**Section 1**

i) The Logit Model (Table 1) yields 4 significant variables out of 41 (ignoring the constant), which are highlighted in red. So treatment and control are not balanced. The dummies are not very concerning, but it is concerning that people with higher past consumption are more likely to be treated.

**Table 1. Logit Regression Results**

==============================================================================

Dep. Variable: T No. Observations: 652

Model: Logit Df Residuals: 610

Method: MLE Df Model: 41

Date: Sat, 11 Apr 2015 Pseudo R-squ.: 0.4853

Time: 17:11:22 Log-Likelihood: -189.27

converged: True LL-Null: -367.74

LLR p-value: 5.188e-52

===============================================================================

coef std err z P>|z| [95.0% Conf. Int.]

-------------------------------------------------------------------------------

kwh\_2009\_08 0.0017 0.003 0.640 0.522 -0.004 0.007

kwh\_2009\_09 0.0102 0.004 2.852 0.004 0.003 0.017

kwh\_2009\_10 -0.0041 0.004 -1.138 0.255 -0.011 0.003

kwh\_2009\_11 0.0038 0.003 1.239 0.216 -0.002 0.010

kwh\_2009\_12 0.0084 0.002 4.208 0.000 0.005 0.012

D\_410\_2 0.3894 0.735 0.530 0.596 -1.052 1.831

D\_410\_3 0.6943 0.751 0.925 0.355 -0.777 2.166

D\_420\_2.0 -0.2471 0.594 -0.416 0.677 -1.411 0.916

D\_420\_3.0 -1.3566 0.682 -1.989 0.047 -2.693 -0.020

D\_420\_4.0 -1.0793 0.815 -1.324 0.186 -2.677 0.519

D\_420\_5.0 -1.8734 1.224 -1.530 0.126 -4.273 0.526

D\_43111\_2.0 -0.6987 0.549 -1.273 0.203 -1.774 0.377

D\_43111\_3.0 -1.5845 0.902 -1.758 0.079 -3.351 0.182

D\_43111\_4.0 -1.1097 2.070 -0.536 0.592 -5.168 2.948

D\_405\_2 -0.3503 0.332 -1.056 0.291 -1.000 0.300

D\_43521\_2.0 0.0527 0.662 0.080 0.936 -1.244 1.350

D\_43521\_3.0 -0.2625 0.582 -0.451 0.652 -1.403 0.878

D\_43521\_4.0 0.3391 0.609 0.557 0.578 -0.855 1.533

D\_43521\_5.0 0.0676 0.748 0.090 0.928 -1.399 1.534

D\_43521\_6.0 0.6399 1.054 0.607 0.544 -1.425 2.705

D\_43521\_7.0 0.3522 0.637 0.553 0.581 -0.897 1.601

D\_470.1\_1 -1.0992 1.116 -0.985 0.324 -3.286 1.087

D\_470.2\_1 -0.0161 0.706 -0.023 0.982 -1.400 1.368

D\_470.3\_1 -0.6860 0.694 -0.988 0.323 -2.046 0.674

D\_470.4\_1 -0.8305 0.532 -1.561 0.119 -1.873 0.212

D\_470.5\_1 0.3733 0.395 0.945 0.344 -0.401 1.147

D\_4701.1\_1 -0.7261 0.531 -1.368 0.171 -1.766 0.314

D\_4701.2\_1 -0.8409 0.315 -2.670 0.008 -1.458 -0.224

D\_4701.3\_1 0.0595 0.928 0.064 0.949 -1.760 1.879

D\_4701.4\_1 -0.4547 0.608 -0.748 0.454 -1.646 0.737

D\_4701.5\_1 -0.3394 0.423 -0.802 0.423 -1.169 0.490

D\_4701.6\_1 -0.8164 0.507 -1.610 0.107 -1.810 0.177

D\_4701.7\_1 0.3318 1.244 0.267 0.790 -2.107 2.771

D\_4701.8\_1 -1.0049 1.910 -0.526 0.599 -4.747 2.738

D\_5414\_2 -0.8294 1.040 -0.797 0.425 -2.868 1.210

D\_5414\_3 -0.3678 0.364 -1.011 0.312 -1.080 0.345

D\_5418\_2 -0.6003 1.097 -0.547 0.584 -2.751 1.550

D\_5418\_3 0.1588 1.072 0.148 0.882 -1.943 2.260

D\_5418\_4 -0.0408 1.058 -0.039 0.969 -2.114 2.033

D\_5418\_5 -0.1977 1.076 -0.184 0.854 -2.306 1.911

D\_5418\_6 -0.1935 1.149 -0.168 0.866 -2.446 2.059

const -6.3048 1.388 -4.543 0.000 -9.025 -3.585

===============================================================================

The Quick Means Comparison also show imbalance, but they show something a little different. Table 2 is a list of the significant variables (|t-stats| > 2), along with their t-stats:

**Table 2. Quick Mean Results**

========================

kwh\_2009\_08 14.739744

kwh\_2009\_09 15.203416

kwh\_2009\_10 14.877134

kwh\_2009\_11 14.984830

kwh\_2009\_12 15.887552

D\_410\_3 5.556768

D\_420\_2.0 2.320252

D\_420\_4.0 2.983374

D\_43111\_2.0 3.392115

D\_405\_2 -6.508454

D\_43521\_4.0 2.419806

D\_5418\_2 -3.577503

========================

(ii) In the Logit imbalance test, August kwh, December kwh, the dummy for having 3 people over the age of 15 in the house (420.3), and the dummy for having an immersion electric water heater (4701.2) are considered significant. In the "Quick Means Comparison", significant variables are listed in Table 2.

