Comparing the effectiveness of weatherization treatments for low-income, American, urban housing stocks in different climates

Article *in* Energy and Buildings · February 2014

DOI: 10.1016/j.enbuild.2013.11.035

[Note: This document is a preprint manuscript for an article that has been accepted for publication (DOI: 10.1016/j.enbuild.2013.11.035) in Energy and Buildings, an Elsevier journal. The final article can be downloaded from ScienceDirect using the following URL: http://www.sciencedirect.com/science/article/pii/S0378778813007329]

Comparing the effectiveness of weatherization treatments for low-income, American, urban housing stocks in different climates[1]

Jonathan L. Bradshaw [a, b, 2,] *, Elie Bou-Zeid [a], and Robert H. Harris [a, c]

[a] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, United States

[b] ENVIRON International Corporation, 4350 North Fairfax Drive, Suite 300, Arlington, VA 22203, United States

[c] ENVIRON International Corporation, 214 Carnegie Center, Princeton, New Jersey, 08540, United States

* Corresponding author. Tel.: +1 207 242 5943. E-mail address: jbradshaw@alumni.princeton.edu (J.L. Bradshaw)

ABSTRACT

This paper presents and demonstrates a method for evaluating how the effectiveness of weatherization treatments varies geographically due to difference in climate and housing stock. American Housing Survey Data was used to describe the low-income urban housing stock in six different cities representing a range of geographical and climatic areas. These data were then used to drive the Home Energy Saver model to simulate current energy consumption and expected energy savings from a combination of

[1] **Abbreviations:** A – Attic insulation; AHS – American Housing Survey; CWP – Conservation Works Program; EIA – Energy Information Administration; GJ – Gigajoule; HES – Home Energy Saver; HDD – Heating Degree Day; LBNL – Lawrence Berkeley National Laboratory; MMBTU – Million British Thermal Unit; MSA – Metropolitan Statistical Area; NWAPE – National Weatherization Assistance Program Evaluation; PRISM – Princeton Scorekeeping Method; RECS – Residential Energy Consumption Survey; S – Air sealing; T – Programmable thermostat; TMY2 – Typical Meteorological Year; WAP – Weatherization Assistance Program; Wx- Weatherization Assistance

[2] **Present Address:** Yang & Yamazaki Environment & Energy Building – MC 4020, 473 Via Ortega, Room 159, Stanford University, Stanford, California, 94305

three weatherization treatments: replacing a standard thermostat with a programmable thermostat, installing attic insulation, and envelope air sealing. Modeled energy savings were compared to observed energy savings. Results show that greater energy saving potential generally exists in cities with colder climates, but the effectiveness of different weatherization treatments also varies with differences in regional housing stock and space conditioning equipment. This study's results and methodology could be used in future research to compare the cost-effectiveness and carbon reductions of potential weatherization programs.

Keywords: weatherization, low-income housing, building energy modeling

1 INTRODUCTION

A substantial amount of energy is consumed to heat and cool houses. In the U.S., residential buildings account for 22% of primary energy consumption, of which space conditioning (i.e., heating and cooling) accounts for 41% [1]. Weatherization treatments can make houses more energy-efficient, which results not only in reduced energy bills, but also in lower carbon emissions, improved air quality [2], job creation, and increased national security [3]. Following the energy crisis of the early 1970s, the Weatherization Assistance Program (WAP) was created in 1976 to help low-income families lower their energy bills by implementing weatherization measures [4]. Low-income households are not only those that could most benefit from lower energy bills, but they are also typically less energy-efficient: low-income houses are on average 20% more energy intensive than non-low-income houses [5], and analysis of a national leakage database determined that leakage is 145% higher in low-income houses than in non-low-income houses [6].

Since WAP's inception, the program has been appropriated approximately $6.5 billion, with an additional $5 billion granted under the American Recovery and Reinvestment Act of 2009 in order to weatherize almost 600,000 houses [4,7]. Should government support for weatherization assistance (Wx) programs continue, it is advantageous to predict where weatherization programs can save the most energy. Prior studies have noted that the design and performance of conditioning systems [8,9] and houses [6,10] varies regionally. Figure 1 demonstrates how space conditioning energy use varies substantially among different Census regions and climate zones, while the amount of energy consumed for water heating, lighting, and appliances remains relatively constant [11].

Because space conditioning energy use varies geographically, it can be expected that retrofit effectiveness will vary as well. Measuring the energy savings expected from a retrofit, however, can prove challenging. The empirical method for measuring energy savings consists of comparing a household's energy consumption before and after retrofitting. These comparisons must be normalized for weather conditions, since energy used for space conditioning depends on outside weather conditions. Because of these and other factors, it is standard practice to use an entire year of energy consumption before and after retrofitting in order to determine energy savings. The industry standard for analyzing these data is the Princeton Scorekeeping Method (PRISM), a statistical model that processes weather data and a year of monthly energy bills to produce a weather-normalized measure of energy consumption [12]. The National Weatherization Assistance Program Evaluation (NWAPE), an evaluation of the measured effectiveness of WAP programs across the country, is currently underway, but the results of this evaluation were unavailable at the time of this study's completion [13].

