summit

complexity simplified.

U.S. Department of Housing and Urban Development

HUD Utility Schedule Model (HUSM)

Task 10 – Final Model Development Report

Contract Number: DU206SF-13-T-00003

July 28, 2015

Table of Contents

Executive Summary

The purpose of this memo is to report the summary and results of Summit's validation and update of the U.S. Department of Housing and Urban Development (HUD) Utility Schedule Model (HUSM), as commissioned by the Office of Policy Development and Research (PD&R).

HUD provides rental housing assistance to low-income households in the United States through Public Housing, the Housing Choice Voucher Program (HCVP), and various project-based rental assistance programs administered by the HUD Office of Housing. In these programs, eligible tenants are generally required to pay a percentage of their income (usually 30 percent) toward "shelter costs," which include rent and select utility costs. HUD pays the balance of the shelter costs for a resident through a subsidy. Covered utility costs include reasonable costs for space heating, water heating, cooling, refrigeration, lighting, appliances, water, and sewer fees. However, unlike rent, which represents a predetermined contractual amount between a landlord and tenant, utility costs vary with consumption. Thus, only the extent of consumption that a Housing Authority (HA) determines as reasonable utility costs are covered under shelter costs. Administrators of HUD rental housing assistance programs require a method of estimating a reasonable allowance for covered utility costs for their eligible tenants. These estimated costs are referred to as "utility allowances."

Program administrators currently have discretion in terms of how to estimate utility allowances. Many of the acceptable methods require consumption data collection or engineering calculations that may be difficult for administrators to obtain. In order to assist program administrators, PD&R developed an Excel-based HUSM tool that program administrators can use, if they choose, to estimate allowances for their tenants. PD&R's current model uses algorithms, based upon regression analyses and correlations, to estimate the expected utility costs for a tenant occupying a housing unit with specified location characteristics, utility services, and utility rates.

For this project, PD&R engaged Summit to validate and improve the current HUSM using additional data sources identified by Summit and HUD, update the existing Excel-based model with enhancement features desired by users, convert the Excel-based model into a user-friendly web-based tool, and update HUD's Public Housing Utility Allowance Guidebook and HUD's Housing Choice Voucher Program Guidebook. These tasks are designed to ensure that the HUSM provides reasonable estimates of utility costs, and that the model is comprehensible and easy to use for program administrators. The development of more accurate and consistent utility schedules will assist HUD in its efforts to more accurately and fairly project utility consumption for reimbursement of tenants, improve energy efficiency and sustainability, and assist in Housing Authority (HA) benchmarking efforts.

This memo, and final deliverable under this project, details the ten delivered tasks and their specific contributions, which include the following:

- exploration of alternative data sources for HUSM, which led to Summit's conclusion that the U.S. Energy Information Administration's (EIA) Residential Energy Consumption Survey (RECS) is the most appropriate source of data for the HUSM;
- validation that the original HUSM used best practice techniques for the set of regressions models that underlie HUSM;
- identification and correction of some errors in the original Excel-based HUSM tool;

- revised HCVP and Public Housing guidebooks, with detailed and clear HUSM instructions for using HUSM;
- development of an online version of HUSM, with an enhanced interface and online-specific features;
- updates and refinements to select algorithms and regression models underlying HUSM, including updates to HUSM's underlying weather data, HA list, methodological changes to calculation of heating and water/sewer usage; and
- additional HUSM functionality, such as zip-code level weather data, two new "Green" discounts, factor adjustment capability for users to increase or decrease the model projected allowances, and a mechanism for comparing local allowances with HUSM-generated state average allowances.

Background

Summit's engagement with the U.S. Department of Housing and Urban Development (HUD) began with the review of the existing HUD Utility Schedule Model (HUSM). HUD had previously involved several contractors in various aspects of development and validation for the HUSM.

HUSM was originally developed in 2002 in a contract awarded to GARD Analytics. Their mandate was to generate a nationally applicable utility schedule model that would provide HAs the ability to estimate the utility costs for their federal-assisted tenants. GARD's Excel-based model was based upon 1997 Residential Energy Consumption Survey (RECS)[1] data.

The RECS dataset is a nationally representative sample of housing units. The survey collects housing unit energy characteristics, usage patterns, and demographic information. RECS is administered every four years[2]. To date, 13 surveys have been conducted; initially in 1978 and most recently in 2009. GARD provided an extensive report[3] that described the model.

In 2004, HUD engaged 2rw+di to validate the allowance model against actual utility bills [4] and revise the work of GARD. They found GARD's model to be reasonably accurate, and updated the results using the 2001 RECS. In addition, 2rw+di added a method of adjusting for building age and for estimating heat pump usage. In 2007, HUD engaged GARD to revise the heat pump efficiency calculation, updating the heating energy use calculation and comparing the model's results with actual PHA allowance data[5].

From 2012 to 2014[6], Riley & Associates (Riley) were engaged to investigate the effects of using an aggregated dataset of RECSs results from 1997 through 2009, rather than using only the most recent

[1] U.S. Energy Information Administration (EIA), "Residential Energy Consumption Survey (RECS)," http://www.eia.gov/consumption/residential/.