More variables are considered significant in the Quick Means Comparison. In particular, the differences in consumption are significant in ALL months, whereas only two months were significant in the Logit. This is probably due to the fact that the consumption values are highly collinear.

iii) The benefit of the Logit is that one can consider the imbalance of one variable, holding other variables constant. It's entirely likely that the groups look different on average, but look similar when you control for observables. This would tell us whether any selection that exists was selection on observables or unobservables. The problem is we do not know how different each variable is, since we are not comparing the original data.

On the other hand, by comparing means we get a sense of that different groups look like as a whole. We know the difference for each variable. However, we are not able to check if one variable is still significantly different between treatment and control after controlling for other variable(s).

iv) Questions 410, 420, and 43111 contain very similar information: namely, the age makeup of the household. In particular, given the answers to questions 420 (How many people over 15 years of age live in your home?) and 43111 (How many people under 15 years of age live in your home?), you can determine the answer to 410 (which asks whether there any children under 15 years of age in the home). I would have wanted to include more variables that are typically associated with consumption, like income and house size.

Also, the potential answers to 410 are not mutually exclusive, making it a bad question. The second two possible answers are "All people in my home are over 15 years of age" and "Both adults and children under 15 years of age live in my home". Clearly one of these must be true, but in the data we observe about 25% of the respondents saying neither is true. It seems that these people must have opted for the other option given, "I live alone." But then the second statement must also be true (since only adults can presumably participate in the program)! In short, this is a terrible question.

**Section 3**

i) The results are in Tables 3 and 4 below. Note that the effect of interest is the effect of treatment during the treatment period, which is captured by the coefficient on the treatment\*trial variable. The un-weighted regression shows no significant effect of this variable: a coefficient of -0.008, with a p-value of 0.54.

By contrast, the weighted regression shows a significant -0.025, or a 2.5% reduction in electricity use attributable to the program. In both regressions, the seasonal dummies show higher use during the winter months, which makes sense because of higher heating demand. The coefficient on the trial period is negative in both regressions.

ii) Without weights, we would conclude (at the 5% and 10% significance level) that the program had no effect on electricity usage. Thus, as a consultant of CER, we must conclude that this treatment had no effect.

iii) With weights, we find a strongly significant effect. Not only is the effect now significant, but the effect is estimated with larger precision. Hence, we

would inform the CER that the program was effective.

iv) Since the data was biased, our estimated effect from the weighted regression depends heavily on our propensity score calculation being correct. If we had chosen different input variables for calculating the propensity scores, we may well have gotten very different answers. Hence, the believability of the answers depends on how sensitive we find the results to be. One approach would be to redo these calculations using different input variables to see if the estimated effect of the treatment changes dramatically. If it does, it would reduce our confidence in the estimate. If the results seem stable, then perhaps we are comfortable with them.

**Table 3. OLS Regression Results**

==============================================================================

Dep. Variable: log\_kwh R-squared: 0.193

Model: OLS Adj. R-squared: 0.192

Method: Least Squares F-statistic: 155.6

Date: Sun, 19 Apr 2015 Prob (F-statistic): 8.07e-265

Time: 10:48:19 Log-Likelihood: 697.35

No. Observations: 5868 AIC: -1377.

Df Residuals: 5859 BIC: -1317.

Df Model: 9

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

trial -0.0201 0.012 -1.626 0.104 -0.044 0.004

trt&trial -0.0080 0.013 -0.617 0.537 -0.034 0.017

ym\_2009\_09 -0.0175 0.012 -1.471 0.141 -0.041 0.006

ym\_2009\_10 0.0774 0.012 6.497 0.000 0.054 0.101

ym\_2009\_11 0.1370 0.012 11.503 0.000 0.114 0.160

ym\_2009\_12 0.2780 0.012 23.344 0.000 0.255 0.301

ym\_2010\_01 0.2858 0.012 23.995 0.000 0.262 0.309

ym\_2010\_02 0.1080 0.012 9.068 0.000 0.085 0.131

ym\_2010\_03 0.1194 0.012 10.028 0.000 0.096 0.143

==============================================================================

Omnibus: 2940.250 Durbin-Watson: 1.194

Prob(Omnibus): 0.000 Jarque-Bera (JB): 401634.697

Skew: -1.374 Prob(JB): 0.00

Kurtosis: 43.437 Cond. No. 4.53

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

**Table 4. Weighted LS Regression Results**

==============================================================================

Dep. Variable: y R-squared: 0.192

Model: OLS Adj. R-squared: 0.190

Method: Least Squares F-statistic: 154.3

Date: Sun, 19 Apr 2015 Prob (F-statistic): 8.05e-263

Time: 10:48:19 Log-Likelihood: -314.58

No. Observations: 5868 AIC: 647.2

Df Residuals: 5859 BIC: 707.3

Df Model: 9

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

trial -0.0222 0.011 -1.940 0.052 -0.045 0.000

trt&trial -0.0253 0.010 -2.615 0.009 -0.044 -0.006

ym\_2009\_09 -0.0299 0.010 -2.944 0.003 -0.050 -0.010

ym\_2009\_10 0.0491 0.010 4.825 0.000 0.029 0.069

ym\_2009\_11 0.0982 0.010 9.657 0.000 0.078 0.118

ym\_2009\_12 0.2296 0.010 22.571 0.000 0.210 0.250

ym\_2010\_01 0.2396 0.010 23.547 0.000 0.220 0.259

ym\_2010\_02 0.0846 0.010 8.313 0.000 0.065 0.105

ym\_2010\_03 0.0855 0.010 8.402 0.000 0.066 0.105

==============================================================================

Omnibus: 2492.920 Durbin-Watson: 1.238

Prob(Omnibus): 0.000 Jarque-Bera (JB): 169546.835

Skew: -1.185 Prob(JB): 0.00

Kurtosis: 29.226 Cond. No. 4.98

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified