z5254903.doc

by George Bai

ECON 3210 REPORT

Individual Group

Author: George Bai

Z ID: z5254903

EVALUATION OF ENERGY AUDITS

PREPARED BY

DH CONSULTING

July 2021

EXECUTIVE SUMMARY

- DH Consulting has been requested by the Department of Energy to evaluate the impact of energy audits on household energy expenditure. The Department of Energy's preliminary study had incomplete data and raised internal modelling concerns. Additional data and desires for improved analysis have motivated a reassessment of the evidence.

- The model indicates that audits reduce energy expenditure at the household level by an average of $25.55, indicating a roughly 2.5% reduction in annual energy consumption.

- Audits have different impacts across districts. In the preliminary study district, audits had on average no impact on energy expenditure. In the other district, audits reduced energy expenditure by an average of $35.14, plausibly reflecting greater impacts in more variable weather conditions.

- Households that received previous audits within the last two years are not impacted on average by another audit. This suggests that audits do not have temporary effects, but rather cause long-term reductions in energy expenditure.

- It is plausible that targeted energy audits in locations with high weather variation accumulate significant energy savings over the years. Hence, targeted implantation of the auditing program is recommended.

Introduction

The Department of Energy has requested DH Consulting to evaluate the impact of energy audits on residential energy consumption, and whether this impact differs across districts. As their own internal study was based on poor modelling and incomplete data, this has motivated a reassessment of the evidence.

Additional data from another district is added to the data sample. This is expected to improve average treatment effects due to stronger impacts in areas with more volatile weather conditions. Moreover, their model included expenditure data from one time period rather than two and lacked satisfactory controls.

Hence, DH Consulting has modelled the average treatment effect in order to give more precise estimates. Furthermore, we have decomposed those treatment effects across districts to determine how location affects the causal impact of energy audits.

Furthermore, we test if weather differences explain the differing impacts of audits across districts. Finally, we determine whether audits have either long-lasting or temporary impacts.

The results suggest that the audit program is plausibly effective when targeted properly across districts. Hence DH consulting suggests implementing the auditing program.

Data Analysis

The econometric analysis involves a study of 5510 households, with 59.74% being treated via the energy audit. Non-treated households represent the control group and act as the baseline with which treated households are compared to.

Randomisation Check

An initial exploratory look into the data finds that in the year before the experiment was conducted, the mean annual energy expenditure in the treated group was $1099. Meanwhile in the control group, the mean annual energy expenditure was $1102.

Moreover, all other variables have similar means and standard deviations except for area, which has more than double the standard deviation in the treatment group than the control group

This is represented in table 1, which separates the treated group and the control group, with corresponding mean and standard deviation for all non-weather variables.

Table 1: Balancing Treatment and Control Groups

	Control Group (N = 2218)	Treatment Group (N = 3292)
expend0	1098.79 (581.31)	1102.25 (586.55)
expend	1252.41 (659.26)	1231.63 (661.57)
area	168.05 (97.94)	174.32 (207.21)
pool	0.11 (0.31)	0.12 (0.32)
aircon	0.26 (0.44)	0.25 (0.43)
rc_aircon	0.16 (0.36)	0.18 (0.38)
cheat	0.16 (0.36)	0.16 (0.37)
dish	0.32 (0.47)	0.33 (0.47)
mwave	0.90 (0.30)	0.91 (0.29)
dryer	0.80 (0.40)	0.80 (0.40)
ftype	3.66 (1.17)	3.58 (1.19)
people	4.47 (2.35)	4.48 (2.39)
inc	5.63 (2.65)	5.61 (2.61)
hinc	0.53 (0.50)	0.53 (0.50)
audit	0.01 (0.09)	0.01 (0.10)

For the variable area, the large differences in standard deviation are driven mostly by a disproportionate number of outliers assigned to the treatment group rather than the control group. If the three largest outliers in terms of area are excluded from the treatment group, then the mean decreases to 168.79, while the standard deviation falls to 87.49.

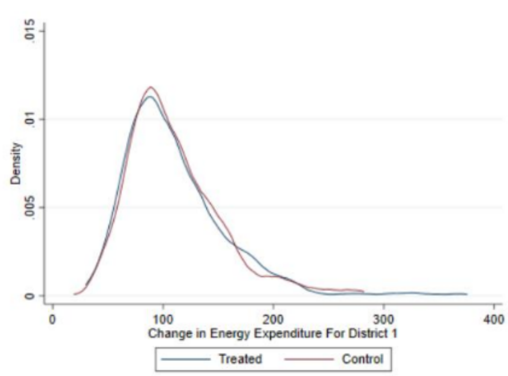
Hence, there is balancing between the treatment and control group, implying proper randomisation within the research design.

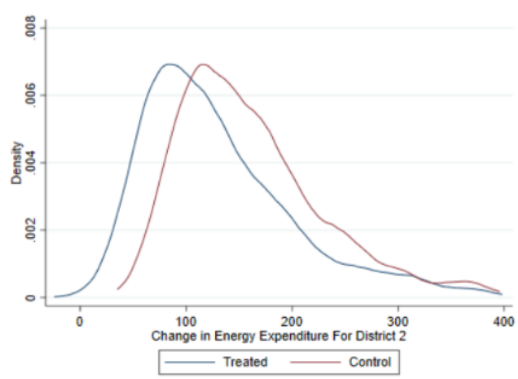
District Data

Households are divided into two districts: District 1 and District 2. In the Department of Energy preliminary report, only District 1 was considered for analysis. Here, District 2 data supplements the analysis.

An initial look into the data for people whose change in energy expenditure is below $400 in the graphs below indicate that in District 1, the treatment appears to have little impact. However, in District 2, the treatment appears to reduce energy expenditure in a systematic manner across households.

This implies that the energy audit impacts District 2 differently due to some district-wide factor, providing suggestive evidence in favour of weather variability.

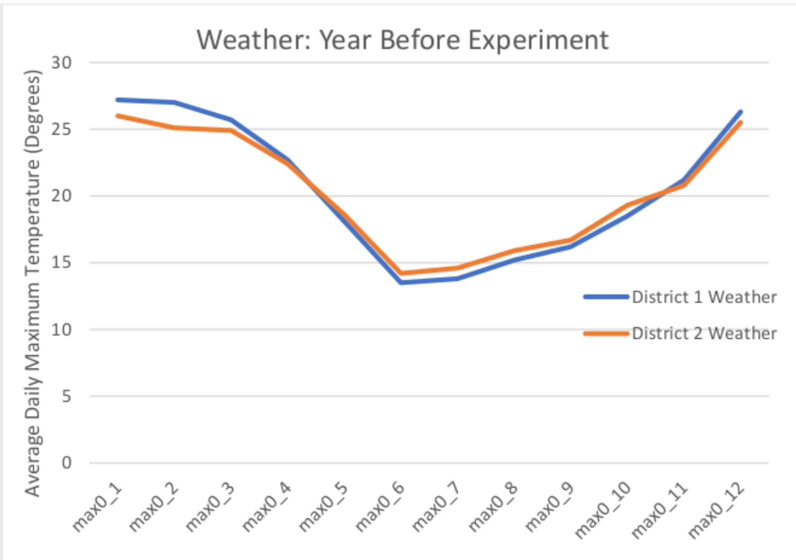




Weather Data

We also note that weather patterns in the data are sorted into six distinct groups, decomposed into two weather patterns for District 1 and four weather patterns for district 2. As such, it is plausible to match locations with similar weather patterns, each in different districts.

Hence, matching a pair of locations where the average maximum daily temperature in the year before the experiment was 27.2℃ in the District 1 location and 26℃ in the District 2 location, we plot their weather patterns below.



Evidently, there are similarities across these weather patterns. These similarities persists in the year after the experiment, and when considering the average minimum daily temperature.

Audit Data

Finally, we consider the subsample that has received an audit within the last two years. There are only 53 observations, with 20 in the treatment group and 33 in the control group. This represents a small sample size, where imbalances unrelated to observable household characteristics or weather may drastically impact on analysis.

Issues are especially plausible due to the self-selection of households who receive energy audits, who may differ from the general population via culture, ability, insight, and many other unobservable characteristics.

Hence, this data is useful only for suggestive purposes.

Econometric Analysis

A difference in difference approach was used in order to determine the impact of the energy audit on energy expenditure.

This requires an assumption of parallel trends, meaning that the differences in energy expenditure in the treatment group and the control group would've remained the same over time, were it not for the energy audit.

As the treatment was allocated randomly, the observable characteristics of the households in the treatment group and the control group are similar in many ways as described in the Data Analysis section.

It is implausible that any time-varying confounders could have impacted the treatment group without impacting the control group as well in a similar manner due to this similarity, except at a minor, idiosyncratic level.

However, the sample has finite size, which is exacerbated by randomisation being conducted at the neighbourhood level rather than the household level. As such, the sample may have covariates that differ by chance from the true population sample.

Hence, the addition of covariates with predictive power to the regression enables more precise estimates of the energy audit's effects to be attained. As such, Double Selection is useful in improving the precision of the estimates as it selects predictive covariates well.

Double Selection Regression

The model produced is in the form of the one below:

$$\Delta expend_i = \beta_0 + \beta_1 treat_i + controls + u_i$$

The dependent variable $\Delta expend_i$ represents the change in energy expenditure from the year before the audit to the year after the audit.