[2] Four-year cycles began in 1993. Prior to the four-year cycle, RECS was published every three years beginning in 1984 and published yearly between 1979 and 1982.

[3] GARD Analytics. 2003. Utility Allowance Model Final Report.

[4] 2rw+di. 2005. Utility Model Evaluation.

[5] GARD Analytics. 2007. Final Report on HUD52667 Spreadsheet Update.

[6] Riley & Associates. 2014. HUD Utility Model (HUSM) Re-benchmarking.

survey data (2009). When Riley recalculated the estimation algorithms using the 2009 RECS alone, their results were statistically consistent with the results using combined data from the 1997, 2001, and 2005 surveys. Riley recommended and implemented the model, using the most recent data alone. They also simplified the model by using fewer structure types and eliminating structure age adjustments, and including a simple 18% discount rate for Energy Star homes.

In 2014, HUD charged Summit with reviewing the calculations used by Riley, providing model updates, and validating the model. In addition, HUD tasked Summit with revising the Utility Allowance Guidebook, which describes the methodologies and decision parameters in estimating utility allowances. Lastly, HUD tasked Summit to deploy the model (using the Excel tool) in a web-based interface.

In the following sections, we provide a brief synopsis of all ten tasks that comprise Summit's work under this engagement.

Task 1 – Conduct Review of Entities Collecting Residential Utility Data

In Task 1, delivered to HUD on January 2, 2014, Summit examined potential data sources that could be used to update the HUSM. Based on our analysis, Summit recommended continuing to use the RECS data for the HUSM, as it is the most comprehensive source of the required data elements with national coverage of all required building types and utility usage types. In addition, Summit found that using structure and bedroom relationships derived from the more common utility types, past RECS surveys, and other sources such as the AHS, can be used to indirectly estimate usage for residences that are less commonly represented in the 2009 RECS.

The HUSM requires detailed information on housing structure (such as building type and number of bedrooms per unit), energy consumption by fuel type (natural gas, fuel oil, electricity, etc.), and geographic location. Summit investigated more than 20 separate data sources from Government agencies, public utilities, non-profit organizations, and private entities. Accessibility, geographic coverage, cost, and property detail were the criteria used to evaluate potential data sets.

The current version of the HUSM is based on 2009 RECS data. The EIA sponsors the RECS survey, collecting data on energy-related characteristics and usage patterns of a sample of housing units. The dataset includes type and number of energy consuming devices, usage patterns, structural characteristics of the home, and household demographics, as well as weather data. Additionally, RECS data includes information on the number of floors and bedrooms in units, and identifies households that receive public housing assistance.

In circumstances where the 2009 RECS data cannot be used to produce reliable estimates (such as for rare building types, units with five or more bedrooms, fuel mixes, or for non-energy utilities), RECS may be augmented by additional data sources. Examples of additional data sources include previous versions of RECS and the American Housing Survey (AHS).

Other datasets did not include sufficient information to be included in the final dataset used for developing the model algorithms. In general, the sample sizes of these datasets were insufficient and some did not contain the necessary information to build the model. Some of the other datasets that were studied include the U.S. Department of Energy (DOE) Low Income Weatherization Program, the 2011 U.S. Census, and the HUD Office of Affordable Housing Preservation (OAHP) Mark to Market

Program/Stewards for Affordable Housing (SAHF), which can be used to help refine model estimates. In addition, Summit considered the HUD Office of Public and Indian Housing (PIH) Benchmarking Utility Consumption & Cost Systems (BUCCSs) and PIH datasets as tools to validate the model.

Task 2 – Develop and Describe Strategies to Estimate Utility Allowances

In Task 2, delivered to HUD on February 28, 2014, Summit developed strategies to update and validate the HUSM, based on a review of the regression modeling methodology described in the Riley HUSM Re-benchmarking reports from 2013. Summit used the HUSM Re-benchmarking study as a basis for validating the model coefficients, specifically those that relate climate- and building-related factors to energy usage.

After a review of potential validation data sources, Summit recommended using the PIH Benchmarking Utility Consumption & Cost System (BUCCS) data to validate that the HUSM model generates accurate estimates of usage in real-world situations.

In addition, Summit recommended the following HUSM enhancements:

- inclusion of regional effects in the regression models used to estimate energy usage;
- a more granular definition of location, which would allow zip code-level estimates of heating degree and cooling degree days;
- “Floor” and “Ceiling” values for all types of utilities;
- an updated definition of and the value of the “Green Discount” currently applied to Energy Star properties; and
- an update of the water usage estimates based on U.S. Geological Survey data.

Task 3 – Develop and Validate the Model’s Estimation Algorithms

In Task 3, delivered to HUD on May 30, 2014, Summit validated the HUSM and implemented an initial round of model enhancements.

HUSM Validation

Summit validated the model-predicted outputs of the regression equations using alternative sources, including BUCCS survey data, HUD Office of Affordable Housing Preservation (OAHP) Utility Consumption Baseline Analysis (UCBA) data, and 2011 American Housing Survey (AHS) data.

- **BUCCS**: BUCCS contains utility usage survey data from different buildings, including both single family and multifamily, over 350 PHAs nationwide from 2004-2007. The data includes 2,680 properties and contains the following variables which are relevant to our validation exercise: property location, construction date, property type, number of bedrooms, and annual electricity, gas, water, and sewer consumption and expenses.
- **UCBA**: The UCBA data contain over 200 HUD-assisted properties with baseline data which includes key information such as property location, number of bedrooms, the annual utility usage for each unit, and the end use for each utility type, when separately metered. The data consists of both pre- and post-retrofit utility consumption baseline analysis. 61 properties had both baselines for our analysis.

- **AHS**: AHS does not have utility consumption information. Instead, it contains unit-level utility expenses. Because the AHS data we have is from the 2011 survey, Summit used the 2011 state-level average utility unit prices from EIA survey data as the utility tariff unit price in the HUSM model to estimate the utility consumption. Summit enhanced AHS by merging the PHA code using county and state, then kept only single family detached, single family attached, lowrise apartment, and larger apartment building unit types with housing assistance to reach a total sample population of 1,795 valid observations covering 25 PHAs in 8 different states.

Summit found that the HDD/CDD implementation and model prediction are in a reasonable range of values based on outside estimates from these datasets.

In addition, Summit also validated the HDD/CDD values for each HA using 1981-2010 National Oceanic and Atmospheric Administration (NOAA) Weather Normals.

HUSM Errors and updates

Summit made corrections to the following errors in Excel-based implementation of the HUSM:

1. Summit corrected an error in how fixed monthly charges were applied to different utility types; that error previous led to underestimation of utility costs in certain cases. More specifically, the monthly charge was allocated exclusively to one end-use. For example, an electric monthly charge was allocated to "Other Electric" and a natural gas monthly charge was entirely found in "Cooking". Summit recognized that not all utility profiles are the same. In the example just described, were a tenant to have a natural gas fixed fee and use an electric stove, their allowance would not account for the monthly charge. Summit implemented a logic into HUSM which adds the monthly charge to the first instance of the utility wherein the tenant is responsible for remittance.

2. Summit fixed the electric water heating expense functionality, which was implemented incorrectly on the final HUD 52667 form. Formerly, the model included both electric resistance and electric heat pump as possible tenant remittances. However, the schedule only provides for "electric" water heat. The incongruence caused the monthly allowance for water heating to equal $0 when the targeted unit uses electric as water heating fuel source, regardless of the actual monthly allowance for electric water heating.

3. In task 3, Summit corrected the engineering-based model coefficients for electric heating in apartments of two to four units and for oil heating across building types. Ultimately, Summit would update the entire coefficient table, as per an updated from the Riley Re-benchmarking 2014 Report[7].

4. The most recent model documentation reviewed by Summit, after the delivery of Task 3 report, indicated that the current HUSM uses a Section 8 tenants' distribution-based weighting logic to calculate average HDD/CDD for each HA. However, Summit found that in the HUSM, each HA was assigned a HDD/CDD for the nearest weather station and that population-based weights were not used in the actual HUSM.

[7] Riley & Associates. 2014. HUD Utility Model (HUSM) Re-benchmarking.

Model Enhancements

Summit conducted the following updates/analyses of the model's algorithms:

1. Summit explored including regional effects in the energy usage models. Our model estimation showed that regional-level geographic information (i.e., region, division, and reportable domain) from RECS 2009 data does not provide statistically significant marginal benefit to the model fit, due to the high correlation between the existing model explanatory variables (heating degree days [HDD]/cooling degree days [CDD]) and geographic variables. Therefore, we recommended that HUD not include any of the three regional variables from RECS in the current model.

2. Summit implemented "Floor" and "Ceiling" functionality in the Excel-based HUSM to take into account real-world situations when any minimum/maximum billing amount for a utility type is required.

3. Summit also built in a "factor adjustment feature" that allows PHAs some flexibility in using the HUSM. A factor adjustment provides for the possibility that HUSM does not particularly represent a particular housing stock or cultural behavior. Such an alteration makes sense when the HUSM estimate is uniformly insufficient across an allowance type (e.g., if a stock of 2-bedroom units using space heating is uniformly over-estimated). The HUSM user could reduce the allowance estimate by an appropriate percentage. When a factor adjustment can be made that provides for reasonable allowance according to the housing regulations, the HUSM remains the preferred form of estimation over the engineering-based methodology.

4. Summit proposed two possible solutions to update the water usage estimate in the HUSM, which were considered by HUD:

    A. Retain the current water usage relationship in the equation, but update indoor water usage coefficients using the updated number from 2005 U.S. Geological Survey (98 gal) for the cold water consumption per capita per day assumption; or,

    B. Replace the equation by implementing an updated regression relationship between the monthly water consumption and various factors (e.g. household size, monthly local minimum temperature, and state), and also update the water consumption assumption.

Task 4 – Update the HUD Utility Allowance Guidebook

In Task 4, delivered to HUD on September 7, 2014, Summit edited, updated, and reorganized two of HUD's utility allowance publications: the 2001 HCVP Guidebook and the 1998 Public Housing (PH) Utility Allowance Guidebook.