To facilitate energy modeling when sufficient energy bill or weather data are unavailable, many different building energy simulation programs have been developed since the 1980s [14,15]; however calibrating and validating these models is a topic of ongoing research [16–19]. For example, a recent evaluation of several popular residential energy simulation programs found that the mean difference between observed and modeled natural gas consumption ranged from approximately -21% to 36% [20].

Following weatherization treatments, discrepancies between modeled and observed energy consumption are classified as “rebound effects,” primarily caused by a combination of shortfall (technical estimation error or improper weatherization treatment installation) and take-back (behavioral energy consumption changes triggered by the increased energy efficiency expected after weatherization treatment) [21–23]. These discrepancies will be empirically accounted for in this study, but their underlying causes and categorizations will not be pursued in depth.

Despite such uncertainty surrounding the quantitative accuracy of energy simulation programs, they are still widely employed by energy auditors as they can still prove to be useful qualitative decision-making tools. This study will use energy modeling software to compare weatherization treatment scenarios for different housing stocks and climates. This method is not intended to replace WAP impact assessments, which empirically measure the energy savings realized in retrofitted buildings (e.g., [24–27]). Rather, the goal of this paper is to develop a method to estimate and compare potential weatherization savings in locations where observational data are unavailable.

2 DATA AND METHODOLOGY

2.1 ENERGY AND RETROFIT MODELING

The Home Energy Saver (HES) software was selected for this study to model expected energy consumption and savings gained from retrofitting treatments with publicly available technology. HES is a freely available web-based residential energy audit tool developed and maintained by Lawrence Berkeley National Laboratory (LBNL). HES relies on user input, housing stock statistics, and the building simulation DOE-2 engine to approximate whole house energy consumption, potential energy savings with various retrofit treatments, and the costs of such treatments. HES was selected over other models because it is readily available, comprehensive, and user-friendly. In an evaluation of three top house energy modeling programs—SIMPLE, REM/Rate, and HES—HES was the publically available software that required the fewest data inputs and the least time for data entry [28]. A comparison of the different models is provided in Table 1. This added complexity of the other publically available software, REM/Rate, while potentially useful, can result in larger modeling error if the needed inputs cannot be estimated accurately (such as when a large number of houses are being simulated). Additionally, in a recent evaluation of these three residential energy simulation programs , HES modeled natural gas consumption more accurately than SIMPLE or REM/Rate; the mean difference between observed and modeled natural gas consumption were -9.6% for HES, -21% for SIMPLE, and 36.1% for REM/RATE [20]. Another recent study found that, when building physical characteristics and occupant behavior are accounted for, energy consumption as modeled by HES is accurate to within 1% of actual values when averaged across a group of homes [29]. Finally, we judged HES an appropriate choice for this study given that past McKinsey & Co. analyses [22,30] have used HES to estimate the energy consumption and possible savings from retrofit treatments in the residential sector. Because our study only considers energy consumed for space conditioning, this discussion of HES is limited to those aspects of the model related to space conditioning.

HES calculates and reports end-use energy savings expected for the modeled house with prescribed retrofitting treatments. HES reports these savings both by end-use category (i.e. space heating, space cooling, water heating, appliances, lighting) and by fuel (i.e. gas, fuel oil, or electricity). Space conditioning energy consumption depends on a significant number of factors including, but not limited to: geographic location; house construction and foundation type; appliance use; the quality, quantity, and location of windows; building orientation; HVAC equipment type and efficiency; insulation levels in the floors, walls, and ceilings; air-tightness of the house envelope; and residents' energy-consumption behavior. HES models the major components of space conditioning that Wx programs frequently address: namely, building envelope insulation and air-tightness, HVAC equipment type and efficiency, and residents' energy-consumption behavior [9].

To model these components, HES sends the relevant equipment and house envelope information to DOE-2 software. DOE-2 is a widely used and accepted building simulation program: the U.S. and other countries have developed building standards on the basis of DOE-2, and many design and consulting firms use DOE-2 as the main engine for energy modeling [31]. DOE-2 performs the thermodynamic modeling required to determine hourly space conditioning energy consumption using Typical Meteorological Year, version 2 (TMY2) weather data. Additionally, DOE-2 models HVAC equipment performance using unpublished LBNL performance data to determine equipment capacity curves and efficiency as a function of outdoor temperature [9,32].

2.2 DATA

2.2.1 American Housing Survey Data

This study used the 2007 American Housing Survey (AHS) national microdata to drive the Home Energy Saver software [33]. The U.S. Census Bureau and the Department of Housing and Urban Development have conducted the American Housing Survey every odd-numbered year since 1981. The survey collects data from a fixed sample of roughly 55,000 houses selected in 1985 using cluster sampling. In each iteration of the survey, the Census Bureau adds to the sample some newly constructed houses and removes houses from the sample if they no longer exist. AHS reports not only house characteristics—such as house vintage, conditioned floor area, number of floors, and space conditioning equipment—but also household characteristics—such as household size and income.

While AHS has less energy-related data than the Energy Information Administration's (EIA's) Residential Energy Consumption Survey (RECS), AHS is more useful for this study because it contains more specific location information for each house in the sample. The only information RECS provides that could be used to determine houses' location is census region, census division, and heating and cooling degree-days. At best, this information allows the user to identify a climate contour along which the house exists within a census division. AHS, on the other hand, reports if a house is within a Metropolitan Statistical Area (MSA) area, as defined by the Office of Management and Budget [33]. AHS also reports if the house exists in an urban or rural area within the MSA. This resolution of AHS data makes possible the isolation of low-income urban houses within a specific metropolitan area.