In the Department of Energy's preliminary study, the variable $expend$ was used, which represented the energy expenditure in the year after the audit. However, accounting for the previous year's energy expenditure improves the balance between the treatment and control group, improving the precision of the estimates.

The initial list of controls included all variables except customer id and expenditure-based variables, as well as the interaction between income and characteristics of the household.

A Double Selection procedure culled those set of controls into the most predictive ones, with the excluded variables unlikely to have too much omitted variable bias due to their low predictive power.

Estimates provided by the Double Selection procedure for β_1 in this model are presented in the table below:

Table 2 Regression Results (Average Treatment Effects)

	OLS (No controls)	OLS (All controls)	Double Selection
treat	[illegible].24***	-26.38***	-25.55***
	([illegible]91)	(1.595)	(1.652)
Observations	5510	5510	5510

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

The results suggest that audits reduce energy expenditure at the household level by an average of $25.55. This average treatment effect is robust to using different models, involving either no controls or the initial set of controls.

All results are statistically significant at the 0.1% level, indicating a precise effect. However, this represents only an approximately 2.5% reduction in annual energy expenditure, using the year before the experiment as a baseline.

This is a larger estimate than the one in the Department of Energy preliminary study. Hence, this indicates that this effect may be driven by audits having improved impact in District 2. As such, a fully moderated model is used in order to test for treatment effect heterogeneity.

FMM Double Selection: By District

The model produced is below:

$$\Delta expend_i = \beta_0 + \beta_1 treat_i + \beta_2 district_i + \beta_3 treat_i \cdot district_i$$
$$+ controls + district_i \cdot controls + u_i$$

Instead of adding in a single interaction between district and treatment as one would do in a single interaction model (SIM), the fully moderated model (FMM) interacts district and all other controls.

Due to this, the model has improved flexibility, meaning it protects against omitted variable bias better than the single interaction model.

Hence, using a Double Selection procedure, estimates for β_1, β_2 and β_3 in this FMM are presented in Table 3.

Table 3 Regression Results (District Average Treatment Effects)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-0.412	-0.250	-0.230
	(2.607)	(2.245)	(2.249)
district=2	56.58***	50.38***	47.62***
	(2.893)	(2.939)	(2.749)
treat=1 # district=2	-31.86***	-34.38***	-34.29***
	(3.756)	(2.990)	(2.998)
Constant	111.5***		
	(2.009)		
Observations	5510	5510	5510

Standard errors in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

In District 1, the audit caused an average $0.23 reduction in energy expenditure. This is statistically insignificant, and the null effect falls within the 95% confidence interval. Meanwhile, in District 2, the audit reduced average household energy expenditure by $34.52, which is statistically significant at the 0.1% level.

These results are robust to changes in specification, as they remain statistically significant even in the single interaction models.

The estimates suggest that the impact of energy audits on the initial Double Selection regression were driven entirely by their District 2 impact. Hence, the cost-effectiveness of energy audits is dependent on how effectively they can be targeted. At an impact of $34.52, the audit program is likely cost-effective if targeted well.

However, there are still unknowns regarding what drives this difference across districts. A plausible suggestion is that more variable weather patterns explain differences in the audit's impact across districts.

We conduct a simple test of that assumption by comparing similarly matched weather patterns across districts.

Weather Across Districts

In the Data Analysis section, the weather patterns corresponding to min0_1 = 27.2℃ and min0_1 = 26℃ matched closely.

Dropping all other observations apart from these two weather patterns, we repeat the Double Selection procedure across districts for the fully moderated model. We find that the audit was $31.38 more impactful in reducing energy expenditure in households in District 2 compared to District 1.

Table 4 Regression Results (Average Treatment Effects On Locations With Similar Weather Patterns)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-0.885	0.284	0.284
	(2.623)	(2.137)	(2.137)
district=2	51.21***	51.44***	51.44***
	(5.338)	(5.140)	(5.140)
treat=1 # district=2	-27.19***	-31.38***	-31.38***
	(6.856)	(5.801)	(5.801)
Constant	107.0***		
	(2.028)		
Observations	1615	1615	1615

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

As such, this provides evidence that average temperatures over a month provide minimal ability to distinguish the reason why audits impact the two districts differently.

However, there is higher variability within the temperature data in District 2 than in District 1, which has mostly homogenous weather conditions. Moreover, in the Data Analysis section, it was demonstrated that energy audits appeared to have a homogenous impact across most households.

Other possible explanations remain implausible. For example, differences in culture across districts seem unlikely to generate such homogenous impacts. There may be differences in local laws between the two districts, but this appears implausible.

Hence, weather still remains the most plausible explanation of the difference in audit impact via differences in intra-day and day to day variations in weather. It is likely that averaging daily data removes weather's explanatory power.

Here, we caution for more research to be done in determining whether temperature differences are truly the differentiator, as effective targeting of the audit program is necessary for the program to be cost-effective.

Another factor required for audits to be cost-effective are long-term effects. Hence, the analysis ends with tests in that regard.

Double Selection FMM: Audit

The final model is considered below:

$$\Delta expend_i = \beta_0 + \beta_1 treat_i + \beta_2 audit_i + \beta_3 treat_i \cdot audit_i$$
$$+ controls + audit_i \cdot controls + u_i$$

The estimates for β_1, β_2 and β_3 in this FMM are presented in Table 5.

Table 5 Regression Results (Average Treatment Effects On Already Audited Households)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-24.12***	-26.19***	-25.63***
	(2.204)	(1.668)	(1.686)
audit=1	21.47	-5.802	-17.10
	(14.91)	(12.00)	(45.53)
treat=1 # audit=1	-13.95	1.488	-5.919
	(18.96)	(16.91)	(16.36)
Constant	153.4***		
	(1.726)		
Observations	5510	5510	5510

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Here, energy audits impact already-audited households by reducing energy consumption by $5.91. This is a statistically insignificant effect.

Interpreting these results, it is plausible that energy-audits have long-lasting impacts and cause continued reductions in energy expenditure, instead of being temporary as households relapse into previous energy-use habits.

However, the sample size of previously audited households is small. Moreover, there is selection bias, as people who paid for an audit may be more prone to keeping expert advice in mind instead of discarding it, as they justify it to themselves more.

Hence, more research is required into the long-term benefits of energy audits. Nevertheless, it appears plausible that household behaviour can remain changed across multiple years, though this remains weakly suggestive evidence.

Conclusion

The impact of energy audits differs across districts, and hence strict targeting is required to make audits cost-effective. However, there are still many unknowns in determining how to properly target energy audits, with further tests of how variability in weather impact the effectiveness of energy audits required.

Nevertheless, the benefits of audits in more responsive districts are large enough to plausibly satisfy cost-benefit calculations. With weak suggestive evidence that the benefits of energy audits last more than two years, the impact of energy audits are large enough that DH consulting support the implantation of the auditing program

Good project! Key features of the data were well characterized and you did a good job in producing and evaluating econometric models, especially the use of the double selection method and checking for the effect of the audit variable. However, it is important to discuss the choices of controls in your models.

Appendix

----

name: <unnamed>

log: \\Client\C$\Users\georg\Downloads\Appendix3.txt

log type: text

opened on: 30 Jul 2021, 03:54:45

.

. cd "\\Client\C$\Users\georg\Downloads\"

\\Client\C$\Users\georg\Downloads

. import excel using en_audit.xls, firstrow

(66 vars, 5,510 obs)

.

. set seed 1511

.

. ** This sets up the basic data generation **

. **<><><><><><><><>**

. * This sets up the randomised data in the base year

. gen expend0 = expendbase + 10*rnormal()

. * It is reasonable to assume that the effect of area on expenditure is better described in percentage increases in area

. * rather than linear terms. Each additional increase of are on expenditure should have decreasing impact

. gen logarea = log(area)

```
. * This sets up the DID regression

. gen change_expend = expend - expend0


. **<><><><><><><><><>**

.

.

. ** This is the beginning of the data exploration**

. **<><><><><><><><><>**

. * This clarifies the sample size of audit across districts

. bysort treat district: summarize audit if audit == 1


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

-> treat = 0, district = 1


   Variable |       Obs        Mean    Std. Dev.       Min        Max

-------------+---------------------------------------------------------

      audit |         4           1           0          1          1


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

-> treat = 0, district = 2


   Variable |       Obs        Mean    Std. Dev.       Min        Max

-------------+---------------------------------------------------------

      audit |        16           1           0          1          1


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

-> treat = 1, district = 1
```

```
   Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      audit |          6           1           0          1          1

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

-> treat = 1, district = 2

   Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      audit |         27           1           0          1          1



.
. * This creates a word file that contains the mean and standard deviations of all non-weather variables
. summtab, by(treat) contvars(expend0 expend area pool aircon rc_aircon cheat dish mwave dryer ftype people inc hinc audit) mean landscape replace word wordname(summary_table5) title(Table 1: Balancing Treatment an
> d Control Groups)
successfully replaced "//Client/C$/Users/georg/Downloads/summary_table5.docx"
Table 'summary_table5' saved in current working directory

.
. * This checks for imbalances within the audited households in the treatment and control group
. preserve

. drop if audit == 0
(5,457 observations deleted)
```

. stddiff expend0 expend area pool aircon rc_aircon cheat dish mwave dryer ftype people inc hinc, by(treat)

	treat=0		treat=1		
	Mean or N	SD or (%)	Mean or N	SD or (%)	Std Diff
expend0	1271	507.63	1156	476.62	0.23331
expend	1446	573.52	1293	541.33	0.27426
area	205.9	135.06	157.1	39.921	0.49008
pool	.15	.36635	.2121	.41515	-0.15867
aircon	.2	.41039	.2424	.43519	-0.10030
rc aircon	.15	.36635	.2121	.41515	-0.15867
cheat	.3	.47016	.2121	.41515	0.19814
dish	.2	.41039	.3939	.4962	-0.42594
mwave	.75	.44426	.9394	.24231	-0.52929
dryer	1	0	.8485	.36411	0.58849
ftype	4.1	.85224	3.818	1.1307	0.28149
people	6.15	1.8994	4.303	2.2843	0.87923
inc	7.8	1.6733	7.515	1.253	0.19270
hinc	1	0	1	0	.