Previous guidebook editions had directed HAs to choose between engineering-based data and utility consumption data, as a methodology for estimating utility allowances. The updated guidebooks still suggest those methodologies. However, Summit added instructions on how and when to use the HUSM model.

For the PH Guidebook, Summit updated all of the facts, figures, tables, appendices, resources, citations, and examples. Most importantly, Summit reformatted the decision-tree section of the text, which was designed to help users choose the appropriate methodology. In addition, Summit added a chapter and

appendix that describe and illustrate use of the HUSM, referencing the web-based HUSM instead of the spreadsheet version.

For the HCVP Guidebook, Summit updated chapter 18, which covers utility allowances, to include clear online HUSM instruction.

Task 5 – Develop and Provide a Web-Based HUSM

In Task 5, delivered to HUD on September 7, 2014, Summit developed a web-based beta version of the HUSM. The web-based model is Section 508-compliant, meets application guidelines, is secure, and is capable of residing on the HUD User website. In keeping with these guidelines, the model is built using a MySQL database. The web page is designed with both HTML and JavaScript. The JavaScript was coded to interact with PHP script on the server.

HUSM users can clearly identify where an input is required and many entries use drop-down menus. Furthermore, since the Excel-based HUSM was divided across multiple sheets within a workbook, Summit reorganized the model into a single form, simplifying the user input process.

Though the web-model was contractually delivered before subsequent updates due in Tasks 7-10, Summit updated the web-model concurrently. Thus, all of the updates described in the following sections are also included in the web version. Currently, the model has been vetted by HUD, using a staging server, and is in the possession of HUDUSER[8].In addition, Summit updated the Excel version to reflect all the HUSM model enhancements made over the course of this project.

Task 6 – Deliver Model Interim Report

On October 27th 2014, Summit delivered its Model Interim Report to HUD. That report detailed Summit's previous deliverables, Tasks 1-5. The remainder of this report similarly describes subsequent deliverables, Tasks 7-10.

Task 7 – Develop sample of PHA generated utility allowance schedules

On December 23rd 2014, Summit delivered to HUD a memo describing the statistical sample of utility allowance data Summit developed and summary statistics of the allowance comparison between HUSM and the statistical sample. Summit gathered a convenience sample of three HAs that publish their allowance online, in each of the nine climatic zones of the U.S., as defined by National Oceanic and Atmospheric Administration (NOAA)[9]. Summit compiled a dataset of the 27 HA allowance schedules, with allowances for each usage, unit type, utility[10], and number of bedrooms, as well as HUSM predicted schedules for the identical 27 HAs[11].

[8] Upon delivery of this report, no public URL is available, as HUDUSER has not published the site.
[9] http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php
[10] The convenience sample only included allowances with electric, natural gas and oil as utility types.
[11] Overall, there were 8,610 utility allowances in the data; 4,305 were from online HA schedules and the other half were produced from HUSM. HUSM schedules were produced using state average utility rates from EIA survey data.

Summit compared HUSM allowances with those of the convenience sample of HAs. While the comparison was intended to be descriptive and was not able to support statistical inference, we found that different allowances exhibit different variation patterns. The variance between HUSM and published allowances for all utility types compared in the analysis remain relatively constant across unit types and the number of bedrooms. Among all compared utility types, electric space heating, regardless of region, has the greatest variance between published allowances and model predictions. In addition, HUSM exhibits the closest prediction for water heating and cooking compared to HA allowances.

Task 8 – Comparison of PHA allowances with those of HUSM and analysis of LIHTC allowances

On June 11th, Summit delivered Task 8[12] to HUD. For Task 8, Summit enhanced its data collection methodology on the published HA allowance data and conducted detailed analysis on the difference between HUSM and HA allowances. Summit also conducted some basic exploratory analysis on the difference between HUSM-generated utility allowances and a set of Low-Income Housing Tax Credit (LIHTC) allowances provided by HUD.

In this task, Summit expanded the statistical sample of published HA allowances from Task 7 and constructed a dataset of 224 HA utility schedules published online, representing 48 states, 9% of all HAs and 32% of all units of public housing.

By comparing the HUSM-generated energy-related utility allowances with actual utility allowances by property type, utility type, number of bedrooms, and climate zone, Summit examined the potential cost impact to HUD of using HUSM-generated utility allowances. Summit also identified potential areas of improvement for the energy consumption estimation function within HUSM.

Major findings of this analysis:

- The median difference between HUSM-generated energy allowances and HA-published energy utility allowances, broken down by end-use, is close to 0%. Although not from a statistically representative sample, this indicates that using HUSM will not lead to a significant difference in the magnitude of overall energy-related utility allowance expenditures paid by HUD.
- HUSM-generated energy allowances are generally slightly higher than published allowances for studio apartments and for units using only natural gas utilities (+$4 to +$6) and generally lower than published allowances for other units with 2 or more bedrooms (-$3 to -$9).
- HUSM generates higher heating allowances than published schedules in warmer climates. This is because, in the current iteration of the HUSM, heating allowances are generated for all properties, even in the warmest climates. Task 9 introduces a HUSM update to address this issue.
- HUSM-generated estimates of “Other” electric use tend to be higher than published schedules.

