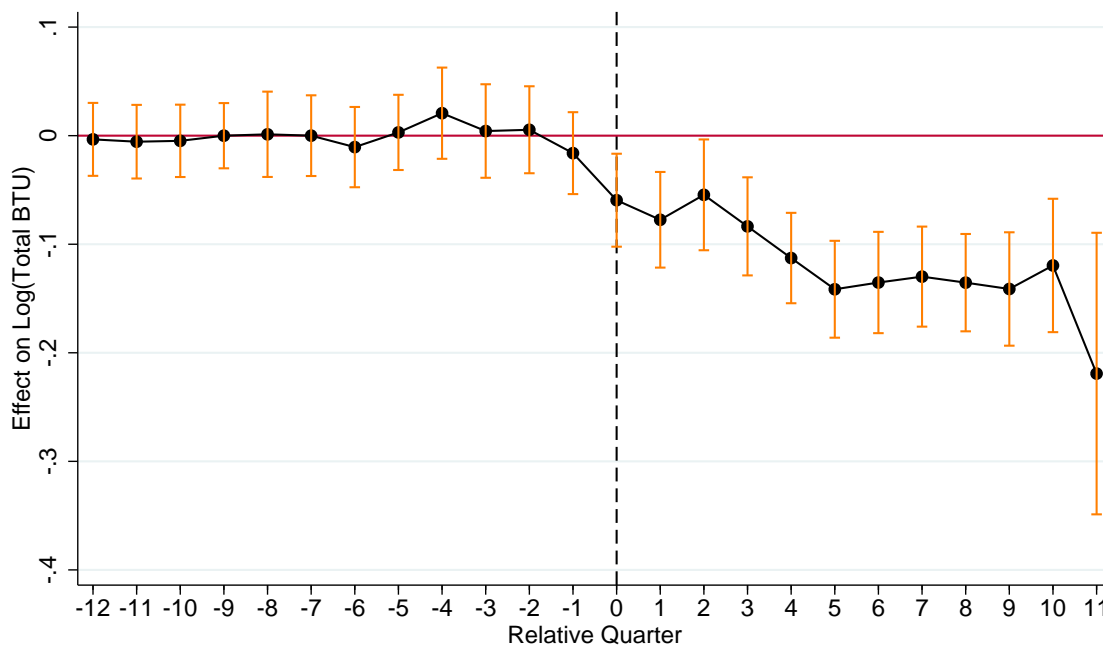
8 Figures

Figure 1: Household-level re-optimization in response to an efficiency improvement



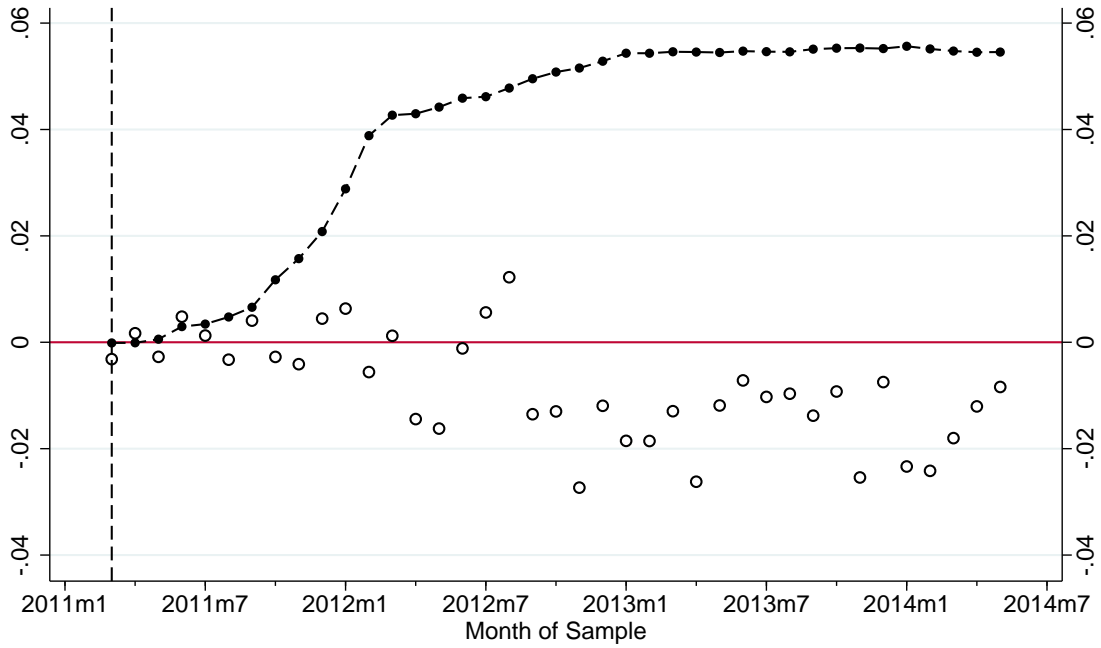
Note: Budget constraints and indifference curves of a representative consumer are plotted in the top quadrant. A linear relationship between heating services and building energy consumption (over a range where heating demand is strictly positive, and heating technology constraints do not bind) are plotted in the bottom quadrant. Please see text for details.