. restore

.

. * This visualises the high degree of standard deviation in weather

. stddiff expend0 expendbase expend area pool aircon rc_aircon cheat dish mwave dryer ftype people inc hinc audit max0_1 max0_2 max0_3 max0_4 max0_5 max0_6 max0_7 max0_8 max0_9 max0_10 max0_11 max0_12 max1_1 max1_2

> max1_3 max1_4 max1_5 max1_6 max1_7 max1_8 max1_9 max1_10 max1_11 max1_12 min0_1 min0_2 min0_3 min0_4 min0_5 min0_6 min0_7 min0_8 min0_9 min0_10 min0_11 min0_12 min1_1 min1_2 min1_3 min1_4 min1_5 min1_6 min1_7 mi

> n1_8 min1_9 min1_10 min1_11 min1_12, by(district)

	district=1		district=2		
	Mean or N	SD or (%)	Mean or N	SD or (%)	Std Diff
expend0	837.1	338.5	1192	622.57	-0.70803
expendbase	837.1	338.09	1192	622.34	-0.70883
expend	948.3	384.31	1341	704.77	-0.69123
area	162.3	63.871	175.1	195.61	-0.08781
pool	.1089	.31164	.114	.31787	-0.01621
aircon	.2298	.42088	.2656	.44172	-0.08294
rc aircon	.1443	.35149	.1794	.38377	-0.09558
cheat	.1612	.36789	.1567	.3636	0.01232
dish	.2843	.45124	.3394	.47355	-0.11903
mwave	.9045	.29397	.9058	.2922	-0.00422
dryer	.7907	.40698	.8042	.39686	-0.03367
ftype	3.623	1.1748	3.606	1.1889	0.01447
people	4.471	2.3538	4.48	2.3762	-0.00390
inc	4.822	2.5656	5.895	2.589	-0.41651
hinc	.4059	.49125	.5708	.49502	-0.33431
audit	.007072	.083828	.0105	.10193	-0.03671
max0 1	26.63	1.1995	25.81	6.0711	0.18826
max0 2	26.63	.77388	26.36	5.5065	0.06933
max0 3	25.13	1.1995	24.85	5.6539	0.06820

max0 4 |	22.3	.85127 |	20.96	5.208 |	0.35842
max0 5 |	18.05	.11608 |	15.72	5.7554 |	0.57105
max0 6 |	13.35	.30955 |	11.48	4.8873 |	0.54047
max0 7 |	13.62	.38694 |	11.44	5.5368 |	0.55505
max0 8 |	15.07	.27086 |	12.5	6.249 |	0.58245
max0 9 |	16.04	.34825 |	14.16	5.3814 |	0.49303
max0 10 |	18.63	.27086 |	17.15	6.472 |	0.32166
max0 11 |	21.16	.077388 |	21.4	5.7956 |	-0.05679
max0 12 |	26.12	.38694 |	25.38	5.9789 |	0.17317
max1 1 |	27.67	.15478 |	27.47	5.9357 |	0.04846
max1 2 |	26.43	.58041 |	25.54	5.9095 |	0.21156
max1 3 |	27.38	.88996 |	25.82	5.9211 |	0.36717
max1 4 |	20.33	.58041 |	19.1	5.3317 |	0.32381
max1 5 |	16.34	.54172 |	15.05	5.0614 |	0.35966
max1 6 |	13.97	.69649 |	12.56	4.8306 |	0.40819
max1 7 |	13.98	.038694 |	11.77	6.4134 |	0.48671
max1 8 |	14.11	.19347 |	11.79	6.3057 |	0.51939
max1 9 |	17.77	.15478 |	15.52	6.7939 |	0.46852
max1 10 |	21.49	.23216 |	20.03	5.928 |	0.34806
max1 11 |	26.31	.61911 |	24.46	6.0444 |	0.43073
max1 12 |	25.42	.38694 |	24.7	5.7521 |	0.17541
min0 1 |	15.11	.61911 |	14.28	3.3294 |	0.34586
min0 2 |	14.64	.54172 |	13.55	2.9038 |	0.52142
min0 3 |	14.21	.19347 |	13.82	2.7943 |	0.19396
min0 4 |	11.85	.11608 |	8.918	2.4979 |	1.65566
min0 5 |	9.608	.19347 |	6.675	3.5485 |	1.16741
min0 6 |	6.408	.19347 |	3.823	2.9809 |	1.22422
min0 7 |	6.635	.34825 |	3.941	3.4539 |	1.09761
min0 8 |	6.308	.19347 |	3.048	3.4356 |	1.33982

```
    min0 9 |    8.035      .34825 |    5.652      3.4485 |   0.97256
   min0 10 |    7.772      .27086 |    5.453      3.8494 |   0.84975
   min0 11 |    9.563     .077388 |    8.748      2.7233 |   0.42323
   min0 12 |    12.66     .077388 |    12.47      2.5675 |   0.10488
    min1 1 |    14.37      .69649 |    14.04      2.2155 |   0.20105
    min1 2 |    12.79      .23216 |    12.51      2.6811 |   0.14582
    min1 3 |    14.38     .038694 |    13.42      3.0521 |   0.44665
    min1 4 |    10.74      .34825 |    8.412      3.1103 |   1.04974
    min1 5 |    7.555      .11608 |    5.109      2.6863 |   1.28662
    min1 6 |    5.418     .038694 |    2.102      2.4662 |   1.90118
    min1 7 |    4.818     .038694 |    1.125      2.9581 |   1.76551
    min1 8 |    5.327      .15478 |    2.383      3.6104 |   1.15210
    min1 9 |    6.872      .27086 |    3.961      3.4515 |   1.18907
   min1 10 |      8.9           0 |    7.174      2.6337 |   0.92661
   min1 11 |    12.62      .38694 |    11.22      3.3756 |   0.58223
   min1 12 |    13.25      .11608 |    13.07      2.7694 |   0.08892
-----------------------------------------------------------------------------

.
. * Provides information of what is the main driver of changes in energy expenditure
. bysort district: summarize change_expend

---------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------
-> district = 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
change_exp~d |      1,414    111.2051    48.15262   19.02051   508.5732
```

```
-------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------

-> district = 2

   Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
change_exp~d |     4,096    148.7789     86.4565  -24.45995   773.5679

. * These two graphs indicate that District 2 shifts their change in energy expenditure to the left

. twoway kdensity change_expend if district == 1 & treat == 1 & change_expend < 400 || kdensity change_expend if district == 1 & treat == 0 & change_expend < 400, legend(order(1 "Treated" 2 "Control")) xtitle("Chan
> ge in Energy Expenditure For District 1") ytitle("Density") graphregion(color(white)) bgcolor(white)

.

. twoway kdensity change_expend if district == 2 & treat == 1 & change_expend < 400|| kdensity change_expend if district == 2 & treat == 0 & change_expend < 400, legend(order(1 "Treated" 2 "Control")) xtitle("Chang
> e in Energy Expenditure For District 2") ytitle("Density") graphregion(color(white)) bgcolor(white)

.

. * this is how the outlier amounts were determined

. summarize area if treat == 1, detail

                            area
-------------------------------------------------------------
      Percentiles      Smallest
 1%           78             70
```

```
 5%        93          70
10%       105          70     Obs            3,292
25%       132          71     Sum of Wgt.    3,292

50%     158.9                 Mean        174.3164
                  Largest     Std. Dev.   207.2137
75%     185.1        3112
90%       228        4259     Variance    42937.53
95%       278        6791     Skewness    27.40639
99%       449        7629     Kurtosis    883.8856

. summarize area if treat == 0, detail

                     area
-------------------------------------------------------------
    Percentiles    Smallest
 1%        77          70
 5%        92          70
10%       104          70     Obs            2,218
25%       131          70     Sum of Wgt.    2,218

50%       160                 Mean        168.0515
                  Largest     Std. Dev.   97.94184
75%     185.7         766
90%       222         852     Variance    9592.605
95%       277        1041     Skewness    19.88135
99%       407        3558     Kurtosis    654.6534

.
```

```
. * This drops the three largest observations of area and finds checks that it is solely the
. * three outliers that impact on such a large standard deviation
. summarize area if treat == 1 & area < 4000, detail

                        area
-------------------------------------------------------------
      Percentiles      Smallest
 1%          78             70
 5%          93             70
10%         105             70       Obs              3,289
25%         132             71       Sum of Wgt.      3,289

50%       158.8                      Mean          168.7962
                        Largest      Std. Dev.     87.49305
75%         185           1022
90%         227           1129       Variance      7655.034
95%         277           1143       Skewness      13.92937
99%         427           3112       Kurtosis      407.7202

. **<><><><><><><><><>**
.
.
. ** This is the beginning of the econometric modelling**
. **<><><><><><><><><>**
. * These are the base controls. It includes all variables and interactions between income and household observable characteristics
. global controls c.max0_1 c.max0_2 c.max0_3 c.max0_4 c.max0_5 c.max0_6 c.max0_7 c.max0_8 c.max0_9 c.max0_10 c.max0_11 c.max0_12 c.max1_1 c.max1_2 c.max1_3 c.max1_4 c.max1_5 c.max1_6 c.max1_7 c.max1_8 c.max1_9 c.ma
> x1_10 c.max1_11 c.max1_12 i.inc##(i.pool i.aircon i.rc_aircon i.cheat i.dish i.mwave i.dryer i.ftype i.people i.inc c.logarea)
```

```
.
. global controls1 $controls i.district i.audit

. global controls2 $controls audit#$controls

. global controls3 $controls district#$controls

.
. * This checks the average treatment effect across a variety of specifications
. * This includes no controls OLS, all controls OLS, and Double Selection.
. * <><><> *
. regress change_expend treat, vce(r)

Linear regression                    Number of obs     =      5,510
                                     F(1, 5508)        =     122.38
                                     Prob > F          =     0.0000
                                     R-squared         =     0.0220
                                     Root MSE          =     79.247

------------------------------------------------------------------------------
             |               Robust
change_exp~d |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       treat |  -24.23723   2.190908   -11.06   0.000    -28.53228   -19.94219
       _cons |   153.6173   1.715538    89.54   0.000     150.2542    156.9805
------------------------------------------------------------------------------

. estimates store m1
```

.

```
. regress change_expend treat $controls1, vce(r)
note: max0_6 omitted because of collinearity
note: max0_7 omitted because of collinearity
note: max0_8 omitted because of collinearity
note: max0_9 omitted because of collinearity
note: max0_10 omitted because of collinearity
note: max0_11 omitted because of collinearity
note: max0_12 omitted because of collinearity
note: max1_1 omitted because of collinearity
note: max1_2 omitted because of collinearity
note: max1_3 omitted because of collinearity
note: max1_4 omitted because of collinearity
note: max1_5 omitted because of collinearity
note: max1_6 omitted because of collinearity
note: max1_7 omitted because of collinearity
note: max1_8 omitted because of collinearity
note: max1_9 omitted because of collinearity
note: max1_10 omitted because of collinearity
note: max1_11 omitted because of collinearity
note: max1_12 omitted because of collinearity
note: 2.district omitted because of collinearity

Linear regression                 Number of obs   =    5,510
                          F(216, 5293)    =    25.67
                          Prob > F        =   0.0000
                          R-squared       =   0.5298
                          Root MSE        =   56.051
```

```
------------------------------------------------------------------------------
             |               Robust
change_expend |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       treat |  -26.38243   1.595084   -16.54   0.000    -29.50945   -23.25541
      max0_1 |  -65.70702   3.655679   -17.97   0.000    -72.87366   -58.54038
      max0_2 |   .9365709   1.581792     0.59   0.554    -2.164393    4.037535
      max0_3 |   69.51058   4.721722    14.72   0.000     60.25406     78.7671
      max0_4 |  -17.37974    4.41093    -3.94   0.000    -26.02698   -8.732498
      max0_5 |   15.78758   2.366942     6.67   0.000      11.1474    20.42777
      max0_6 |          0  (omitted)
      max0_7 |          0  (omitted)
      max0_8 |          0  (omitted)
      max0_9 |          0  (omitted)
     max0_10 |          0  (omitted)
     max0_11 |          0  (omitted)
     max0_12 |          0  (omitted)
      max1_1 |          0  (omitted)
      max1_2 |          0  (omitted)
      max1_3 |          0  (omitted)
      max1_4 |          0  (omitted)
      max1_5 |          0  (omitted)
      max1_6 |          0  (omitted)
      max1_7 |          0  (omitted)
      max1_8 |          0  (omitted)
      max1_9 |          0  (omitted)
     max1_10 |          0  (omitted)
     max1_11 |          0  (omitted)
```

max1_12	0 (omitted)					
inc						
2	48.59847	48.04855	1.01	0.312	-45.59649	142.7934
3	31.93646	51.0007	0.63	0.531	-68.04594	131.9189
4	74.82485	53.72291	1.39	0.164	-30.49421	180.1439
5	-122.9509	68.99334	-1.78	0.075	-258.2063	12.30447
6	-157.6457	62.64243	-2.52	0.012	-280.4507	-34.84072
7	-103.0148	74.96129	-1.37	0.169	-249.9698	43.94024
8	-145.0132	80.22579	-1.81	0.071	-302.2889	12.26237
9	29.04063	102.4158	0.28	0.777	-171.7365	229.8177
10	-142.0126	72.87215	-1.95	0.051	-284.8721	.8468272
1.pool	2.802571	15.39397	0.18	0.856	-27.37596	32.98111
1.aircon	20.89457	12.75946	1.64	0.102	-4.119233	45.90838
1.rc_aircon	14.02624	16.33993	0.86	0.391	-18.00676	46.05924
1.cheat	-4.714729	5.084791	-0.93	0.354	-14.68302	5.253558
1.dish	-18.32107	13.87591	-1.32	0.187	-45.52358	8.881447
1.mwave	4.220975	7.70098	0.55	0.584	-10.87612	19.31807
1.dryer	4.124276	7.564476	0.55	0.586	-10.70522	18.95377
ftype						
2	28.35663	13.58786	2.09	0.037	1.71882	54.99444
3	-3.231029	16.18673	-0.20	0.842	-34.96369	28.50163
4	-1.804096	15.3594	-0.12	0.907	-31.91484	28.30665
5	3.232157	11.82098	0.27	0.785	-19.94185	26.40616
people						
2	-1.930794	13.8041	-0.14	0.889	-28.99251	25.13093

3	-2.754148	10.67677	-0.26	0.796	-23.68501	18.17672
4	-8.274451	12.09857	-0.68	0.494	-31.99263	15.44373
5	-1.635484	12.9183	-0.13	0.899	-26.96068	23.68971
6	-13.72631	14.39433	-0.95	0.340	-41.94514	14.49252
7	6.922969	21.21549	0.33	0.744	-34.66814	48.51408
8	-6.093037	15.80395	-0.39	0.700	-37.0753	24.88922
9	1.708107	15.89576	0.11	0.914	-29.45413	32.87034
logarea	30.8771	8.221158	3.76	0.000	14.76024	46.99396
inc#pool						
2 1	1.365851	18.45169	0.07	0.941	-34.80707	37.53877
3 1	8.019788	19.13629	0.42	0.675	-29.49524	45.53481
4 1	-23.74099	19.44258	-1.22	0.222	-61.85647	14.37449
5 1	-14.4172	21.83854	-0.66	0.509	-57.22974	28.39534
6 1	-12.0559	24.83309	-0.49	0.627	-60.739	36.62719
7 1	-19.70183	23.96026	-0.82	0.411	-66.67382	27.27015
8 1	2.180591	23.69865	0.09	0.927	-44.27852	48.63971
9 1	-7.294849	29.69058	-0.25	0.806	-65.50062	50.91093
10 1	-43.58548	19.5469	-2.23	0.026	-81.90547	-5.2655
inc#aircon						
2 1	-41.1989	18.52291	-2.22	0.026	-77.51145	-4.886352
3 1	-22.0968	16.90243	-1.31	0.191	-55.23253	11.03893
4 1	13.14036	18.3849	0.71	0.475	-22.90162	49.18233
5 1	20.99186	20.51854	1.02	0.306	-19.23294	61.21666
6 1	3.664494	25.37205	0.14	0.885	-46.07519	53.40418
7 1	-18.24707	22.78203	-0.80	0.423	-62.90923	26.41509
8 1	22.53032	20.45118	1.10	0.271	-17.56242	62.62306

9 1	-8.622719	26.70742	-0.32	0.747	-60.98027	43.73483
10 1	4.278202	18.59214	0.23	0.818	-32.17005	40.72646
inc#rc_aircon						
2 1	-1.133645	16.9786	-0.07	0.947	-34.4187	32.15141
3 1	-2.331654	18.4638	-0.13	0.900	-38.52831	33.865
4 1	28.61274	23.27023	1.23	0.219	-17.00651	74.23199
5 1	87.83379	21.14046	4.15	0.000	46.38978	129.2778
6 1	77.88166	22.31884	3.49	0.000	34.12754	121.6358
7 1	74.40166	21.28131	3.50	0.000	32.68152	116.1218
8 1	72.14435	19.47249	3.70	0.000	33.97023	110.3185
9 1	67.80707	21.31778	3.18	0.001	26.01543	109.5987
10 1	87.03878	19.29494	4.51	0.000	49.21275	124.8648
inc#cheat						
2 1	14.97129	7.305598	2.05	0.040	.6493023	29.29327
3 1	-2.176494	8.078836	-0.27	0.788	-18.01434	13.66135
4 1	13.49207	7.066339	1.91	0.056	-.3608639	27.34501
5 1	17.2603	8.856255	1.95	0.051	-.1016095	34.62221
6 1	24.27676	7.718726	3.15	0.002	9.144878	39.40865
7 1	15.27634	9.358365	1.63	0.103	-3.069915	33.62259
8 1	17.15416	9.624529	1.78	0.075	-1.713883	36.02221
9 1	26.76976	11.16597	2.40	0.017	4.879843	48.65967
10 1	22.68671	9.796134	2.32	0.021	3.48225	41.89117
inc#dish						
2 1	44.9469	19.38782	2.32	0.020	6.938784	82.95502
3 1	20.54292	17.89312	1.15	0.251	-14.53497	55.62082
4 1	-1.228346	18.38931	-0.07	0.947	-37.27898	34.82229

5 1	-1.028207	21.34447	-0.05	0.962	-42.87217	40.81575
6 1	16.01888	25.36786	0.63	0.528	-33.71259	65.75034
7 1	37.77669	23.62926	1.60	0.110	-8.546401	84.09979
8 1	-6.299546	21.08133	-0.30	0.765	-47.62764	35.02855
9 1	16.30011	27.87742	0.58	0.559	-38.35112	70.95135
10 1	10.43217	21.59686	0.48	0.629	-31.90658	52.77093
inc#mwave						
2 1	.1383898	9.603189	0.01	0.989	-18.68782	18.9646
3 1	-1.11824	11.19678	-0.10	0.920	-23.06855	20.83207
4 1	7.378586	9.4396	0.78	0.434	-11.12692	25.88409
5 1	-8.505036	12.35479	-0.69	0.491	-32.72552	15.71545
6 1	-35.26025	13.98438	-2.52	0.012	-62.67539	-7.845109
7 1	-6.003154	13.16143	-0.46	0.648	-31.80497	19.79867
8 1	7.776008	13.55386	0.57	0.566	-18.79514	34.34716
9 1	-14.19898	17.05932	-0.83	0.405	-47.64228	19.24432
10 1	15.82497	14.1878	1.12	0.265	-11.98896	43.6389
inc#dryer						
2 1	-9.402406	9.110651	-1.03	0.302	-27.26304	8.458227
3 1	-4.888745	9.185357	-0.53	0.595	-22.89583	13.11834
4 1	-10.55363	10.18081	-1.04	0.300	-30.51221	9.404944
5 1	-30.53201	10.57547	-2.89	0.004	-51.26429	-9.799728
6 1	-15.43131	10.95896	-1.41	0.159	-36.9154	6.052769
7 1	-15.88561	12.1442	-1.31	0.191	-39.69325	7.92203
8 1	-40.67251	12.1908	-3.34	0.001	-64.57151	-16.77351
9 1	-27.49843	14.08728	-1.95	0.051	-55.11531	.1184384
10 1	-27.89125	11.62511	-2.40	0.016	-50.68125	-5.101248

inc#ftype							
2 2	-25.48841	17.02217	-1.50	0.134	-58.85889	7.882065
2 3	4.24772	19.82514	0.21	0.830	-34.61774	43.11318
2 4	-9.708335	18.53884	-0.52	0.601	-46.0521	26.63543
2 5	-.5734174	14.77697	-0.04	0.969	-29.54237	28.39553
3 2	-18.44498	18.46102	-1.00	0.318	-54.63619	17.74623
3 3	-19.44488	21.32461	-0.91	0.362	-61.24992	22.36015
3 4	3.116461	19.58255	0.16	0.874	-35.27341	41.50633
3 5	-15.62901	15.84048	-0.99	0.324	-46.68288	15.42487
4 2	-41.38949	19.16978	-2.16	0.031	-78.97017	-3.808811
4 3	-8.406765	21.64492	-0.39	0.698	-50.83972	34.02619
4 4	-23.37078	20.78153	-1.12	0.261	-64.11114	17.36958
4 5	-26.96764	16.66627	-1.62	0.106	-59.6404	5.70512
5 2	-42.65296	20.74869	-2.06	0.040	-83.32895	-1.976968
5 3	21.62027	23.26513	0.93	0.353	-23.98897	67.22951
5 4	-18.95839	21.4903	-0.88	0.378	-61.08825	23.17146
5 5	-21.22262	17.29809	-1.23	0.220	-55.13401	12.68877
6 2	34.53252	27.92469	1.24	0.216	-20.21138	89.27642
6 3	29.25214	31.40095	0.93	0.352	-32.30667	90.81095
6 4	53.34852	28.20384	1.89	0.059	-1.942637	108.6397
6 5	32.1314	23.6217	1.36	0.174	-14.17688	78.43967
7 2	-82.90224	22.51291	-3.68	0.000	-127.0368	-38.76765
7 3	-74.92182	26.1451	-2.87	0.004	-126.177	-23.66665
7 4	-48.38067	23.06332	-2.10	0.036	-93.59428	-3.167059
7 5	-53.31595	18.55569	-2.87	0.004	-89.69275	-16.93916
8 2	-48.99388	25.51769	-1.92	0.055	-99.01908	1.031322
8 3	8.237985	28.063	0.29	0.769	-46.77707	63.25304
8 4	-2.73122	25.45032	-0.11	0.915	-52.62435	47.16191
8 5	8.211896	21.46372	0.38	0.702	-33.86583	50.28963

9 2	-35.32394	27.79485	-1.27	0.204	-89.81331	19.16543
9 3	10.39321	31.30824	0.33	0.740	-50.98386	71.77027
9 4	-11.25939	28.30918	-0.40	0.691	-66.75704	44.23827
9 5	-9.958012	23.20834	-0.43	0.668	-55.45593	35.53991
10 2	20.35107	27.93182	0.73	0.466	-34.40682	75.10896
10 3	46.22553	27.47763	1.68	0.093	-7.641963	100.093
10 4	18.40456	25.22941	0.73	0.466	-31.05549	67.8646
10 5	15.56683	22.64711	0.69	0.492	-28.83084	59.9645
inc#people						
2 2	12.71756	16.4548	0.77	0.440	-19.54064	44.97576
2 3	6.783625	13.7214	0.49	0.621	-20.11597	33.68322
2 4	24.31924	15.15468	1.60	0.109	-5.390176	54.02865
2 5	21.05503	16.28111	1.29	0.196	-10.86265	52.97272
2 6	28.65425	17.36988	1.65	0.099	-5.397886	62.70638
2 7	5.668003	24.06904	0.24	0.814	-41.51724	52.85324
2 8	17.25268	19.49166	0.89	0.376	-20.95901	55.46437
2 9	.6222333	20.14729	0.03	0.975	-38.87476	40.11923
3 2	1.295733	18.4497	0.07	0.944	-34.87329	37.46476
3 3	-22.19096	14.97551	-1.48	0.138	-51.54914	7.167214
3 4	-7.290927	16.43861	-0.44	0.657	-39.51738	24.93553
3 5	-15.00092	16.91849	-0.89	0.375	-48.16814	18.1663
3 6	3.63636	18.13571	0.20	0.841	-31.9171	39.18982
3 7	-28.27387	24.66978	-1.15	0.252	-76.63682	20.08908
3 8	-11.19135	19.91423	-0.56	0.574	-50.23145	27.84876
3 9	-32.92238	20.46575	-1.61	0.108	-73.04368	7.19893
4 2	-5.360028	17.35727	-0.31	0.757	-39.38743	28.66737
4 3	4.748138	13.80808	0.34	0.731	-22.32139	31.81766
4 4	12.78119	15.06989	0.85	0.396	-16.762	42.32439

4 5	8.88612	16.06791	0.55	0.580	-22.61362	40.38586
4 6	20.60478	18.11772	1.14	0.255	-14.91342	56.12298
4 7	2.664328	25.70777	0.10	0.917	-47.7335	53.06216
4 8	13.40461	20.24647	0.66	0.508	-26.28681	53.09603
4 9	12.12564	21.81671	0.56	0.578	-30.6441	54.89539
5 2	-14.44377	18.9832	-0.76	0.447	-51.65868	22.77114
5 3	13.65313	15.92135	0.86	0.391	-17.55928	44.86554
5 4	15.52858	17.52901	0.89	0.376	-18.8355	49.89267
5 5	11.01719	17.98741	0.61	0.540	-24.24555	46.27993
5 6	31.37826	20.31643	1.54	0.123	-8.450323	71.20684
5 7	1.252504	25.66573	0.05	0.961	-49.06291	51.56792
5 8	19.16339	21.86231	0.88	0.381	-23.69575	62.02252
5 9	27.0032	23.37196	1.16	0.248	-18.81549	72.82188
6 2	-18.39915	29.17423	-0.63	0.528	-75.59267	38.79437
6 3	-43.29996	25.18677	-1.72	0.086	-92.67642	6.076492
6 4	-38.65128	25.52268	-1.51	0.130	-88.68626	11.3837
6 5	-20.58462	26.41707	-0.78	0.436	-72.37295	31.20372
6 6	-16.16036	27.61826	-0.59	0.558	-70.30354	37.98281
6 7	-56.71205	32.25169	-1.76	0.079	-119.9387	6.51457
6 8	-23.00591	29.36114	-0.78	0.433	-80.56584	34.55403
6 9	-29.33928	30.33116	-0.97	0.333	-88.80086	30.12231
7 2	32.06038	21.54672	1.49	0.137	-10.18007	74.30083
7 3	7.550499	15.62843	0.48	0.629	-23.08768	38.18867
7 4	22.12573	17.30619	1.28	0.201	-11.80154	56.053
7 5	20.95456	18.78467	1.12	0.265	-15.87114	57.78026
7 6	38.94454	21.52592	1.81	0.070	-3.255142	81.14423
7 7	10.04841	27.36192	0.37	0.713	-43.59223	63.68906
7 8	2.839722	23.25501	0.12	0.903	-42.74969	48.42913
7 9	32.77805	24.45112	1.34	0.180	-15.15622	80.71232

8 2	-9.800696	23.7408	-0.41	0.680	-56.34245	36.74106
8 3	18.76872	22.239	0.84	0.399	-24.82889	62.36633
8 4	8.142049	21.50001	0.38	0.705	-34.00684	50.29093
8 5	-.1027463	22.87134	-0.00	0.996	-44.94001	44.73452
8 6	20.45903	23.79858	0.86	0.390	-26.19599	67.11406
8 7	-6.356739	31.1448	-0.20	0.838	-67.41339	54.69991
8 8	5.227823	27.06472	0.19	0.847	-47.83019	58.28583
8 9	18.77868	31.5388	0.60	0.552	-43.05037	80.60772
9 2	-11.98868	25.93501	-0.46	0.644	-62.83199	38.85462
9 3	14.57565	20.08549	0.73	0.468	-24.80018	53.95149
9 4	23.38127	21.88158	1.07	0.285	-19.51564	66.27818
9 5	17.8768	24.41041	0.73	0.464	-29.97766	65.73127
9 6	23.88362	24.83131	0.96	0.336	-24.79599	72.56322
9 7	2.800352	30.09796	0.09	0.926	-56.20406	61.80476
9 8	16.24775	27.22342	0.60	0.551	-37.12138	69.61687
9 9	3.551954	27.49256	0.13	0.897	-50.34479	57.4487
10 2	-56.12704	22.04148	-2.55	0.011	-99.33744	-12.91664
10 3	-29.07587	19.89727	-1.46	0.144	-68.08272	9.930975
10 4	-20.71324	20.59247	-1.01	0.315	-61.08297	19.65649
10 5	-33.82017	21.40309	-1.58	0.114	-75.77905	8.138707
10 6	-16.42924	22.85796	-0.72	0.472	-61.24027	28.38179
10 7	-26.89157	28.50545	-0.94	0.346	-82.774	28.99087
10 8	-26.02878	24.34048	-1.07	0.285	-73.74615	21.6886
10 9	-33.5166	26.24221	-1.28	0.202	-84.96216	17.92896
inc#c.logarea						
2	-11.17491	9.669432	-1.16	0.248	-30.13098	7.781164
3	-3.17745	10.16509	-0.31	0.755	-23.10521	16.75031
4	-13.53778	10.77592	-1.26	0.209	-34.66302	7.587458

```
         5  |  35.17111  13.72317    2.56  0.010     8.26804    62.07419
         6  |  40.14237  12.84833    3.12  0.002    14.95435     65.3304
         7  |  34.60963  14.97745    2.31  0.021    5.247645    63.97161
         8  |  42.69309   16.2505    2.63  0.009    10.83541    74.55076
         9  |   10.0317  21.11534    0.48  0.635   -31.36308    51.42647
        10  |   40.2606  14.65963    2.75  0.006    11.52169    68.99952
            |
 2.district |         0  (omitted)
    1.audit |  -3.856253   8.56681   -0.45  0.653   -20.65073   12.93823
      _cons |   21.64483  42.15235    0.51  0.608   -60.99116   104.2808
------------------------------------------------------------------------------

. estimates store m2

.
. dsregress change_expend treat, rseed(10101) controls($controls1)

Estimating lasso for change_expend using plugin
Estimating lasso for treat using plugin

Double-selection linear model         Number of obs               =      5,510
                                      Number of controls          =        357
                                      Number of selected controls =         21
                                      Wald chi2(1)                =     239.25
                                      Prob > chi2                 =     0.0000

------------------------------------------------------------------------------
             |               Robust
change_exp~d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
       treat |  -25.54659   1.651618   -15.47   0.000    -28.7837   -22.30948
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
      of interest jointly equal to zero. Lassos select controls for model
      estimation. Type lassoinfo to see number of selected variables in each
      lasso.

. estimates store m3

. lassocoef (., for(change_expend)), display(coef, postselection)


----------------------------
             | change_expend
--------------+-------------
      max0_1 |    -1.672257
             |
         inc |
          8  |     12.18631
          9  |     12.63882
             |
   rc_aircon |
          0  |    -66.89824
          1  |            e
             |
      0.dish |    -5.663445
     logarea |     53.05631
             |
    inc#pool |
```

```
         8 0  |     21.61407
         9 0  |     8.750091
        10 0  |     31.68189
              |
   inc#aircon |
         1 0  |    -5.768775
         4 0  |    -3.505105
              |
inc#rc_aircon |
         2 1  |    -55.73196
         9 0  |      4.69098
              |
    inc#dryer |
         3 1  |      -13.482
              |
    inc#ftype |
         1 4  |    -13.39473
              |
inc#c.logarea |
           1  |     -4.48819
           2  |    -6.518431
           3  |    -5.095943
           4  |    -7.111369
              |
   1.district |    -20.58511
        _cons |    -16.29568
-----------------------------
Legend:
  b - base level
```

```
  e - empty cell
  o - omitted

. lassocoef (., for(treat)), display(coef, postselection)

-----------------------
             |    treat
-------------+----------
       _cons |  .5974592
-----------------------
Legend:
  b - base level
  e - empty cell
  o - omitted

.
. esttab m1 m2 m3, label title("Table 2 Regression Results (Average Treatment Effects)")
nonumbers mtitles("OLS (No controls)" "OLS (All controls)" "Double Selection")
modelwidth(25) keep(treat) se replace
```

Table 2 Regression Results (Average Treatment Effects)

	OLS (No controls)	OLS (All controls)	Double Selection
treat	-24.24***	-26.38***	-25.55***
	(2.191)	(1.595)	(1.652)
Observations	5510	5510	5510

Standard errors in parentheses

```
* p<0.05, ** p<0.01, *** p<0.001

. * <><><> *
.
. * This checks the average treatment effect across districts across a variety of specifications
. * This includes no controls SIM, Double Selection SIM, and Double Selection FMM.
. * <><><> *
. regress change_expend treat##district, vce(r)

Linear regression                    Number of obs     =      5,510
                                     F(3, 5506)        =     186.54
                                     Prob > F          =     0.0000
                                     R-squared         =     0.0710
                                     Root MSE          =     77.251

------------------------------------------------------------------------------
             |               Robust
change_expend |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
      1.treat |  -.4118897   2.607002    -0.16   0.874    -5.522643    4.698863
   2.district |   56.57701   2.893186    19.56   0.000     50.90522     62.2488
             |
treat#district |
         1 2  |  -31.86399   3.755973    -8.48   0.000    -39.22718    -24.5008
             |
       _cons |   111.4524   2.008685    55.49   0.000     107.5146    115.3902
------------------------------------------------------------------------------

. estimates store m4
```

```
.

. dsregress change_expend treat##district, rseed(10101) controls($controls i.audit)

Estimating lasso for change_expend using plugin
Estimating lasso for 1bn.treat using plugin
Estimating lasso for 2bn.district using plugin
Estimating lasso for 1bn.treat#2bn.district using plugin

Double-selection linear model        Number of obs            =     5,510
                                     Number of controls       =       355
                                     Number of selected controls =     27
                                     Wald chi2(3)             =    467.56
                                     Prob > chi2              =    0.0000

------------------------------------------------------------------------------
               |               Robust
 change_expend |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
       1.treat |   -.250257   2.244693    -0.11   0.911    -4.649774     4.14926
    2.district |    50.3779   2.939137    17.14   0.000      44.6173    56.13851
               |
treat#district |
           1 2 |  -34.38154   2.990129   -11.50   0.000    -40.24208     -28.521
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
      of interest jointly equal to zero. Lassos select controls for model
      estimation. Type lassoinfo to see number of selected variables in each
      lasso.
```

```
. estimates store m5

. lassocoef (., for(change_expend)), display(coef, postselection)

----------------------------
           | change_expend
--------------+--------------
     max0_1 |     -1.676889
            |
        inc |
         8  |     9.348527
         9  |     11.54808
            |
   rc_aircon |
         0  |    -70.71265
         1  |           e
            |
     0.dish |     -6.527941
    logarea |     52.78746
            |
    inc#pool |
       8 0  |     27.05677
       9 0  |     8.996538
      10 0  |     33.70834
            |
  inc#aircon |
       4 0  |    -2.382813
            |
```

```
inc#rc_aircon |
        2 1  |     -60.97055
        3 1  |     -54.90531
        9 0  |      9.182898
             |
   inc#dryer |
        3 1  |     -7.358099
             |
   inc#ftype |
        1 4  |     -13.24599
             |
inc#c.logarea |
          1  |    -5.842103
          2  |    -6.971584
          3  |    -5.352049
          4  |    -7.685223
             |
       _cons |     -16.15368
----------------------------
Legend:
  b - base level
  e - empty cell
  o - omitted

.
. dsregress change_expend treat##district, rseed(10101) controls($controls3)

Estimating lasso for change_expend using plugin
Estimating lasso for 1bn.treat using plugin
```

```
Estimating lasso for 2bn.district using plugin
Estimating lasso for 1bn.treat#2bn.district using plugin

Double-selection linear model        Number of obs            =     5,510
                        Number of controls        =       355
                        Number of selected controls =        25
                        Wald chi2(3)            =    462.54
                        Prob > chi2             =    0.0000

------------------------------------------------------------------------------
             |             Robust
 change_expend |     Coef.   Std. Err.     z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
     1.treat |  -.2302492   2.24887   -0.10   0.918   -4.637954   4.177455
   2.district |   47.61929   2.748638   17.32   0.000    42.23205   53.00652
             |
treat#district |
        1 2  |  -34.29187   2.997583   -11.44   0.000   -40.16703   -28.41672
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
    of interest jointly equal to zero. Lassos select controls for model
    estimation. Type lassoinfo to see number of selected variables in each
    lasso.

. estimates store m6

.

. esttab m4 m5 m6, label title("Table 3 Regression Results (District Average Treatment
Effects)") nonumbers mtitles("SIM OLS (No controls)" "SIM Double Selection" "FMM Double
Selection") modelwidth(25) se replace d
```

> rop(0.treat 1.district 0.treat#1.district 0.treat#2.district 1.treat#1.district)

Table 3 Regression Results (District Average Treatment Effects)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-0.412	-0.250	-0.230
	(2.607)	(2.245)	(2.249)
district=2	56.58***	50.38***	47.62***
	(2.893)	(2.939)	(2.749)
treat=1 # district=2	-31.86***	-34.38***	-34.29***
	(3.756)	(2.990)	(2.998)
Constant	111.5***		
	(2.009)		
Observations	5510	5510	5510

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

. * <><><> *

.

. * This checks the average treatment effect across already-audited across a variety of specifications

. * This includes no controls SIM, Double Selection SIM, and Double Selection FMM.

. * <><><> *

. regress change_expend treat##audit, vce(r)

```
Linear regression                     Number of obs     =      5,510
                                      F(3, 5506)        =      41.90
                                      Prob > F          =     0.0000
                                      R-squared         =     0.0223
                                      Root MSE          =     79.249

------------------------------------------------------------------------------
             |               Robust
change_exp~d |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.treat |  -24.11904   2.204391   -10.94   0.000    -28.44052   -19.79757
     1.audit |   21.47228   14.91397     1.44   0.150    -7.764986    50.70955
             |
 treat#audit |
        1 1  |  -13.94752    18.9647    -0.74   0.462    -51.12582    23.23077
             |
       _cons |   153.4237    1.72566    88.91   0.000     150.0407    156.8067
------------------------------------------------------------------------------

. estimates store m7

.
. dsregress change_expend treat##audit, rseed(10101) controls($controls i.district)

Estimating lasso for change_expend using plugin
Estimating lasso for 1bn.treat using plugin
Estimating lasso for 1bn.audit using plugin
Estimating lasso for 1bn.treat#1bn.audit using plugin
```

```
Double-selection linear model        Number of obs             =     5,510
                             Number of controls        =       355
                             Number of selected controls =       185
                             Wald chi2(3)              =    248.66
                             Prob > chi2               =    0.0000

------------------------------------------------------------------------------
             |               Robust
change_exp~d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.treat |  -26.18627   1.668174   -15.70   0.000    -29.45583   -22.91671
     1.audit |  -5.801791   12.00394    -0.48   0.629    -29.32907    17.72549
             |
 treat#audit |
        1 1  |   1.487942   16.91244     0.09   0.930    -31.65984    34.63572
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
      of interest jointly equal to zero. Lassos select controls for model
      estimation. Type lassoinfo to see number of selected variables in each
      lasso.

. estimates store m8


.

. dsregress change_expend treat##audit, rseed(10101) controls($controls2)


Estimating lasso for change_expend using plugin
Estimating lasso for 1bn.treat using plugin
```

```
Estimating lasso for 1bn.audit using plugin

Estimating lasso for 1bn.treat#1bn.audit using plugin


Double-selection linear model        Number of obs            =     5,510

                             Number of controls         =       355

                             Number of selected controls =        33

                             Wald chi2(3)               =    234.68

                             Prob > chi2                =    0.0000


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

             |             Robust

change_exp~d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

-------------+----------------------------------------------------------------

     1.treat |  -25.62577   1.686207   -15.20   0.000    -28.93068   -22.32087

     1.audit |  -17.09873   45.53144    -0.38   0.707    -106.3387    72.14125

             |

 treat#audit |

        1 1  |   -5.91916   16.35719    -0.36   0.717    -37.97866    26.14034

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

Note: Chi-squared test is a Wald test of the coefficients of the variables

      of interest jointly equal to zero. Lassos select controls for model

      estimation. Type lassoinfo to see number of selected variables in each

      lasso.


. estimates store m9


.

. esttab m7 m8 m9, label title("Table 5 Regression Results (Average Treatment Effects On
Already Audited Households)") nonumbers mtitles("SIM OLS (No controls)" "SIM Double
Selection" "FMM Double Selection") modelw
```

> idth(25) se replace drop(0.treat 0.audit 0.treat#0.audit 0.treat#1.audit 1.treat#0.audit)

Table 5 Regression Results (Average Treatment Effects On Already Audited Households)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-24.12***	-26.19***	-25.63***
	(2.204)	(1.668)	(1.686)
audit=1	21.47	-5.802	-17.10
	(14.91)	(12.00)	(45.53)
treat=1 # audit=1	-13.95	1.488	-5.919
	(18.96)	(16.91)	(16.36)
Constant	153.4***		
	(1.726)		
Observations	5510	5510	5510

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

. * <><><> *

.

. * This checks whether weather has explanatory power for differing treatment effect across districts

. * <><><> *

. preserve

```
. keep if max0_1 == 27.2 | max0_1 == 26
(3,895 observations deleted)

. regress change_expend treat##district, vce(r)

Linear regression                  Number of obs    =     1,615
                                   F(3, 1611)       =     42.31
                                   Prob > F         =    0.0000
                                   R-squared        =    0.1027
                                   Root MSE         =     51.19

------------------------------------------------------------------------------
             |               Robust
change_expend |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
      1.treat |  -.8851492   2.623261    -0.34   0.736    -6.030512    4.260214
   2.district |    51.2096   5.338304     9.59   0.000     40.73885    61.68035
              |
treat#district |
         1 2  |  -27.19073   6.856348    -3.97   0.000    -40.63903   -13.74243
              |
        _cons |    106.978   2.028383    52.74   0.000     102.9995    110.9566
------------------------------------------------------------------------------

. estimates store m10

. dsregress change_expend treat##district, rseed(10101) controls($controls i.audit)

Estimating lasso for change_expend using plugin
```

```
note: max0_1 dropped because of collinearity with another variable
note: max0_2 dropped because of collinearity with another variable
note: max0_3 dropped because of collinearity with another variable
note: max0_4 dropped because of collinearity with another variable
note: max0_5 dropped because of collinearity with another variable
note: max0_6 dropped because of collinearity with another variable
note: max0_7 dropped because of collinearity with another variable
note: max0_8 dropped because of collinearity with another variable
note: max0_9 dropped because of collinearity with another variable
note: max0_10 dropped because of collinearity with another variable
note: max0_11 dropped because of collinearity with another variable
note: max0_12 dropped because of collinearity with another variable
note: max1_1 dropped because of collinearity with another variable
note: max1_2 dropped because of collinearity with another variable
note: max1_3 dropped because of collinearity with another variable
note: max1_4 dropped because of collinearity with another variable
note: max1_5 dropped because of collinearity with another variable
note: max1_6 dropped because of collinearity with another variable
note: max1_7 dropped because of collinearity with another variable
note: max1_8 dropped because of collinearity with another variable
note: max1_10 dropped because of collinearity with another variable
note: max1_11 dropped because of collinearity with another variable
Estimating lasso for 1bn.treat using plugin
note: max0_1 dropped because of collinearity with another variable
note: max0_2 dropped because of collinearity with another variable
note: max0_3 dropped because of collinearity with another variable
note: max0_4 dropped because of collinearity with another variable
note: max0_5 dropped because of collinearity with another variable
note: max0_6 dropped because of collinearity with another variable
```