Summit also received from HUD a dataset of the LIHTC program utility allowances. However, as opposed to the published HA allowances that provide detail regarding the individual components of each

[12] Task 8, a comparison report of the HA published schedules and HUSM, was originally delivered, as per contract, on February 27th. HUD provided Summit the LIHTC data the day before, and extended the deliverable schedule to include a LIHTC analysis.

allowance, LIHTC allowance data includes only the total utility allowances without any information about the composition of those allowances. Therefore, it is impossible to determine with certainty what the energy-related, water-related, and other portions of the allowance are. Instead, Summit compared LIHTC total utility allowances with HUSM predicted total by assuming that all LIHTC allowances include natural gas space heating, water heating, cooking, and other electric consumption. Summit found that HUSM-generated allowances on average are about 80% of the total allowances granted within the LIHTC program.

Task 9 – HUSM Refinement Report

Task 9, delivered to HUD on May 29th 2015, reported on Summit's implementation of updates and model refinements as a result of model refinement and validation studies conducted in Tasks 3 and 8. These model updates and refinements are described below.

Summit implemented the HUSM changes into both online and Excel-based versions.

"Green Discount"

With increased public housing investment in energy savings, Summit included two new parameters for "Green discounts" in the model: a discount of 25% for buildings which meet the Leadership in Energy and Environmental Design (LEED) standards and a discount of 18% for those that have undergone significant retrofit. The additional 25% and 18% savings were determined based on the following outside studies:

- **LEED**: The LEED rating system, developed by the U.S. Green Buildings Council, awards performance in five key areas of development: sustainable site development, water savings, energy efficiency, materials selection, and indoor environmental quality. In 2008, the New Building Institute (NBI) released a report[13] that measured energy performance for 121 LEED-certified buildings. The report estimated LEED whole building performance, on average, including energy and water usage, to be 25-30% better than the baseline. Based on these findings, Summit implemented a conservative discount of 25% on the total allowance prediction for water and energy usage in HUSM for LEED-certified properties.
- **Significant Green Retrofit**: Many properties have undergone energy- and water-efficiency retrofits that can be defined not by a certification but in terms of their rehabilitation. These include improvements to heating, cooling, lighting, domestic hot water (DHW) systems, appliances, building envelope, conservation water measures, and on-site generation encapsulate retrofit projects that increase energy- and water-efficiency. To determine a reasonable adjustment, Summit researched average savings ranges seen in affordable housing retrofit. One of the most recent and relevant reports was a 2014 study on the HUD Multifamily Green Retrofit Program (GRP), undertaken by Stewards of Affordable Housing for the Future (SAHF).[14] GRP began in 2009 and involved the retrofit of 227 affordable housing properties. The study reported that whole building energy consumption and water consumption were each reduced by 18%, which was adopted by Summit, with little variation as to the required GRP investment.

[13] http://www.usgbc.org/Docs/Archive/General/Docs3930.pdf
[14] http://www.sahfnet.org/multifamilyretrofitreport_2_1287596736.pdf

Regression-based Water Consumption Estimates

The water usage and sewer calculation used in the previous version of HUSM had not been updated since 2003 and provided a fixed consumption per number of bedrooms. In Task 3, Summit determined that this method underestimated actual water usage for most of the states. In this task, Summit implemented water and sewer consumption calculations based on a statistical relationship with property-related factors, including number of bedrooms and location. Summit's regressed the average monthly water consumption per person on each state, using the 2005 U.S. Geological Survey[15], and then adjusted consumption based on an estimated number of people per bedroom.

Adjustment of Heating Consumption Estimates

Summit's validation work illustrated that the HUSM model over-predicted heating consumption in very warm climates, and identified that this was a result of the particular form of the heating consumption regression equation. Summit maintained the underlying heating regression equation for most climates and applied an adjustment in the HUSM to provide zero heating allowance for areas with fewer than 150 heating degree days (HDD) per year. Summit targeted this threshold by considering the overall share of HAs affected and areas that traditionally do not need heating, including review of published allowances from warm climates. Less than 3% of HAs are below 150 HDD when taking their 30-year averages. These include some but not necessarily all HAs in Puerto Rico, Hawaii, and Florida.

Data Updates and Maintenance

Other data updates and maintenance implemented in this final version of the HUSM include the following:

- HUD requested that Summit provide non-HA users the ability to generate utility estimates based on zip code, not HA. Summit used NOAA's climate database and census zip code data to assign the closest weather station to each census zip code according to the distance calculated based on their longitude and latitude. Summit added approximately 30,000 zip codes with their corresponding HDD/CDD to HUSM. Thus, HUSM users can compare/generate HUSM estimates according to their HA code or their zip code.
- Summit also updated the HUSM with the latest NOAA 30-year weather data. HUSM had formerly included 30-year HDD and CDD data from 1971-2000. In the updated HUSM, HDD and CDD data are from 1981-2010.
- Finally, Summit made some routine maintenance updates. Most notably, Summit updated the list of HAs in the model using a list from HUD's website.[16] The update includes 2,105 HAs whereas the previous version of HUSM includes 2,362 HAs. The difference can be ascribed to a change in HA status; either the HA is no longer active on HUD's list or the HA no longer administers utility allowances.

Task 10 – Final Report and State Average Update

In Task 10, besides providing HUD a abovementioned summary of all the work have been done in previous tasks and recommendations for future improvements, Summit also added another update to

[15] http://water.usgs.gov/watuse/data/2005/index.html
[16] http://portal.hud.gov/hudportal/HUD?src=/program_offices/public_indian_housing/pha/contacts/

HUSM that is described in this section. Summit added a HUSM-generated average state utility schedule for HUSM users to use as a basis for comparing their results.

State Average Utility Schedule

Utility allowance administrators are tasked, according to HUD guidelines, with estimating a "reasonable" utility allowance. These guidelines are relatively straightforward when the administrator has access to engineering or consumption data, as this information forms the basis of other utility allowance estimation methodologies. However, many HAs, for whom HUSM is most accessible, do not have other forms of data. Summit developed a state average HUSM-generated allowance table that allows a user to compare their specific HUSM schedule against a HUSM schedule that includes state average HDD and utility rate information. This can provide users with some context for their utility allowance calculation, and also flag user potential unreasonable allowances.

To generate these state averages, Summit first calculated average HDD/CDD, (Cooling Degree Days) by state and month. Summit also collected data from two sources to establish state-wide utility rates: The Energy Information Administration (EIA) publishes electric[17] and natural gas[18] utility rates by state annually, and the American Water Works Association (AWWA) surveys both water and waste water utility rates by state annually[19]. EIA does not publish annual rates for oil or bottled gas. AWWA data required a calculation to extract a utility rate; Summit took the AWWA's median expected variable portion[20] of a water/sewer bill and divided that by the expected consumption for a typical bill. Appendix 1 includes a table of those 2014 annual rates. Summit developed a functionality in HUSM to automatically calculate the state average utility schedule based on the state average HDD data and the state average tariff for each state, utility type, building type, and number of bedrooms. Summit also added a worksheet in Excel and a button online for users to quickly compare their HUSM generated schedule with their state's average HUSM-generated schedule.

Recommendations for Future HUSM Enhancements

In the following sections we outline our recommendations for the future of the HUSM. Our recommendations include two principle types of improvements: modeling recommendations that address HUSM's ability to predict utility usage with precision and user optimization recommendations that help the HUSM administrator use HUSM. This section is divided accordingly.

Modelling Recommendations

1. **Summit recommends a nationally randomized sample of public housing utility consumption, across building types, utility types, apartment size, and end-use.**

   Even though Summit has concluded in the project that RECS and USGS data was the best available source for the HUSM model estimation thus far, Summit identified that the current

[17] http://www.eia.gov/electricity/monthly/current_year/february2015.pdf (Page 124)
[18] http://www.eia.gov/dnav/ng/ng_pri_sum_a_EPG0_PRS_DMcf_a.htm
[19] http://www.awwa.org/resources-tools/water-and-wastewater-utility-management/water-wastewater-rates.aspx (2014 Interactive database)
[20] The variable portion of a bill, defined in the 2014 AWWA Water and Wastewater Rate Survey, includes the expected portion of a bill that is not fixed. Summit did include fixed fees in estimating average allowances, but they are not appropriate in determining average rates.

HUSM has suffered from issues of thin data in uncommon unit and utility types and information imprecision due to the aggregated level of data. Over the next five years, HUD is planning on including as part of data collection and analysis effort a Utility Allowance Comparison Study. This large-scale data collection effort is a unique opportunity for HUD to address particular issues in utility allowance estimation, and to improve the HUSM through targeted data collection. In order to leverage this opportunity for the most benefit of HUD, we recommend that HUD consider the following issues when implementing the study:

- *Stratify the sample collection based on variables known to impact utility usage estimation.* Rental subsidy calculation errors in HUD QC study are generally the results of poor oversight or management, or random calculation errors. These calculation errors are generally not related to aspects of physical plant or geography. However, we have seen in our HUSM validation work that variance between utility allowances and actual energy consumption is often directly related to regional or property-related factors. Summit recommends stratifying projects by building type, heating source, and climate zone when selecting units for the Utility Allowance Comparison Study. In addition to explicit stratification on these variables, it may also be useful to implicitly stratify the sample (i.e., sort the sample prior to selection to promote maximum variation) by additional variables of interest. These may include method of utility allowance calculation, and whether utilities are tenant- or owner-paid.
- *Oversample specific types of units that are under-represented in the RECS.* As part of the model validation process, Summit concluded that the 2009 RECS was the best available source for the HUSM model estimation thus far. While RECS is a high-quality source of data that works well for estimating most types of utility consumption, the sample size in RECS is limited for particular types of housing stock that are relatively uncommon. Specifically, building types with small sample sizes in RECS include each of the following:
    a) Manufactured homes (not found in the public housing stock, but are found in the HCV program)
    b) Properties using bottled gas
    c) Multifamily properties using fuel oil
    d) Units in warm climates that use natural gas heat
    e) Studio bedroom apartments

  For these types of units, the HUSM consumption estimation algorithms were based on less precise, less recent data sources, including the AHS and previous years of RECS data. The HUSM energy estimation algorithms would benefit greatly from additional data for these particular types of units. Therefore, we recommend that HUD attempt to gather at least 10 samples from each of the above-listed types of properties. These data can be used to validate and potentially increase the precision of the utility consumption estimates for these types of properties.

- *Collect additional data for use in estimating water consumption in the utility allowance context.* Prior to the current HUSM update, the methodology for estimating water consumption had not been updated since 2003. In the current implementation of HUSM, Summit used the best available resource for estimating water consumption, data on the U.S. Geological Survey (USGS). However, the USGS only allows provides average household consumption at the state level, and does not include variables that could aid in estimating consumption more precisely. In particular, we recommend that HUD gather each of the

following variables on water consumption for all units participating in the Utility Allowance comparison Study:

a) End uses covered
b) Household size
c) Tariff
d) Building type

Although end-uses for water are not specifically included in the HUD-52667, information on end uses is particularly useful, as household water consumption depends heavily on types of consumption that are often not included in utility allowances, including lawn maintenance and clothes washing. Utilizing detailed unit-level data from HUD's actual portfolio, as opposed to national average water consumption data, will ensure that HUD is not overestimating water consumption.

2. **Summit recommends a multilevel model, as the basis for utility consumption estimation.** HUSM currently uses Ordinary Least Squares (OLS) regressions to estimate consumption for each iteration of building type and utility type. Where data are thin, HUSM uses other sources of data or on-top engineering adjustments to fill in the data gaps[21]. While the Summit validation process confirmed that this regression methodology is a valid practice approach, Summit believes another modelling approach would produce greater precision in utility consumption estimates.

   For example, multilevel models is a regression technique that can potentially mitigate the thin data issue. This regression technique allows construction of models which take account of the structure of the data. Specifically, this model accounts for data contained in groups, such as regions and building types. The groupings in multilevel models can be thought of as separate models per region or per building type. However, instead of entirely separate regression models, a multilevel model will also take the data from other regions into account simultaneously. This helps improve estimates overall, because the data is not partitioned into separate datasets to be analyzed entirely separately. Multilevel models also help where there is little data; in these cases, the model relies more heavily on data from related groupings. Thus, the technique safeguards extreme and most likely inaccurate estimates, simply due to lack of data.

User Optimization Recommendations

1. **Summit recommends including functionality that produces more salient utility schedules.** HUSM is intended to generate example HUD-52667 forms, which require an HA to produce a utility allowance estimate for a particular unit, expressed as the sum as a series of predefined end uses. However, many administrators are interested in producing comprehensive utility allowance schedules, as opposed to individual HUD-52667 forms. In the current model configuration, Summit implemented the functionality in HUSM to produce a utility allowance schedule based on the particular unit a HA uses to generate the example HUD-52667 form. Although a utility schedule is similar in most respects to the HUD-52667 form, there are some subtleties that are hard to address in the HUD-52667 format. For example, in the HUD-52667

[21] Presumably, if HUD considers Summit's first recommendation, the thin data problem will no longer be an issue. However, this does not negate considering a new modelling methodology.

form, a fixed fee for a particular utility must be assigned to a particular end use, as there is not a "fixed fee" end use category on the form. Because the HUD-52667 form is an OMB-cleared form, changes cannot be made outside the overall form development schedule. However, HAs have latitude when publishing comprehensive utility allowance schedules, which are intended to work broadly across an HA's housing stock, and therefore can include customizations to reflect the particular needs of an HA, which may include dealing with multiple utility providers, "floor" and "ceiling" fees not related to a direct end-use, and other costs that may be applicable to particular tenants and not to others. Adding functionality, specific to the schedule, which produces a legend that delineates fees to users, would solidify HUSM as both a tool for developing comprehensive schedules and for individual allowances needed for the HUD-52667 Forms.

2. **Summit recommends implementing the capability for user log-in and storage of previous utility allowance estimates in the web-based HUSM tool.** As with many other web-based tools, a platform that would allow HUSM administrators to establish and log in to individual accounts that provide access to one or more stored utility allowances would open up the possibility for additional functionality that would benefit users. These types of functionality include:

    a) HUSM administrators could save allowances from previous years, and compare historical allowances with those currently implemented.
    b) Data from previous HUD-52667 forms could be autofilled from previous versions or previous years, allowing administrators to avoid re-entering information. This decreases administrator burden significantly, and would decrease data entry error.
    c) One administrator could access multiple HUD-52667 forms through one interface.
    d) Users could enter limited information (for example, just changes in rates), and easily determine, per HUD guidelines, when they should republish allowances because the change in total allowances between periods exceeds ten percent.
    e) Users could track when allowances were last updated, and the source of the change (e.g., HUSM estimates, consumption study, engineering study).

    Summit has designed the web version of HUSM with this additional functionality in mind. Even though the HUDUSER platform is, as of yet, not designed to support this capability, PD&R could consider hosting the tool on a platform that could support such storage capabilities.