Figure 2: Event study analysis: Matched quasi-experimental sample



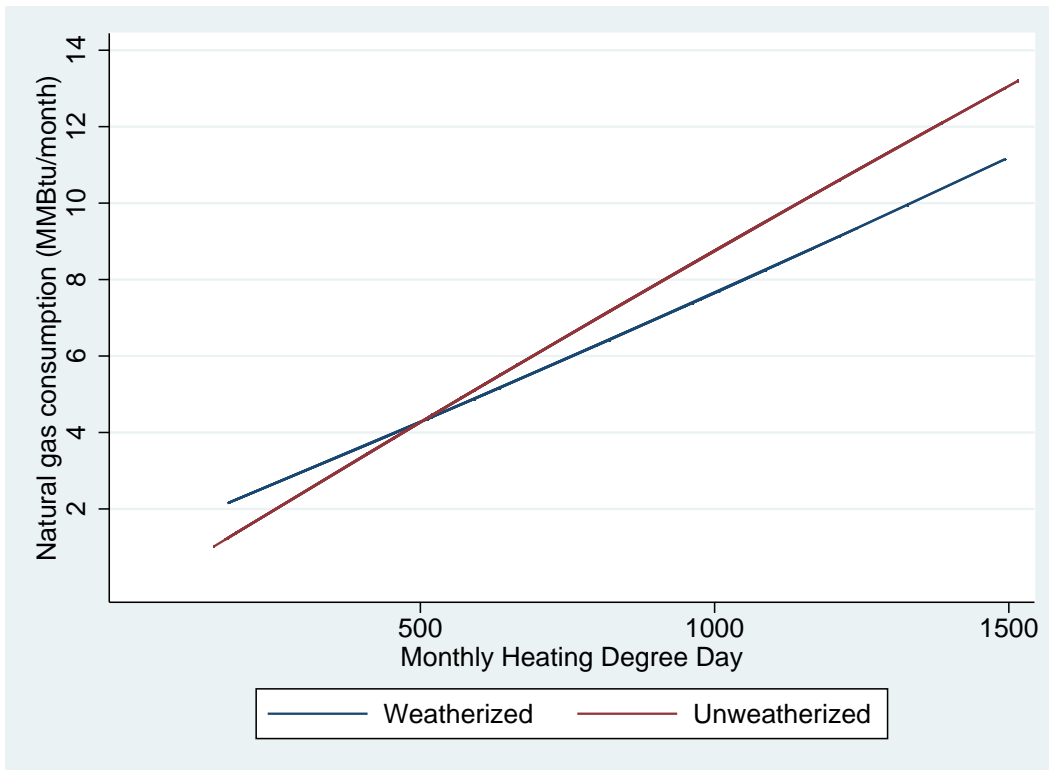
Notes: This figure reports estimated weatherization effects by quarter before and after the weatherization was completed based on the quasi-experimental estimates reported in column (1) of Table 3. The time zero effect captures energy consumption in the month of the weatherization. The effects are also allowed to vary by the realized weather in the quarter. See text for details.

Figure 3: Effect of encouragement on participation and energy consumption



Notes: This figure provides an overview of the local average treatment effect estimates. The broken line tracks the cumulative difference in participation rates across the encouraged and control groups. The circular markers plot the monthly estimates of the intent to treat effects on household energy consumption (in logs).

Figure 4: Building energy performance at weatherized versus unweatherized homes



Notes: This figure plots the estimated relationship between monthly natural gas consumption and heating degree days during winter months at weatherized and unweatherized households, respectively. See Equation (3) in the text.

9 Tables

Table 1: Randomized encouragement intervention

| Encouragement activity | |
|-------------------------------|---------|
| Encouraged group (households) | 8,648 |
| Initial home visits | 6,694 |
| Robo-calls | 23,500 |
| Personal calls | 9,171 |
| Follow up appointments | 2,720 |
| Average cost/encouraged hh | \$55.00 |

Note: The table summarizes efforts to encourage a group of Michigan households to take up weatherization assistance. These households were selected randomly from a sub-population of households who were located in the service territory of our partner utility and presumptively eligible based on ex ante available income information.

Table 2: Differences in Sample Means Between Groups of Households

| | Experimental encouraged (1) | Experimental control (2) | (1) - (2) (3) | All weatherized (4) | Unweatherized applicants (5) | (4) - (5) (6) |
|-----------------------------------------------------------------|-----------------------------------|--------------------------------|------------------|---------------------------|------------------------------------|---------------------|
| Panel A: Pre-treatment period monthly energy consumption | | | | | | |
| Winter gas (MMBtu) | 10.40 | 10.38 | 0.02 (0.07) | 9.88 | 11.63 | -1.75** (0.16) |
| Summer gas (MMBtu) | 2.84 | 2.79 | 0.06 (0.03) | 1.80 | 2.16 | -0.36** (0.05) |
| Winter electricity (MMBtu) | 2.12 | 2.10 | 0.02 (0.02) | 2.24 | 2.30 | -0.06 (0.04) |
| Summer electricity (MMBtu) | 2.17 | 2.17 | 0.00 (0.02) | 2.23 | 2.20 | 0.03 (0.04) |
| Panel B: Demographics and dwelling characteristics | | | | | | |
| Household income (\$) | | | | 19,617 | 17,509 | 2,108** (417.16) |
| Percent of poverty (%) | | | | 115 | 104 | 11** (1.96) |
| Household size (#) | | | | 2.56 | 2.47 | 0.09 (0.06) |
| Children (share of households) | | | | 0.24 | 0.15 | 0.10** (0.01) |
| Reported disability (share of households) | | | | 0.04 | 0.03 | 0.01* (0.00) |
| Elderly (share of households) | | | | 0.23 | 0.13 | 0.09** (0.01) |
| Heat with natural gas (share) | | | | 0.79 | 0.57 | 0.22** (0.01) |
| Age of home (years) | | | | 59.15 | 62.08 | -2.93 (5.34) |
| Households | 7,549 | 21,339 | | 2,074 | 2,973 | |

Note: Columns numbered (1), (2), (4) and (5) report average values. Columns (3) and (6) report differences in means (with standard errors in parentheses). Columns (1) and (2) report sample means for the randomized encouraged and the experimental control groups, respectively. Column (4) reports the sample means for all weatherized households in the quasi-experimental sample while column (5) reports the the sample means for households in the quasi-experimental sample that applied for weatherization but did not receive assistance as of April 2014. Household counts summarize energy consumption data. Panel B focuses exclusively on program applicants because this demographic data is not available for the majority of households in the experimental sample that did not apply to the program.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 3: Quasi-experimental estimated impacts of weatherization on household energy consumption

| Panel A: Dependent variable is monthly energy consumption (in logs) | | | | | | |
|----------------------------------------------------------------------------|-------------------|-------------------|----------------------------|-------------------|-------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| WAP | -0.08** (0.01) | -0.09** (0.01) | -0.08** (0.01) | -0.09** (0.01) | -0.10** (0.01) | -0.10** (0.01) |
| Average consumption control group (MMbtu/month) | | 8.48 [8.29] | | 12.20 [10.36] | | 9.79 [7.32] |
| month-of-sample FE | Y | N | Y | N | Y | N |
| month-of-sample x county FE | N | Y | N | Y | N | Y |
| P-score matched sample | N | N | N | N | Y | Y |
| Adjusted R-squared | 0.85 | 0.86 | 0.83 | 0.83 | 0.80 | 0.81 |
| Households | 5,013 | 5,013 | 3,334 | 3,334 | 3,404 | 3,404 |
| Observations | 282196 | 282196 | 183353 | 183353 | 188287 | 188287 |
| Panel B: Present value of (discounted) savings | | | | | | |
| Time Horizon | 3 percent | | Discount rate 6 percent | | 10 percent | |
| 10 years | \$1,321 | | \$1,140 | | \$952 | |
| 16 years | \$1,946 | | \$1,565 | | \$1,212 | |
| 20 years | \$2,304 | | \$1,777 | | \$1,319 | |

Note: Panel A reports estimates of the reduction in monthly energy consumption following weatherization. The dependent variable is the log of monthly household energy consumption (electricity and natural gas) measured in MMBtu. All columns include household-by-month-of sample fixed effects. Columns (1) and (2) use data from all weatherization applicants while columns (3) and (4) use a sample limited to implementing agencies that participated in the experiment as well as applicants that applied after the encouragement intervention was initiated. Columns (5) and (6) report estimates comparable to columns (1) and (2) reweighted by the propensity score. Standard errors (in parentheses) are clustered at the household level. Panel B reports the net present value of energy savings implied by the preferred estimate reported in column (6). Reductions in energy bills associated with the estimates in column (6) are assumed to accrue over the life of the measure using a range of discount rates and assumed time horizons.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 4: Randomized encouragement: return on effort

| | Application (1) | Efficiency audit (2) | Weatherization complete (3) |
|---------------|------------------------|----------------------------|-----------------------------------|
| Base Rate | 0.02** (< 0.01) | 0.01** (< 0.01) | 0.01** (< 0.01) |
| Encouragement | 0.13** (< 0.01) | 0.05** (< 0.01) | 0.05** (< 0.01) |
| Households | 28889 | 28889 | 28889 |

Note: The table shows the effect of our encouragement on program applications, efficiency audits, and weatherization. Indicators of program participation status are regressed on an encouragement indicator and a constant. The unit of observation is a household.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 5: Experimental estimated impacts of weatherization on household energy consumption

| Panel A: Dependent variable is monthly energy consumption (in logs) | | | | |
|----------------------------------------------------------------------------|-------------------|------------------|-------------------|-----------------|
| | Total Energy | | Gas | Electricity |
| | (1) | (2) | (3) | (4) |
| | OLS-FE | IV-FE | IV-FE | IV-FE |
| WAP | -0.10** (0.01) | -0.20* (0.08) | -0.21** (0.08) | -0.10 (0.10) |
| Imputed baseline consumption MMbtu/month | | 7.52 | 6.38 | 2.19 |
| F-statistic | . | 267.41** | 260.10** | 266.78** |
| Households | 27,990 | 27,229 | 26,054 | 27,115 |
| Observations | 1662781 | 1653583 | 1528526 | 1638337 |
| Panel B: Present value of (discounted) savings | | | | |
| Time Horizon | Discount rate | | | |
| | 3 percent | 6 percent | 10 percent | |
| 10 years | \$2,003 | \$1,728 | \$1,443 | |
| 16 years | \$2,949 | \$2,373 | \$1,837 | |
| 20 years | \$3,493 | \$2,693 | \$1,999 | |

Note: Dependent variable measures log of monthly household energy consumption. Panel A reports regression coefficients. With the exception of the first column, all specifications are estimated using 2SLS. Standard errors (in parentheses) are clustered by household. Panel B reports the net present value of energy savings implied by the savings estimates in columns (3) and (4) using a range of discount rates and assumed time horizons. Reductions in energy bills associated with the estimates are assumed to accrue over the life of the measure. All regressions include month-of-sample and household-month fixed effects.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 6: Indoor temperature survey results

| | Thermometer | | Thermostat | |
|---------------------------|-------------------|-------------------|-------------------|-------------------|
| | (1) | (2) | (3) | (4) |
| Base temperature | 72.36** (0.95) | 72.17** (1.24) | 69.26** (0.96) | 68.91** (1.29) |
| Weatherized home | 0.57 (0.41) | 0.65 (0.44) | -0.57 (0.29) | -0.56 (0.33) |
| Heating Degree Days | -0.16** (0.03) | -0.15** (0.04) | 0.04 (0.03) | 0.05 (0.04) |
| Propensity Score Weights? | N | Y | N | Y |
| R-squared | 0.02 | 0.02 | 0.01 | 0.01 |
| Observations | 1359 | 1359 | 899 | 899 |

Note: The table reports measured indoor temperature differentials across weatherized (WAP) and unweatherized households. Columns (1) and (2) have the indoor thermometer temperature reading as a dependent variable while columns (3) and (4) use the survey thermostat readings. Columns (2) and (4) are weighted so that surveyed population better represents total quasi-experimental sample. Standard errors clustered at the household level.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 7: Estimated returns on investments in energy efficiency

| Time horizon | Ex ante projections (NEAT) (1) | Empirical estimates (2) |
|--------------------------------------------------------------------------------------------------|--------------------------------|-------------------------|
| Panel A: Private internal rate of return | | |
| 10 years | 7.0% | -10.5% |
| 16 years | 11.8% | -2.2% |
| 20 years | 12.8% | 0.3% |
| Panel B: Private internal rate of return, adding avoided emissions damages | | |
| 10 years | 11.3% | -8.8% |
| 16 years | 15.5% | -0.8% |
| 20 years | 16.4% | 1.5% |
| Panel C: Social internal rate of return | | |
| 10 years | -1.0% | -20.0% |
| 16 years | 5.4% | -9.5% |
| 20 years | 7.0% | -6.1% |
| Panel D: CO_2 abatement cost - 3 percent discount (\$/ton CO_2) | | |
| 10 years | \$29 | \$552 |
| 16 years | -\$19 | \$329 |
| 20 years | -\$35 | \$255 |
| Panel E: CO_2 abatement cost - 7 percent discount (\$/ton CO_2) | | |
| 10 years | \$61 | \$701 |
| 16 years | \$14 | \$484 |
| 20 years | \$0 | \$417 |

Note: All calculations use the average retrofit cost of \$4,581. This is the ex post realized average expenditure per household as reported by the implementing agencies. Column (1) reflects engineering projections of annual energy savings. In Panels B and C, column (1) also incorporates the value of estimated emissions reductions (valued using a social cost of carbon value of \$38 per ton CO_2 and values for avoided local pollutants as described in the text). Column (2) replaces the engineering estimates of energy savings with our experimental estimates of energy savings. Column (2) also incorporates the upper bound on the net welfare gain from increased heating demand using our very small and statistically insignificant point estimate of the upper bound on the efficiency-induced increase in welfare associated with warmer indoor air temperatures.

10 Appendix

10.1 Background on Experimental Research Design

Our experimental research required close collaboration with both our partner utility and the community action agencies that serve households in the utility’s service territory. We established these formal partnerships with two large implementing agencies. The experimental research design includes only households living in the counties served by our partner agencies.

To implement the recruitment and assistance, we issued a request for proposals from organizations specializing in grassroots organizing and mobilization. We selected FieldWorks LLC, a private company that specializes in designing communications strategies, running neighborhood canvassing operations, and managing outreach campaigns. Prior to our project, their staff had generated millions of phone calls and knocked on millions of doors in previous engagements. We found them to be highly focused and innovative in their approach to educating households about weatherization assistance, and helping individuals to navigate the application and enrollment process.

To select the study sample, we first identified census blocks within the counties served by our partner utility and implementing agencies that had high rates of home ownership, high rates of natural gas heating, and household incomes that would qualify for weatherization assistance. FieldWorks then used additional data purchased from InfoUSA to identify specific households within these census blocks that they believed met the income qualifications and owned their home. FieldWorks also focused exclusively on relatively dense neighborhoods to reduce the costs of their canvassing operations. These households comprised our study population. From this population, we randomly selected 34,161 households for our study sample. Approximately one quarter of these households were randomly assigned to our encouragement “treatment.” These treated households were encouraged to apply for WAP and offered extensive application assistance. We provide no recruitment, outreach, or assistance to the remaining 25,513 households. For households assigned to the control groups, we simply observe energy consumption and program participation decisions.

We worked closely with FieldWorks to develop a persuasive recruit and assist strategy for the approximately one quarter of the sample households that were assigned to the treatment group. Fieldworks personnel, together with our field manager, coordinated field operations in Michigan. Individuals selected from the targeted communities were hired to conduct the bulk of the door-to-door canvassing and outreach activities.

10.2 Modeling Selection into the Weatherization Assistance Program

Although our household by month-of sample fixed effects will control for any time invariant differences between households, differences in time-varying factors could confound the quasi-experimental comparisons across weatherized and unweatherized households. This creates the potential for bias in the quasi-experimental estimates. We thus take some additional steps to ensure that the comparison groups are similar along dimensions we can observe. To the extent that observable differences across households are correlated with omitted, time-variant determinants of energy consumption, balancing covariates across comparison groups can mitigate omitted variable bias.

We estimate the probability of receiving weatherization assistance as a function of observable covariates that presumably play a role in determining program take-up: number of children in the home, elderly residents, income level, primary heating fuel, and trends in historic energy use. Estimated propensity scores are used to construct inverse probability weights. Equation (1) is

re-estimated as a weighted regression, using inverse probability weights for the control households. Estimated treatment effects are robust to alternative specifications of the selection equation.

10.3 Fuel-Specific Treatment Effects

In the paper, we report the quasi-experimental results that define the dependent variable in terms of total monthly energy consumption (i.e. natural gas plus electricity). To compute the discounted present value of energy savings, we need to estimate fuel-specific savings in levels. There are two ways to construct these estimates. We can estimate these directly, estimating equation separately for electricity and natural gas, defining the dependent variable in equation in terms of levels. Or we can estimate these two equations using the log transformed dependent variable and calculate levels as a percentage of the control group consumption. Because the timing of the weatherization retrofits vary across participant households, we can at best approximate baseline average consumption.

Appendix Table 2 reports estimation results for specifications in which the dependent variable is defined in levels and in logs. The table also reports the average consumption (natural gas and electricity, respectively) in the control groups in 2012-2014. We use this to approximate the average counterfactual energy consumption. We obtain very similar estimates of monthly reductions in natural gas and electricity consumption using either approach. This is reassuring insofar as our savings estimates are robust to alternative functional form assumptions.

10.3.1 Comparing Experimental and Quasi-Experimental Estimates of Energy Consumption Impacts

We estimate that efficiency improvements reduced monthly energy consumption by 20% on average among those households who were induced by our encouragement to seek out weatherization assistance. This is twice as large as the quasi-experimental estimates of the average effect among all households receiving weatherization assistance from these same agencies. The difference in these point estimates is economically significant, but not statistically significant due to the imprecision of the experimental estimate. This appendix summarizes steps we have taken to investigate possible explanations for this difference.

One possible explanation is that relatively more funds were invested in energy efficiency retrofits at households in our experimental sample. However, when we compare measures of dollars spent and measures installed, we find that the average investment per household was significantly smaller among weatherized households in our encouraged group (as compared to the larger quasi-experimental sample). Thus, it is not the case that differences in energy consumption impacts can be explained by higher investment rates at complier homes.

Another explanation is that the effects of weatherization on energy consumption vary significantly and systematically across households, and that compositional differences in the complier households versus other weatherized households explains the observed differences in average treatment effects. Appendix Table 3 compares households receiving weatherization in the encouraged group with weatherized households that were not part of the experimental sample. The table reveals significant differences along important dimensions. For example, we find that weatherized households in the encouraged group are significantly more likely to report using natural gas as their primary heating source. As mentioned above, we drew our experimental sample from counties associated with high degrees of gas heating so that we could track heating fuel consumption in our natural gas and electricity data. Among weatherized households in the experimental sample, over

96% report using natural gas as the primary heating fuel, 3% use electric heat, and the remaining 1% report heating with a fuel other than natural gas or electricity (i.e., propane, fuel oil, or wood). In contrast, almost 18% of weatherized households outside the experimental sample report using a heating fuel other than electricity or natural gas as their primary heating fuel. Because we do not observe propane, wood, or fuel oil consumption in our data, we observe only a fraction of the weatherization-induced energy savings at these households. This will bias our quasi-experimental estimates towards zero. We also find that weatherized households outside the experimental sample are larger, less likely to receive a furnace replacement as part of the weatherization, and less likely to report an elderly resident as compared to weatherized households in the encouraged group.

To assess whether these observable differences could explain the differences in treatment effect estimates, we estimate a more flexible specification of Equation (1) that allows both average energy consumption in the post-treatment period and the average effect of weatherization on energy consumption to vary along a number of observable dimensions. Appendix Table 4 summarizes the results of estimating these alternative specifications. To make the comparison between quasi-experimental and experimental treatment effect estimates as direct as possible, we use a subset of the quasi-experimental data that includes only those households that applied for weatherization with the two agencies we partnered with in our experiment.

We find that the estimated average treatment effects varies significantly with several of the variables that are distributed differently in the experimental and quasi-experimental groups of weatherized households. In particular, quasi-experimental treatment effect estimates are significantly higher among households that heat with natural gas. We also find that savings are significantly higher among households that report an elderly resident. We find no significant differences in estimated effects in other dimensions such as household size. Column (3) reports a specification that includes variables that are not randomly assigned, but could determine savings (furnace replacement and retrofit costs).

The estimated coefficients in column (2) are used to construct a treatment effect for a household that resembles the average complier household along observable dimensions. This yields an average treatment effect of -0.13. These results suggest that less than a third of the difference between the quasi-experimental and experimental treatment effect estimates can be explained by observable differences across the different groups of weatherized households.

11 Appendix Tables

Table 1: Projected costs, savings, retrofit measures at weatherized households

| | Weatherized households |
|-----------------------------------------------|---------------------------|
| Energy savings (projected) | |
| Natural gas savings: heating (annual MMBtu) | 50.36 (45.96) |
| Electricity savings: cooling (annual MMBtu) | 4.22 (50.45) |
| Projected energy savings/ baseline | 0.46 (0.20) |
| Investment costs and projected savings | |
| Investment cost: reported (\$) | 5,151 (3,137) |
| Investment cost: projected (\$) | 5,306 (2,823) |
| Projected NPV savings (\$) | 10,689 (11,857) |
| Projected savings:investment ratio | 2.01 (2.48) |
| Key measures | |
| Furnace replacement | 0.34 (0.47) |
| Attic insulation | 0.85 (1.06) |
| Wall insulation | 0.44 (1.16) |
| Infiltration reduction | 0.76 (0.43) |
| Households | 1,638 |

Notes: This table summarizes data from all weatherized households that could be exactly matched with audit data. Average values reported, standard deviations appear in parentheses.

Table 2: Quasi-experimental estimated impacts of weatherization on household gas and electricity consumption

| Panel A: Dependent variable is monthly natural gas consumption | | | | | | |
|-----------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| WAP | -0.12** (0.01) | -0.97** (0.08) | -0.12** (0.01) | -1.07** (0.09) | -0.11** (0.01) | -0.97** (0.08) |
| Average consumption control group (MMBtu/month) | | 8.74 [8.48] | | 8.47 [8.25] | | 7.52 [7.10] |
| Trimmed sample | N | N | Y | Y | N | N |
| P-score matched sample | N | N | N | N | Y | Y |
| Dep. variable | log | levels | log | levels | log | levels |
| Adjusted R-squared | 0.85 | 0.86 | 0.85 | 0.85 | 0.84 | 0.84 |
| Households | 3,405 | 3,406 | 2,864 | 2,864 | 3,404 | 3,404 |
| Observations | 183851 | 188291 | 153339 | 156594 | 183848 | 188287 |
| Panel B: Dependent variable is monthly electricity consumption | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| WAP | -0.06** (0.01) | -0.17** (0.03) | -0.05** (0.01) | -0.13** (0.04) | -0.06** (0.01) | -0.12** (0.04) |
| Average consumption control group (MMBtu/month) | | 2.57 [1.86] | | 2.53 [1.77] | | 2.33 [1.44] |
| Trimmed sample | N | N | Y | Y | N | N |
| P-score matched sample | N | N | N | N | Y | Y |
| Dep. variable | log | levels | log | levels | log | levels |
| Adjusted R-squared | 0.66 | 0.55 | 0.64 | 0.53 | 0.63 | 0.63 |
| Households | 4,909 | 4,914 | 3,308 | 3,313 | 3,300 | 3,303 |
| Observations | 276022 | 293548 | 181565 | 194384 | 182113 | 194372 |

Note: Table reports estimates of the reduction in monthly energy consumption following weatherization. The unit of observation is a household-month. The dependent variable is monthly household energy consumption (electricity or natural gas) measured in MMBtu. The coefficient reported is the coefficient on the weatherization indicator, which switches to one from zero after a household weatherization is completed. All specifications include household-by-month fixed effects. Standard errors (in parentheses) are clustered at the household level.