2.2.2 Philadelphia observation data

To evaluate the accuracy of our model, modeled energy savings for low-income urban houses in Philadelphia were compared to energy savings calculated from retrofits of similar housing stock. Observed energy savings data came from an impact evaluation of Philadelphia Gas Works' Conservation Works Program (CWP), a Wx program for low income households in Philadelphia, Pennsylvania [25]. Using the same methods as PRISM, M. Blasnik & Associates (hereafter referred to as "Blasnik") analyzed pre- and post- treatment energy bills to calculate weather-normalized energy consumption in houses that received treatment to determine gross energy savings. To account for any non-program related trends in energy consumption, Blasnik also examined energy bills from a comparison group—a group of houses that did not receive treatment but were physically similar to those treated. The net energy savings were calculated as the gross savings within the treatment group minus the average change in consumption within the non-treatment comparison group.

The Wx program's retrofitting measures included three main retrofitting treatments—programmable thermostat, blower-door guided air sealing, and roof insulation. HES can model these first two measures, but it cannot model the effects of installing roof insulation. However, HES can model the effects of installing attic insulation, which we assumed would have relatively similar effects. With three different treatment elements, there are seven different treatment scenarios of a single treatment or combination of multiple treatments. Table 2 lists each treatment scenario's symbol abbreviation used throughout this report, along with the number of houses that received that treatment according to Blasnik's evaluation.

2.3 ANALYSIS SUMMARY

2.3.1 Analysis methods

For this study, we limited our analysis to occupied, low-income, one-unit buildings within the urban areas of an MSA. One-unit buildings include both attached and detached housing units, but exclude mobile homes and buildings with more than one unit such as apartments or multi-family houses. In this analysis, a household income of 150% of the federal poverty line classified a household as low-income. The federal government uses this income level to determine eligibility for many assistance programs, including LIHEAP, a federal heating and cooling assistance program.

Because our reference data—Blasnik's evaluation—contains combinations of only three treatments, our analysis also only considered these same treatments. Specifically, we modeled that air-sealing would reduce infiltration by 25% and installing attic insulation would increase attic insulation from R-0 to R-38, which is the level of insulation recommended by the International Code Council for houses in moderate climates [34]; these insulation values were selected to model the effectiveness of installing moderately high levels of insulation (R-38) to a previously uninsulated attic (R-0). For treatment scenarios that did not include adding attic insulation, we modeled that the houses had some attic insulation and used the HES default of R-11. In all simulations, the HES default R-value for the walls and roof, respectively R-3 and R-0, was used. The 25% infiltration reduction is consistent with average reductions measured by some specific contractors in Pennsylvania Wx programs, but this reduction estimate is conservative compared to the 40% reductions delivered by contractors in other Wx programs in Pennsylvania [35] or the roughly 27-39% infiltration reduction found in Ohio Wx programs [27].

For this study, we used AHS to provide input to HES for house vintage, conditioned floor area, number of floors, if a house were attached or detached, foundation type, heating equipment fuel and type, air conditioning type, and number of residents in the house. Based on these inputs, HES determined typical values for the remaining input variables based on expected parameters for single family detached houses as described in RECS.

To determine the average energy consumption and energy savings for a city's low-income housing stock, we selected a representative sample of low-income houses from AHS in each city and calculated the energy consumption and energy savings for each treatment in each of the modeled houses. Then, we derived city-averaged energy consumption and savings by weighting the results of each modeled house according to the weights provided in AHS; these weights indicate how many houses in the metropolitan area each house in the sample represents. We used these average energy consumption and average energy savings to describe the results of our analysis. Several of the plots in this paper feature error bars indicating 90% confidence intervals. Blasnik's analysis included calculating these error bars assuming a Student's t-distribution. For the modeled energy savings, we also assumed that the average energy savings followed the Student's t-distribution, and we calculated 90% confidence intervals accordingly, with sample size equal to the number of houses modeled.

2.3.2 City selection

As discussed in Section 1, energy consumption data demonstrate that consumption varies with Census region and climate zone because of housing stock trends and different weather-driven space conditioning demands. In order to investigate how potential energy savings varied among geographical and climatic regions, we selected which cities to analyze based on census region, climate zone, and data availability. With the goal of selecting cities representative of their region's housing stock, we analyzed AHS data to compare the vintage of each MSA's urban housing stock to the vintage of regional urban housing stock. We selected vintage to describe housing stock because factors important to space conditioning (e.g. insulation levels, air-tightness) depend on the building technology available and building practices are largely determined by when the house was built. Additionally, new houses are generally tighter and better insulated than older houses because of improvements in building materials and practices [6].

To measure regional representativeness, we formed a cumulative distribution function (CDF) for house vintage in each city and compared it to the urban housing stock CDF for the Census region. We calculated representativeness as the sum of absolute deviation from the regional CDF, where most representative cities were those with the lowest sum.

