Analysis Plan

Project Name: Using social norms to decrease energy use in public housing

Project Code: 1808

Date Finalized: 9/24/2018

Data and Data Structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

Outcome Variables to Be Analyzed:

New York City Housing Authority (NYCHA) residents will be mailed four energy-feedback letters. The first set of letters were mailed on Friday, August 17th or Monday, August 20th and expected to arrive in residents' mailboxes on Tuesday, August 21st (based on the arrival of a letter sent to a known resident living in the same neighborhood). The two primary outcomes are energy use during two outcome periods:

- One week: the average daily energy use for one week after expected receipt of the first letter (August 22 - 28, 2018); and
- Eight week: the average daily energy use for eight weeks after expected receipt of the first letter (August 22 - October 16, 2018), which includes the letter mailing period and approximately two weeks after the last mailing.

Because mail would not arrive until the morning or afternoon on Tuesday, August 21st and the letters would have no chance at affecting energy use behavior until after residents' open the mail, the outcome period starts on Wednesday, August 22. In addition, some residents may not open their mail every day. To capture the same days of the week and eliminate any risk that the baseline period includes a treatment day, the baseline period will cover Wednesday, August 1 - Tuesday, August 14th.

Secondary outcomes are energy use during shorter time periods (e.g., day, morning/afternoon, by week) and past the eight week outcome period used in the primary analysis.

Raw data come from Wireless Energy Modules (WEMs) that record energy use over time similar to an odometer, meaning the number on the readout is increasing over time. The data include observations of unit-level energy readings at 15-minute intervals for a sample of public housing units in New York City. The analysis will use the 15-minute interval data to calculate energy use (i.e., by subtracting earlier readings from later readings). If for an unforeseen reason the

15-minute interval data are unavailable, the analysis will use WEM data that records unit-level energy use by day.

Transformations of Variables:

Energy use during baseline and outcome observation periods - Outcome variables of interest will be calculated from the raw 15-minute interval WEM readings. The WEMs record cumulative kilowatt-hour (kWh) readings at 15-minute intervals. Energy use (*Y*) over a given time period ($T = t_1 - t_0$) is calculated as:

$$Y = kWh_{t1} - kWh_{t0}.$$

The endpoints for calculating cumulative energy use over our outcome periods of interest are:

*t*0 = 00:00 hrs on the first day of the observation period

*t*1 = 00:00 hrs on the day following the last day of the observation period

Table 1: Observation Periods

Observation Period	Start (HH:MM day)	Stop (HH:MM day)
Baseline (August 1 - August 14)	00:00 August 1	00:00 August 15
One week (August 22 - August 28)	00:00 August 22	00:00 August 29
Eight week (August 22 - October 16)	00:00 August 22	00:00 October 17

Some units may not have WEMs readings for every 15-minute interval, including for the observation period endpoints. In these cases endpoint WEMs readings for these units will be chosen as following: $t^\star_0$ = the first valid WEMs reading following 00:00 hrs on the first day of the observation period; for $t^\star_1$, the latest valid WEMs reading prior to 00:00 hrs on day following the last day of the observation period. For these units define:

$$T^\star = t^\star_1 - t^\star_0$$

$$Y^\star = kWh_{t^\star 1} - kWh_{t^\star 0}.$$

Total energy use during the observation period for units with missing endpoints is calculated as:

$$Y^\star = Y_{T^\star} \times (T/T^\star).$$

The calculation assumes the average energy use over the observed period is the same as over the missing periods. The threshold for making the extrapolation is discussed in the Section: Treatment of Missing Data.

Bedrooms for blocking - The data field for the number of bedrooms for each unit takes the following values: 'EF' (efficiency), '02' (studio), '03' (1 bedroom), '04' (2 bedroom), '05' (3 bedroom), '06' (4 bedroom), and '07' (5 bedroom). Bedroom blocks are created by grouping units into the following bedroom groups:

Table 2: Bedroom Blocks

Block (bedroom group)	Bedrooms
1	efficiency; studio; 1 bedroom
2	2 bedroom
3	3 bedroom
4	4 bedroom; 5 bedroom

Imported Variables:
In addition to raw energy use data, imported data will include blocking variables (i.e., development, bedroom groups -- efficiency/studio or one bedroom, two bedrooms, three bedrooms, and four or five bedrooms), a dichotomous indicator for assignment to the treatment or control group, baseline energy use, and an indicator for imputed baseline energy use.

Transformations of Data Structure:
For the primary analysis, energy use data will be aggregated to the unit level for the outcome periods of interest. A measure of average daily energy use will be created by dividing aggregate energy use by the number of days in the outcome period. A discussion of handling missing or partially missing energy use data follows below.

Data Exclusion:
As in the pilot study, we will exclude energy use records that we expect are due to malfunctioning sensors. We will examine the distribution of energy use data (blind to treatment status) to determine the criteria for exclusion. Excluded observations and missing data are treated the same way (see below).

Treatment of Missing Data:
Missing data will influence the study at three points: (a) for the creation of the baseline covariate on energy use, (b) for the creation of energy feedback letters, and (c) for the creation of the outcome variables on energy use. For units with missing data for our primary analysis:

- When units are missing baseline energy data for a given day (including excluded data due to malfunctioning sensors), energy use will be imputed as the mean energy use for units of the same size in the same building that had valid data (or development if all same size units are missing energy use data) for the day energy use is missing. A dichotomous missing indicator will be created for each baseline day. The value will be set to one for units missing data for that day.

- Feedback letters for units missing data will include a horizontal bar chart with the energy use of efficient neighbors and all neighbors. Because these letters will not include

personalized energy use feedback, they also cannot include injunctive norm feedback. Instead, the letters will encourage residents to "Be an energy-efficient neighbor!"

For the primary analysis, units sent non-personalized letters (or control units with missing data) will be excluded from analysis of one-week effects, and units sent more than two non-personalized letters (or control units with missing data for more than two periods) will be excluded from analysis of eight-week effects. For sensitivity checks, these units will be included to test the robustness of the primary findings.

- The analytic sample will exclude units with missing data for more than 75 percent of the outcome period of interest. That is, included units are those such that $T^{\star} \geq T \times 0.75$. For example, for the one-week outcome period $T = 10,080$ minutes, so non-missing units are those where $T^{\star} \geq 7,560$ minutes. Since the outcome periods differ for the two primary research questions, the analytic sample may also differ. For instance, in the extreme, a unit may have no missing data in the first week and then missing data for the following seven weeks or vice versa. To maximize the use of available data the sample exclusion criteria will be based on the outcome period for the specific research question (one or eight weeks). As a sensitivity check, the analytic sample will be adjusted such that the sample includes the same units for both primary research questions. That is for the sensitivity check, the analytic sample will be limited to units with 75 percent or more data in the first week and in the eight week outcome periods. In addition, to bound estimates of the effects, missing/excluded data will be imputed at the 1st and 99th percentile for units of the same size in the same building (or development if all same size units are missing energy use data).

Statistical Models & Hypothesis Tests

This section describes the statistical models and hypothesis tests that will make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models:

Research question 1: Do public housing residents reduce average daily energy use in the week after they are sent the first feedback letter on their energy use?

Research question 2: Do public housing residents reduce average daily energy use in the eight week period when they are sent four feedback letters on their energy use?

The primary statistical model for both research questions is:

$$Y_{ib} = \beta_0 + \beta_1(T_{ib}) + \pi Z'_{ib} + \alpha_b + \varepsilon_{ib}$$

where *i* indexes units in blocks *b*, and

Y_{ib}: is average daily energy use during the one-week or eight-week outcome period;
T_{ib}: is an indicator for random assignment to the letter group;
Z'_{ib}: is a vector of unit-level average daily energy use during the baseline period and an indicator for missing baseline energy use;
α_b: are block fixed effects (development by bedroom size group by missing baseline energy use); and
ε_{ib}: is an error term.

The parameter of interest is β_1 which is the average intent to treat (ITT) effect of being mailed energy-use feedback letter(s) to public housing residents on average daily energy use. A covariate for energy use during the baseline period is included to enhance precision and statistical power. An indicator for imputed baseline energy is included to account for possible differences in energy use between units where baseline energy data were available and units where baseline energy data needed to be imputed. Block fixed effects are included in the model, because randomization occurred within blocks (development by bedroom size group) to enhance precision.

Heteroskedasticity robust (HC2) standard errors will be used.

Inference Criteria, Including Any Adjustments for Multiple Comparisons:
Standard inference criteria will be used: two-tailed tests and three threshold *p*-values: 1%, 5%, and 10%.

The primary analysis examines the effects of the letters on energy use during one week and eight week outcome periods. No adjustments for multiple comparisons will be made, since the first week of data is embedded in the eight weeks of data and only two statistical tests are planned.

A meaningful effect size would result in energy savings that at least cover the cost of printing and mailing the intervention. Although the exact "break-even" point is difficult to estimate because of uncertainties about utility costs associated with units mailed letters, we expect that any effect size greater than about 2.5 percent will represent a cost-effective intervention.

Limitations:
A primary limitation to this study is uncertainty about the size of the analytic sample, since WEMs sometimes do not operate properly, which results in missing data. Missing data could limit precision and reduce statistical power to detect treatment effects.

A second limitation is that energy use for this sample is driven by window air conditioning units. As a result, households use the most energy and have the most opportunity to reduce energy use on hot days. Because letters were mailed later than originally anticipated when weather may be cooler, residents may have fewer actions to take to reduce energy use in response to the letters, which could attenuate the effectiveness of the letters at reducing energy use. Moreover, during

the baseline period New York City had a hot weather emergency. Hot weather during the baseline period may result in baseline energy use having a lower correlation with energy use during the outcome period that in the pilot study. If this is the case, minimal detectable effects would be larger than anticipated.

A third limitation is that residents may not open the energy feedback letters. However, residents are expected to receive the letters since mailing addresses, which come directly from the public housing authority (PHA), should be accurate and as a result letters should be unlikely to be returned to the sender. In addition, messaging on the envelope may affect residents' energy use behaviors even when residents do not open the letters.

Exploratory Analysis:

Exploratory analysis will examine if feedback letters affect energy use by day or week during the outcome period, by outside temperature (e.g, relatively hot or cold days), and by time of day. In particular, exploratory analysis will examine the attenuation of effects over time.

Subgroup analysis will examine whether effects differ based on baseline energy use, bedroom size, development, and, if data are available by household demographics (e.g., work status, language spoken, number of residents, etc.). Since some mailings will include a coloring sheets for kids, additional subgroup analysis will examine whether effects differ when a household has or does not have a child 10 years old or younger.