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 3: Differences in sample means between groups of weatherized households

| | Experimental control (1) | Experimental encouraged (2) | (2) - (1) (3) | Other weatherized (4) | (4) - (2) (5) |
|-------------------------------------------------------------------------|--------------------------------|-----------------------------------|-----------------------|-----------------------------|----------------------|
| Panel A: Pre-treatment period monthly energy consumption (MMBtu) | | | | | |
| Winter gas (MMBtu) | 10.42 (3.93) | 9.80 (3.47) | -0.62 (0.32) | 9.82 (3.86) | 0.03 (0.23) |
| Summer gas (MMBtu) | 2.94 (1.47) | 2.97 (1.50) | 0.03 (0.13) | 1.32 (1.62) | -1.65** (0.09) |
| Winter electricity (MMBtu) | 1.99 (0.93) | 2.21 (1.10) | 0.22* (0.09) | 2.28 (1.36) | 0.07 (0.07) |
| Summer electricity (MMBtu) | 2.04 (1.02) | 2.21 (1.11) | 0.16 (0.10) | 2.26 (1.35) | 0.05 (0.07) |
| Panel B: Demographics and dwelling characteristics | | | | | |
| Household income (\$) | 17,048 (8,840) | 19,783 (12,172) | 2,735** (1,016) | 20,041 (12,386) | 259 (708) |
| Household size (# people) | 1.99 (1.30) | 2.37 (1.60) | 0.38** (0.14) | 2.71 (1.73) | 0.34** (0.10) |
| Children (share of hh) | 0.19 (0.40) | 0.27 (0.44) | 0.07 (0.04) | 0.24 (0.43) | -0.03 (0.02) |
| Elderly (share of hh) | 0.28 (0.45) | 0.38 (0.49) | 0.11* (0.04) | 0.17 (0.38) | -0.21** (0.02) |
| Age of home(yrs) | 62.87 (18.80) | 58.92 (20.73) | -3.94* (1.94) | 58.90 (29.21) | -0.02 (1.62) |
| Floor area (sq. ft.) | 1780 (596) | 1734 (594) | -26 (57) | 1736 (739) | 2 (42) |
| Furnace replacement | 0.44 (0.50) | 0.44 (0.50) | -0.00 (0.05) | 0.29 (0.45) | -0.15** (0.03) |
| Gas heat | 0.94 (0.24) | 0.97 (0.18) | 0.03 (0.02) | 0.72 (0.45) | -0.24** (0.02) |
| Reported cost (total) | 5428.24 (2828.30) | 4635.99 (2609.83) | -792.26** (281.74) | 5262.04 (3294.01) | 626.06** (209.28) |
| Proj. savings (MMBtu) | 63.71 (44.11) | 55.36 (41.83) | -8.35* (4.10) | 54.84 (49.00) | -0.53 (2.82) |
| Households | 180 | 436 | | 1,473 | |

Note: Columns numbered (1), (2), and (4) report average values and standard deviations (in parentheses). Columns (3) and (5) report differences in means (standard errors are in parentheses).

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 4: Heterogeneity in weatherization impacts on energy consumption

| Dependent variable is monthly energy consumption (in logs) | | | |
|------------------------------------------------------------|-------------------|---------------------|---------------------|
| | (1) | (2) | (3) |
| WAP | -0.09** (0.01) | < 0.01 (0.02) | 0.04 (0.02) |
| Gas heat x WAP interaction | | -0.10** (0.02) | -0.10** (0.03) |
| Elderly x WAP interaction | | -0.05* (0.02) | -0.04* (0.02) |
| Income x WAP interaction | | < 0.01 < (0.01) | < 0.01 < (0.01) |
| Age of house x WAP interaction | | < -0.01 < (0.01) | < 0.01 < (0.01) |
| Furnace x WAP interaction | | | -0.07** (0.02) |
| Retrofit cost x WAP interaction | | | < -0.01 < (0.01) |
| Adjusted R-squared | 0.83 | 0.84 | 0.84 |
| Households | 3,334 | 2,541 | 1,117 |
| Observations | 183,353 | 141,938 | 64,477 |

Note: The unit of observation is a household-month. The dependent variable is the log of monthly household energy consumption measured in MMBtu. All specifications include un-interacted covariates, household-month fixed effects, and month-year-county fixed effects (not shown). Standard errors (in parentheses) are clustered at the household level.

* Significant at the 5 percent level

** Significant at the 1 percent level