To the extent possible, we selected cities to model based on this measure of representativeness. In all cases, however, the most representative cities in each region yielded a very small sample size of low-income household. In the South Census region, for instance, the urban housing stock in the Fort Worth-Arlington, TX metropolitan area was most representative of the region's urban housing stock, but the query for low-income houses in Fort Worth-Arlington identified only one house. Mindful of both sample size and regional representativeness, for modeling purposes we selected the low-income housing stock in Orlando, Florida; Los Angeles-Long Beach, California; Seattle, Washington; Philadelphia, Pennsylvania; Detroit, Michigan; and Milwaukee, Wisconsin.

Building characteristic data for the low-income housing stock in each of these five metropolitan areas in presented in Table 3. This table also includes descriptions of each city's climate. Based on the AHS data, the following significant regional differences in structural design patterns were identified within the various cities' housing stock:

- Basements dominate the foundation type in northern cities (Seattle, Philadelphia, Detroit, Milwaukee), while cities in the south (Orlando and Los Angeles-Long Beach) are predominantly built on concrete slabs.
- Orlando and Los Angeles-Long Beach have newer housing stock than the other cities, suggesting that these houses may initially be better insulated and more tightly sealed.
- The Orlando housing stock relies entirely on electric heating, but most houses in the other cities use natural gas, which may be relevant since electric furnaces generally are more efficient than natural gas furnaces [9].
- Orlando is also the only city where 100% of the houses identified were equipped with air conditioning. Air conditioning ownership in other cities ranged from 24% (Seattle) to 87% (Milwaukee). Although natural gas-powered air conditioning units exist, all air conditioning units in the modeled sample contain either an electric room air conditioner or electric central air conditioning system. The combination of climate and air conditioning ownership will cause the contribution of space cooling to total space energy consumption to vary significantly across the cities we model.
- With the exception of Milwaukee at 1945 ft^2, the average floor area among the different cities is relatively similar, ranging from 1514 ft^2 (Detroit) to 1699 ft^2 (Seattle). However, the number of floors varies substantially among the different cities. The number of floors may be relevant because houses of comparable floor space but with fewer floors will have larger attics than houses with more floors; as such, houses with fewer floors can be expected to lose or gain more heat through uninsulated attics.

3 RESULTS AND DISCUSSION

3.1 MODEL EVALUATION

As discussed in Section 2, our proposed model consisted of driving HES with AHS data in order to predict energy savings for low-income housing stock. To test the accuracy of this approach, we compared modeled energy savings with actual energy savings measured by the CWP. Specifically, we used our model to emulate the measured results in the CWP program. Emulating these results included first analyzing how accurately our model estimated pre-retrofit energy consumption and then analyzing how accurately our model estimated the effectiveness of different retrofitting treatments.

3.1.1 Pre-weatherization energy consumption

We first examined how well our model simulated pre-retrofit energy consumption for space conditioning. The information available in the CWP evaluation limited this analysis in two respects. First, the CWP evaluation includes information on natural gas pre- and post- weatherization consumption, which provides us information about space heating energy, but the evaluation does not include information about the pre- or post- weatherization electricity consumption, so we have no information about the space cooling energy demand or how consumption changes after weatherization. Second, the CWP evaluation uses energy bills to determine natural gas consumption, but because energy bills do not itemize consumption by end-use, we could not precisely isolate natural gas consumption for the space heating end-use. Because we could not compare observed to modeled space heating energy consumption, we instead chose to compare observed and modeled total natural gas consumption. HES provides individual energy end-use estimates for the two major natural gas end-uses—space heating and water heating—and we calculated the total natural gas consumption as the sum of the these two end-uses (thus the model results exclude cooking and other uses, which on average represent a minor amount, approximately 6%, of residential natural gas use [1]). This approach models mean natural gas consumption as 139.9 ± 25.6 GJ (132.6 ± 24.3 MMBTU, 90% confidence interval), which is within 5% of the CWP observed value 126.4 ± 1.6 GJ (126.4 ± 1.5 MMBTU) in 2006. In previous years' CWP evaluations, the average household pre-weatherization natural gas consumption ranged from 128.7 to 197.6 GJ (122.0 to 187.3 MMBTU), so both the modeled and observed energy consumption values appear relatively representative of gas usage within the Philadelphia low-income housing stock [25]. This mean error of 5% compares to a mean error of -9.6% calculated in prior studies using HES [20].

3.1.2 Energy savings

We considered natural gas consumption to be a reasonable proxy for space heating energy consumption, so establishing that the model accurately emulates natural gas consumption suggests HES accurately describes heating loads and thermal exchanges (which are also the main driver of cooling loads) when driven with AHS data. After validating the model's ability to emulate energy consumption, we analyzed how well the model emulated post-retrofit energy savings. Figure 2 shows how average retrofit effectiveness compares between Blasnik's observed results and our study's modeled results for low-income urban housing in Philadelphia. The observed savings, as a percentage of pre-retrofit space conditioning energy consumption, are roughly consistent with those reported in impact evaluations of other Wx programs in colder climate zones [36,37,27,26,35,38].

The CWP evaluation indicates that gas savings range from a minimum of 5% for air sealing and programmable thermostat installation to a maximum of 13% for all of the treatments—programmable thermostats, attic insulation and air sealing. Compared to the observed savings, the model fairly accurately predicted savings when the only treatment was a programmable thermostat (5.4% observed vs. 4.5% modeled) or attic insulation (12.0% observed vs. 10.7% modeled). For each of the other treatment scenarios, the modeled energy savings were more inaccurate (e.g. 6.0% observed vs. 13.3% modeled for the combination of air sealing and thermostat). The error bars plotted in Figure 2 (indicating a 90% confidence interval) suggest the wide variability of energy savings for each treatment scenario. While there is almost certainly some error associated with the energy savings calculation process, these error bars indicate that treatment effectiveness can vary significantly from house-to-house. Such variation occurs because the sample—the Philadelphia housing stock—is heterogeneous. Houses can vary by many physical factors including shape, size, materials, construction type and many other factors. Over their lifetimes, the physical characteristics of houses will change depending on local weather conditions and resident wear-and-tear. And beyond the physical characteristics of the house, there is a behavioral component to energy usage that will cause energy consumption to vary from person-to-person, so take-back may skew the measure of energy savings.

As discussed in Section 1, discrepancies between observed and modeled energy savings fall under the category of shortfall and take-back. Take-back—a behavioral change towards consuming more energy after a weatherization retrofit—is possible in all scenarios, and addressing it is beyond the scope of our research. As for shortfall, inaccurate leakage modeling appeared to be a large source of technical error. As shown in Figure 2, the largest discrepancies between observed and modeled energy savings arose when the treatment scenarios included air sealing. The only combined treatment scenario not to include air sealing—the scenario with attic insulation and thermostats installed—features a discrepancy that is relatively small (11.2% observed vs. 14.8% modeled) compared to the other combination treatment scenarios (e.g. 10.1% observed vs. 18.6% modeled for the combination of attic insulation and air sealing). As discussed in Subsection 2.3.1, the modeled 25% leakage reduction expected from air sealing is comparable to leakage reductions observed from air sealing in other retrofitting programs in Pennsylvania. If we are willing to accept that the 25% leakage reduction represents the actual leakage reductions achieved in the CWP program, then the model overestimated either pre-retrofit leakage, energy losses due to leakage, or both. It is also possible that improper air sealing resulted in lower than expected energy savings. An impact evaluation of the Ohio Weatherization Assistance Program identified improper or inadequate air sealing as the most frequent source of lower-than-expected energy savings [27]. This evaluation found that inadequate air sealing typically around the chimney, plumbing bypasses, wall tops, windows and knee wall bottoms resulted in 30% less reduction in leakage than expected.

3.2 City analysis

3.2.1 Pre-weatherization energy consumption

As discussed in Subsection 2.3.2, for the purposes of analyzing the effectiveness of retrofit treatments varied for different regions, we modeled the energy use and energy savings for low-income urban housing stock in six different metropolitan areas: Milwaukee, Detroit, Philadelphia, Seattle, Los-Angeles-Long Beach, and Orlando. Figure 3 shows the modeled average household space conditioning energy consumption for each of the six cities we examined. Figure 3 also displays the heating and cooling degree days, from which we can see that that energy consumption is primarily driven by heating demand, measured by heating degree day (HDD) [39]. This trend confirms our expectations that the housing stock in colder climates consumes more energy than the housing stock in warmer climates. Among the cities we analyzed, Orlando is the only city that does not follow this trend. Despite the fact that Orlando has the fewest HDDs of any of the cities we modeled, average energy consumption for space conditioning is higher in Orlando than in Los Angeles, a city with more than twice as many HDDs than Orlando. This anomaly may exist because of the difference between the two cities' cooling loads. Orlando has five times as many cooling degree days as Los Angeles, and space cooling constitutes roughly half its space conditioning loads, meaning the difference in cooling demands greatly exceeds the difference in the two cities' heating demand. Additionally, while all of the houses from the Orlando sample were equipped with air conditioning, only 44% of the Los Angeles sample had any air conditioning, making space cooling's contribution to Los Angeles' energy consumption less than it would be if 100% of the housing stock had air conditioning, as is the case for Orlando.

3.2.2 Energy savings

After examining what the model describes as the current status of space conditioning in these cities, we then analyzed the projected effectiveness of the six treatment scenarios. Figure 4 displays the average annual energy savings for an average low-income house in each of the six cities and for each of the seven treatment scenarios considered for this study. These results follow the trend discussed in both Sections 1 and 3.2.1 that houses in cities in colder climates generally consume more energy for space conditioning than cities in warmer climates, so it is expected that there is a greater potential for energy savings in houses that initially consume more energy for space conditioning. Error bars (representing a 90% confidence interval) can be interpreted as describing the variability resulting from physical differences in the housing stock. Although it is difficult to isolate any one factor, variables such as vintage, building materials, space conditioning equipment, and other region-specific building practices all contribute to the energy efficiency of a city's housing stock.

Somewhat unexpectedly, Figure 4 shows that inter-city savings do not remain constant or proportional for different treatment scenarios. For example, attic insulation saves the average Milwaukee house about twice as much energy as it saves the average Orlando house, but air sealing saves the average Milwaukee house about ten times as much energy as the average Orlando house. Such a variation could relate to a modeling error, but our model evaluation discussion in Section 3.1 suggests that the model is capable of estimating energy consumption and savings within reasonable accuracy. A more likely explanation is that the modeled geographic differences in energy consumption and savings are due to differences in climate and housing stock. This is related to the fact that the savings accomplished by the different treatment options do not have the same functional dependence on main driver of energy consumption, the temperature difference ($T_{outdoor} - T_{indoor}$), and on other climatic parameters that vary between the different cities. Attic insulation savings, for example, are proportional to ($T_{outdoor} - T_{indoor}$) in our building model (conductive losses), while reducing air infiltration by 25% results in savings that depend on the initial air infiltration rate and on the changes in indoor humidity resulting from the changes in air infiltration. As such, energy savings related to air sealing are dependent, but not linearly proportional to ($T_{outdoor} - T_{indoor}$) [40]. These results, therefore, are an indication that our thorough modeling approach is well-suited for determining how the effectiveness of treatments vary geographically.

We recall from Section 3.1 that our model exhibited significant error for some treatment scenarios when compared to observed energy savings. As a first-order error correction method, we multiplied the expected energy savings in each treatment scenario by a correction factor derived from our model evaluation in Section 3.1. We developed two different correction factors, a mean error and an extreme error correction factor, for each treatment scenario to account for different levels of model error.

We defined the mean error correction factor as the ratio of the observed mean energy savings from the CWP evaluation to the modeled mean energy savings for Philadelphia's low-income housing stock. For treatment scenarios in which observed savings exceeded modeled savings, we set the correction coefficient to 1, leaving the energy savings estimate unaltered. The mean error correction values range from 0.45 (for the S & T scenario) to 1 (for the A and the T scenarios).

We defined the extreme error correction factor as the ratio of upper 90% confidence limit for the modeled results and the lower 90% confidence limit of the observed results. These values ranged from 0.28 (for the S & T scenario) to 0.66 (for the T scenario). By its definition, the extreme error correction factor is greater than the mean correction factor, and thus the extreme error correction factor provides a more conservative energy savings estimate as it implies greater model overestimation error. This extreme correction factor corresponds to a very low confidence in our results and could be interpreted as a reasonable worst-case scenario factor. Because they are calculated based on Philadelphia data alone, these correction factors may incorporate climatic or region-specific building technology variables that may not apply to other cities. A more robust correction factor could be developed for each city by comparing model results with city-specific energy savings observation, which may become available, for example, when the results of the NWAPE program are released. However, in lieu of this additional city-specific data, applying the Philadelphia-based error correction factor to other cities can still provide insight into the range of possible results. Figure 5 plots the average expected energy savings when applying the mean error correction factor. In this figure, the low end of the error bar represents the savings when applying the extreme error correction factor. The upper end of the error bar represents the modeled mean energy savings with no error correction applied (i.e., the values plotted in Figure 4). Figure 5 demonstrates the greater range of uncertainty for different treatments scenarios (e.g., A &T compared to the individual A and T scenarios) and also emphasizes the model's apparent shortcomings in predicting energy savings for treatment scenarios that consist of multiple treatments, either for technical reasons or for failure to include take-back effects.

4 CONCLUSIONS AND FUTURE WORK

This paper presented a methodology for evaluating the effectiveness of different weatherization treatments when applied to low-income houses in six American cities in different climate and geographic regions. This method, which consisted of driving the HES software using AHS data, modeled average pre-retrofit space conditioning energy consumption with reasonable accuracy when compared to observed energy consumption based on a Philadelphia Wx program evaluation. The results of this study identified that greater energy savings potential generally exists in colder climate; houses in these regions tend not only to consume more energy for space conditioning, but the housing stock is also older than in warmer regions, in which houses are typically more energy-efficient due to advances in building technology over time. However, inter-city energy savings do not remain constant or proportional for different treatment scenarios, indicating that the effectiveness of different weatherization treatments also varies with differences in climate, housing stock, and space conditioning equipment. This is the first comprehensive Wx study spanning several regions with different climates and housing stocks, as well as multiple weatherization options; as such, the findings provide a unique comparative analysis for WAP decision makers. Moreover, the modeling methodology developed here could be used as a reference to be further refined and adapted to develop databases for the effectiveness of a larger array of weatherization options and to cover all US states. There are, however, limitations to the results that are important to recognize.

First, there are occasionally large discrepancies between modeled and observed energy savings for different weatherization treatments. This discrepancy does not necessarily indicate that the HES model is incorrect, only that modeled potential energy results do not reflect observed actual results. Energy modeling literature attributes this to a variety of causes, including insufficient data, modeling errors, improper retrofitting installation errors, or changes in residents' energy consumption behavior; in particular, the accuracy of HES is strongly related to the completeness of the model inputs for house physical characteristics and the residents' behavior [29]. If this is the case and assuming that our data are sufficient, modeled savings should therefore be interpreted as the energy savings possible with proper installation and assuming no changes in residents' energy consumption behavior. Proper installation training and resident education are necessary to ensure that a higher percentage of potential energy savings is realized.

Because the HES model calculates energy savings using DOE-2, a comprehensive physical model—as opposed to statistical models, for instance—it seems unlikely that the model itself is a primary source of error. The model's usefulness, however, hinges upon the data used to drive it, so modeling error can be minimized by driving HES with more detailed information than AHS provides. It is expected that these errors are systematic and affect each modeling run similarly, so it is unlikely that any such error would significantly change the qualitative results of our analysis (e.g. which housing stocks are more effective to weatherize), but it would affect the quantitative details of the results (e.g. the exact quantities of energy saved). Thus, while the simulations in this study may not produce results of sufficient geographic resolution and accuracy to confidently inform specific Wx program decision making; this study's methodology is still useful for comparing weatherization effectiveness among different housing stocks, and it is expected that future research will improve the model's resolution and accuracy. For example, following the completion of the upcoming NWAPE, data and results from that study may be useful to further explore and refine the capabilities of the method presented in this paper.

It is important to emphasize that this approach should only be used to examine the average effectiveness of a specific housing stock. Because the physical characteristics of houses vary depending on design, construction methods, and use, any housing stock will be heterogeneous. Even if weatherization treatments were applied uniformly—which evaluations show they are not—the actual energy savings can vary widely even within a designated housing stock. While our approach may poorly model any one house, we expect that it reasonably models the average energy consumption and savings from many houses.

Errors in this assumption could be decreased by calculating average energy consumption and savings from larger sample sizes. The AHS National microdata we used may not provide the larger sample size required for a more thorough analysis, but data in the AHS Metropolitan supplement might.

This supplement provides the same information as the National microdata we used, except for a larger sample size in 41 metropolitan areas. When this study's simulations were performed, using HES for such modeling was time-consuming, as it required inputting data one value at a time. However, since that time, batch capability has been developed, thereby facilitating the modeling process.

Although our study was designed to model the low-income urban housing stock, our approach is not limited to this application. This approach could be extended to model the cost-effectiveness of weatherization treatments on housing stocks on any level, including upper-income stock and non-urban housing stock, although the usefulness of the average energy consumption and savings determined from such modeling will depend on how the housing stock is defined.

With additional data, this methodology could be used to estimate the carbon savings and cost-effectiveness of different weatherization treatments. Specifically, information regarding energy fuel mix, energy prices, and weatherization treatment costs is required and will vary geographically. For example, housing that rely on coal or fuel oil for primary space conditioning energy will have a higher carbon content per unit energy than houses that use natural gas. The cost of performing weatherization treatments will also vary geographically due to differences in housing stock and labor markets. For instance, houses with smaller roofs will require less attic insulation. Comparing the expected cost-effectiveness of weatherization treatments among different housing stocks could be useful for prioritizing which housing stocks to target for lowering residential energy consumption, household energy bills, and residential carbon emissions. Such analyses of cost-effectiveness and carbon savings based on the energy savings simulated in this paper have been explored [41] and may be further developed in future work.

5 ACKNOWLEDGMENTS

The authors would like to thank Gregory Horman, Principal Research Associate at the Lawrence Berkeley National Laboratory for his help using and understanding the Home Energy Saver software. The authors also thank Michael Blasnik of M. Blasnik and Associates for his insights into the state of residential building science and for providing several weatherization impact assessments. Additionally, the authors thank Michael Donnelly, Alex Acs, and Phillip Connor at Princeton University's Data and Statistical Services lab for their help in understanding how to query American Housing Survey data. Support for this research was provided by the Princeton Environmental Institute's Siebel Energy Grand Challenges program and Isles, Inc. E. Bou-Zeid is supported by the US Department of Energy through Pennsylvania State University's Energy Efficiency Building Hub under grant No. DE-EE0004261.

6 REFERENCES

[1] US DOE/EIA, Annual Energy Outlook 2012 with Projections to 2035, Washington, D.C., 2012.

[2] J.I. Levy, Y. Nishioka, J.D. Spengler, The public health benefits of insulation retrofits in existing housing in the United States, Environ Health. 2 (2003).

[3] M. Schweitzer, B. Tonn, Nonenergy Benefts from the Weatherization Assistance Program--A Summary of Findings from Recent Literature, Oak Ridge National Laboratory, 2002.

[4] US DOE, History of the Weatherization Assistance Program, (2010).

[5] D&R International, Ltd, 2008 Buildings Energy Data book, US DOE, Washington, D.C, 2009.

[6] J. McWilliams, M. Jung, Development of a Mathematical Air-Leakage Model from Measured Data, Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, 2006.

[7] US DHHS, Low-Income Energy Program Funding, 1977 - 2012, (n.d.).

[8] V.P. Shah, D.C. Debella, R.J. Ries, Life cycle assessment of residential heating and cooling systems in four regions in the United States, Energy and Buildings. 40 (2008) 503–513.

[9] E. Mills, The Home Energy Saver: Documentation of Calculation Methodology, Input Data, and Infrastructure, Lawrence Berkeley National Laboratory, Berkeley, 2008.

[10] M.H. Sherman, N.E. Matson, Residential Ventilation and Energy Characteristics, ASHRAE Transactions. 103 (1997) 717–730.

[11] EIA, 2005 Residential Energy Consumer Survey, Energy Information Administration, Washington, D.C, 2009.

[12] M. Fels, PRISM: An Introduction, Energy and Buildings. 9 (1986).

[13] Energy Center of Wisconsin, National Weatherization Assistance Program Evaluation, National Weatherization Assistance Program Evaluation. (2012). <http://www.ecw.org/weatherization/index.php>

[14] E. Mills, Inter-comparison of North American residential energy analysis tools, Energy and Buildings. 36 (2004) 865–880.

[15] D.B. Crawley, J.W. Hand, M. Kummert, B.T. Griffith, Contrasting the capabilities of building energy performance simulation programs, Building and Environment. 43 (2008) 661–673.

[16] P. Raftery, M. Keane, J. O'Donnell, Calibrating whole building energy models: An evidence-based methodology, Energy and Buildings. 43 (2011) 2356–2364.

[17] H. Zhao, F. Magoulès, A review on the prediction of building energy consumption, Renewable and Sustainable Energy Reviews. 16 (2012) 3586–3592.

[18] E.M. Ryan, T.F. Sanquist, Validation of building energy modeling tools under idealized and realistic conditions, Energy and Buildings. 47 (2012) 375–382.

[19] B. Polly, N. Kruis, D. Roberts, Assessing and Improving the Accuracy of Energy Analysis for Residential Buildings, National Renewable Energy Laboratory, 2011.

[20] D. Roberts, N. Merket, B. Polly, M. Heaney, S. Casey, J. Robertson, Assessment of the U.S. Department of Energy's Home Energy Scoring Tool, National Renewable Energy Laboratory, 2012.

[21] S. Sorrell, The Rebound Effect: An Assessment of the Evidence for Economy-wide Energy Savings from Improved Energy Efficiency, UK Energy Research Centre, 2007.

[22] H.C. Granade, J. Creyts, A. Derkach, P. Farese, S. Nyquist, K. Ostrowski, Unlocking Energy Efficiency in the U.S. Economy, McKinsey & Company, 2009.

[23] P. Hoes, J.L.M. Hensen, M.G.L.C. Loomans, B. de Vries, D. Bourgeois, User behavior in whole building simulation, Energy and Buildings. 41 (2009) 295–302.

[24] M. Blasnik & Associates, Colorado Energy $avings Partners Impact Evaluation Report, M. Blasnik & Associates, Boston, 2006.

[25] M. Blasnik & Associates, Impact Evaluation of Philadelphia Gas Works' Conservation Works Program Calendar Year 2006 and Comprehensive Treatment Pilot, Boston, 2008.

[26] M. Blasnik & Associates, NJ Comfort Partners Impact Evaluation Report, M. Blasnik & Associates, Boston, 2004.

[27] M.S. Khawaja, A. Lee, M. Perussi, E. Morris, Ohio Home Weatherization Assistance Program Impact Evaluation, Quantec, LLC, 2006.

[28] Earth Advantage Institute, Conservation Services Group, Energy Performance Score 2008 Pilot: Finding & Recommendations Report, 2009.

[29] D. Parker, E. Mills, L. Rainer, N. Bousarra, G. Homan, Accuracy of the Home Energy Saver Energy Calculation Methodology, in: American Council for an Energy-Efficiency Economy, Washington, D.C., 2012.

[30] J. Creyts, A. Derkach, S. Nyquist, K. Ostrowski, J. Stephenson, Reducing U.S. Greenhouse Gas Emissions: How Much at What Cost? U.S. Greenhouse Gas Abatement Mapping Initiative, McKinsey & Company, 2007.

[31] K. Ellington, DOE-2 : Software : Resources : Environmental Energy Technologies Division, Lawrence Berekely National Laboratory. (2010).

[32] LBNL, Calculation Methodology, Home Energy Saver: Engineering Documentation. (2012). <https://sites.google.com/a/lbl.gov/hes-public/Home-Energy-Saver/calculation-methodology>

[33] US Census Bureau, HUD/U.S., American Housing Survey for the United States: 2007, US Census Bureau & Department of Housing and Urban Development, Washington, D.C., 2008.

[34] International Code Council, International Energy Conservation Code 2009, International Code Council, Inc, Washington, D.C., 2009.

[35] M. Blasnik & Associates, Impact Evaluation of Columbia Gas of Pennsylvania's Warm Choice Program Calendar Year 2005, M. Blasnik & Associates, Boston, 2007.

[36] M. Blasnik, What Saves Energy & Why: US Program Measured Results, (2009).

[37] APPRISE, PPL Electric Utilities Winter Relief Assistance Program, Princeton, 2006.

[38] M. Blasnik & Associates, Colorado Low Income Energy Efficiency Retrofit Program Energy Impacts: First Response and Energy Savings Partners, M. Blasnik & Associates, Boston, 2009.

[39] National Oceanic and Atmospheric Administration/ National Climatic Data Center, Climatography of the U.S. No. 81, Monthly Station Normals, 1971-2000, (2002).

[40] American Society of Heating, Refrigeration and Air-Conditioning Engineers, Inc., 2009 ASHRAE Handbook - Fundamentals (SI Edition), American Society of Heating, Refrigeration and Air-Conditioning Engineers, Inc., Atlanta, 2009.

[41] J. Bradshaw, Cost-Effectiveness of Weatherization in Low-Income Urban Housing Stock, Princeton University, 2010. <http://efm.princeton.edu/pubs/Bradshaw_Thesis%20FINAL.pdf>