Appendix 1 – EIA and AWWA 2014 State Average Utility Rates

The following table describes, by state, average 2014 utility rates for electricity[22], natural gas[23], water and sewer. In the case of water and sewer, data was published by AWWA.org[24], some utility bills also include a fixed fee; these fees are also included in the data. Where data was not published, not available (N/A) was recorded.

State	$/KWhr	$/gasMCF	$/Water Gallon	Water Fixed Fee	$/Sewer Gallon	Sewer Fixed Fee
AK	$0.19	N/A	$0.01	$37.17	$0.01	$7.87
AL	$0.12	$14.51	$0.00	$9.03	$0.00	$8.58
AR	$0.09	$10.35	$0.00	$4.85	$0.00	$5.77
AZ	$0.12	$17.31	$0.00	$15.54	$0.00	$2.50
CA	$0.16	$11.53	$0.00	$17.70	$0.00	$23.02
CO	$0.12	N/A	$0.00	$11.72	$0.00	$9.12
CT	$0.20	$14.14	N/A	N/A	N/A	N/A
DC	$0.13	N/A	$0.00	$3.86	$0.01	N/A
DE	$0.13	$13.21	N/A	N/A	N/A	N/A
FL	$0.12	$18.97	$0.00	$9.47	$0.00	$14.66
GA	$0.12	$14.69	$0.00	$9.85	$0.00	$5.50
HI	$0.37	$47.51	N/A	N/A	N/A	N/A
IA	$0.11	$9.99	$0.00	$8.50	$0.00	$9.44
ID	$0.10	$8.69	$0.00	$7.94	N/A	N/A
IL	$0.11	N/A	$0.00	$7.30	$0.00	$9.85
IN	$0.11	N/A	$0.00	$9.63	$0.00	$7.37
KS	$0.12	$10.57	$0.00	$10.60	$0.00	$10.24
KY	$0.10	$10.62	$0.00	$7.76	$0.01	$17.11
LA	$0.09	$10.91	$0.00	$4.58	$0.00	$14.04
MA	$0.17	$14.47	$0.00	$5.00	N/A	N/A
MD	$0.14	$12.33	$0.01	$9.13	$0.01	$19.18
ME	$0.15	$16.89	$0.00	$9.79	$0.01	$8.42
MI	$0.15	$9.29	$0.00	$11.80	$0.00	$11.25
MN	$0.12	N/A	$0.00	$3.50	$0.00	$1.98
MO	$0.11	$10.56	$0.00	$11.90	$0.00	$12.86
MS	$0.11	$9.43	N/A	N/A	N/A	N/A
MT	$0.10	N/A	N/A	N/A	N/A	N/A
NC	$0.11	N/A	$0.00	$6.93	$0.00	$7.22
ND	$0.09	$8.85	N/A	N/A	N/A	N/A

[22] http://www.eia.gov/electricity/monthly/current_year/february2015.pdf (Page 124)
[23] http://www.eia.gov/dnav/ng/ng_pri_sum_a_EPG0_PRS_DMcf_a.htm
[24] http://www.awwa.org/resources-tools/water-and-wastewater-utility-management/water-wastewater-rates.aspx (2014 Interactive database)

State	$/KWhr	$/gasMCF	$/Water Gallon	Water Fixed Fee	$/Sewer Gallon	Sewer Fixed Fee
NE	$0.10	$9.30	$0.00	$13.34	$0.00	$10.50
NH	$0.18	N/A	N/A	N/A	N/A	N/A
NJ	$0.16	N/A	$0.00	$8.45	N/A	N/A
NM	N/A	N/A	N/A	N/A	N/A	N/A
NV	$0.13	$11.44	$0.00	$15.16	$0.00	$14.00
NY	$0.20	$12.54	$0.00	$8.33	$0.01	$22.10
OH	$0.12	$10.14	$0.00	$8.00	$0.01	$6.52
OK	$0.10	$10.10	$0.00	$4.81	$0.01	$4.91
OR	$0.10	N/A	$0.00	$12.60	$0.00	$24.95
PA	$0.13	$11.68	$0.00	$6.13	$0.01	$7.31
PR	N/A	N/A	N/A	N/A	N/A	N/A
RI	$0.18	$15.14	$0.00	$12.62	N/A	N/A
SC	$0.12	$12.63	$0.00	$9.70	$0.00	$12.00
SD	$0.11	$9.35	N/A	N/A	N/A	N/A
TN	$0.10	$10.22	$0.00	$9.60	$0.00	$14.14
TX	$0.12	$11.02	$0.00	$11.57	$0.00	$11.80
UT	$0.11	$9.49	$0.00	$24.25	$0.00	$7.65
VA	$0.11	$12.26	$0.00	$7.62	$0.01	$15.48
VT	$0.18	$14.68	N/A	$246.00	N/A	N/A
WA	$0.09	$10.67	$0.00	$13.83	$0.00	$50.30
WI	$0.14	N/A	$0.00	$7.90	$0.00	$9.46
WV	$0.09	$10.17	$0.00	N/A	$0.00	N/A
WY	$0.11	$9.36	$0.00	$5.16	$0.00	$4.27