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

Estimating lasso for 2bn.district using plugin

note: max0_1 dropped because of collinearity with another variable

note: max0_2 dropped because of collinearity with another variable

note: max0_3 dropped because of collinearity with another variable

note: max0_4 dropped because of collinearity with another variable

note: max0_5 dropped because of collinearity with another variable

note: max0_6 dropped because of collinearity with another variable

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

Estimating lasso for 1bn.treat#2bn.district using plugin

note: max0_1 dropped because of collinearity with another variable

note: max0_2 dropped because of collinearity with another variable

note: max0_3 dropped because of collinearity with another variable

note: max0_4 dropped because of collinearity with another variable

note: max0_5 dropped because of collinearity with another variable

note: max0_6 dropped because of collinearity with another variable

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

```
note: max1_7 dropped because of collinearity with another variable
note: max1_8 dropped because of collinearity with another variable
note: max1_10 dropped because of collinearity with another variable
note: max1_11 dropped because of collinearity with another variable

Double-selection linear model      Number of obs           =    1,615
                                   Number of controls      =      355
                                   Number of selected controls =   15
                                   Wald chi2(3)            =   111.69
                                   Prob > chi2             =   0.0000

------------------------------------------------------------------------------
               |               Robust
 change_expend |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
       1.treat |   .2841881   2.137063     0.13   0.894    -3.904378    4.472755
    2.district |   51.44027    5.13953    10.01   0.000     41.36697    61.51356
               |
treat#district |
          1 2  |  -31.38468    5.80139    -5.41   0.000    -42.75519   -20.01416
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
      of interest jointly equal to zero. Lassos select controls for model
      estimation. Type lassoinfo to see number of selected variables in each
      lasso.

. estimates store m11

. lassocoef (., for(change_expend)), display(coef, postselection)
```

```
----------------------------
             | change_expend
-------------+-------------
      max1_9 |     322.766
             |
         inc |
          1  |   -15.44611
          3  |   -27.52399
          4  |   -18.12237
             |
  0.rc_aircon |   -30.66556
     logarea |    44.29538
             |
    inc#pool |
       10 0  |    25.23146
             |
   inc#cheat |
        2 0  |   -8.472109
        4 0  |   -12.57901
             |
    inc#dish |
        3 0  |   -.1048343
             |
   inc#ftype |
        1 4  |   -11.93157
             |
inc#c.logarea |
          2  |   -3.811814
```

```
8 |      6.53673
9 |     5.634403
  |
_cons |    -5794.243
-----------------------------
Legend:
  b - base level
  e - empty cell
  o - omitted

. dsregress change_expend treat##district, rseed(10101) controls($controls3)

Estimating lasso for change_expend using plugin
note: max0_1 dropped because of collinearity with another variable
note: max0_2 dropped because of collinearity with another variable
note: max0_3 dropped because of collinearity with another variable
note: max0_4 dropped because of collinearity with another variable
note: max0_5 dropped because of collinearity with another variable
note: max0_6 dropped because of collinearity with another variable
note: max0_7 dropped because of collinearity with another variable
note: max0_8 dropped because of collinearity with another variable
note: max0_9 dropped because of collinearity with another variable
note: max0_10 dropped because of collinearity with another variable
note: max0_11 dropped because of collinearity with another variable
note: max0_12 dropped because of collinearity with another variable
note: max1_1 dropped because of collinearity with another variable
note: max1_2 dropped because of collinearity with another variable
note: max1_3 dropped because of collinearity with another variable
note: max1_4 dropped because of collinearity with another variable
```

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

note: 1.district#c.max0_1 dropped because of collinearity with another variable

note: 2.district#c.max0_1 dropped because of collinearity with another variable

Estimating lasso for 1bn.treat using plugin

note: max0_1 dropped because of collinearity with another variable

note: max0_2 dropped because of collinearity with another variable

note: max0_3 dropped because of collinearity with another variable

note: max0_4 dropped because of collinearity with another variable

note: max0_5 dropped because of collinearity with another variable

note: max0_6 dropped because of collinearity with another variable

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

note: 1.district#c.max0_1 dropped because of collinearity with another variable

note: 2.district#c.max0_1 dropped because of collinearity with another variable

Estimating lasso for 2bn.district using plugin

note: max0_1 dropped because of collinearity with another variable

note: max0_2 dropped because of collinearity with another variable

note: max0_3 dropped because of collinearity with another variable

note: max0_4 dropped because of collinearity with another variable

note: max0_5 dropped because of collinearity with another variable

note: max0_6 dropped because of collinearity with another variable

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

note: 1.district#c.max0_1 dropped because of collinearity with another variable

note: 2.district#c.max0_1 dropped because of collinearity with another variable

Estimating lasso for 1bn.treat#2bn.district using plugin

note: max0_1 dropped because of collinearity with another variable

note: max0_2 dropped because of collinearity with another variable

note: max0_3 dropped because of collinearity with another variable

note: max0_4 dropped because of collinearity with another variable

note: max0_5 dropped because of collinearity with another variable

note: max0_6 dropped because of collinearity with another variable

note: max0_7 dropped because of collinearity with another variable

note: max0_8 dropped because of collinearity with another variable

note: max0_9 dropped because of collinearity with another variable

note: max0_10 dropped because of collinearity with another variable

note: max0_11 dropped because of collinearity with another variable

note: max0_12 dropped because of collinearity with another variable

note: max1_1 dropped because of collinearity with another variable

note: max1_2 dropped because of collinearity with another variable

note: max1_3 dropped because of collinearity with another variable

note: max1_4 dropped because of collinearity with another variable

note: max1_5 dropped because of collinearity with another variable

note: max1_6 dropped because of collinearity with another variable

note: max1_7 dropped because of collinearity with another variable

note: max1_8 dropped because of collinearity with another variable

note: max1_10 dropped because of collinearity with another variable

note: max1_11 dropped because of collinearity with another variable

note: 1.district#c.max0_1 dropped because of collinearity with another variable

note: 2.district#c.max0_1 dropped because of collinearity with another variable

Double-selection linear model   Number of obs   =   1,615

Number of controls   =   355

Number of selected controls =   15

```
                         Wald chi2(3)       =    111.69
                         Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |               Robust
change_expend |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
     1.treat |   .2841881   2.137063     0.13   0.894    -3.904378    4.472755
  2.district |   51.44027    5.13953    10.01   0.000     41.36697    61.51356
             |
treat#district |
        1 2  |  -31.38468    5.80139    -5.41   0.000    -42.75519   -20.01416
------------------------------------------------------------------------------
Note: Chi-squared test is a Wald test of the coefficients of the variables
    of interest jointly equal to zero. Lassos select controls for model
    estimation. Type lassoinfo to see number of selected variables in each
    lasso.

. estimates store m12

. lassocoef (., for(change_expend)), display(coef, postselection)

----------------------------
             | change_expend
-------------+--------------
     max1_9 |      322.766
             |
         inc |
          1  |    -15.44611
```

```
          3  |    -27.52399
          4  |    -18.12237
             |
 0.rc_aircon |    -30.66556
     logarea |     44.29538
             |
    inc#pool |
       10 0  |     25.23146
             |
   inc#cheat |
        2 0  |    -8.472109
        4 0  |    -12.57901
             |
    inc#dish |
        3 0  |    -.1048343
             |
   inc#ftype |
        1 4  |    -11.93157
             |
inc#c.logarea |
          2  |    -3.811814
          8  |      6.53673
          9  |     5.634403
             |
       _cons |    -5794.243
-----------------------------
Legend:
  b - base level
  e - empty cell
```

o - omitted

.

. esttab m10 m11 m12, label title("Table 4 Regression Results (Average Treatment Effects On Locations With Similar Weather Patterns)") nonumbers mtitles("SIM OLS (No controls)" "SIM Double Selection" "FMM Double Se

> lection") modelwidth(25) se replace drop(0.treat 1.district 0.treat#1.district 0.treat#2.district 1.treat#1.district)

Table 4 Regression Results (Average Treatment Effects On Locations With Similar Weather Patterns)

	SIM OLS (No controls)	SIM Double Selection	FMM Double Selection
treat=1	-0.885	0.284	0.284
	(2.623)	(2.137)	(2.137)
district=2	51.21***	51.44***	51.44***
	(5.338)	(5.140)	(5.140)
treat=1 # district=2	-27.19***	-31.38***	-31.38***
	(6.856)	(5.801)	(5.801)
Constant	107.0***		
	(2.028)		
Observations	1615	1615	1615

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

```
. restore

. * <><><> *
.
. log close
     name:  <unnamed>
      log:  \\Client\C$\Users\georg\Downloads\Appendix3.txt
 log type:  text
closed on:  30 Jul 2021, 03:56:07
---------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------
```

z5254903.doc

GRADEMARK REPORT

FINAL GRADE

89/100

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

Comment 1

this is confusing. I don't get what you mean here

Comment 2

don't have evidence for this recommendation

PAGE 3

Comment 3

additionally

PAGE 4

Comment 4

good, would have been better if having a column presenting differences in means and doing t-test to check if the differences are statistically significant

Comment 5

good

Comment 6

good

what about other households where the changes are above 400?

PAGE 5

Comment 7

what about min temp?

PAGE 6

Comment 8

did you include both weather before and after the experiment?

Comment 9

any justification for the inclusion of controls?

PAGE 7

Comment 10

this one could be used as a starting point for the motivation for proposed econometric approach

Comment 11

good

PAGE 8

Comment 12

what about differences in sample size? moreover, observations with these two variables of weather variables might be systematically different to the rest?

PAGE 9

Comment 13

why? the explanation isn't convincing.

PAGE 10

Comment 14

good

Comment 15

how come? both audit and treat*audit are insignificant

Comment 16

this is good

Comment 17

but it doesn't seem to have evidence in favor of this?

Comment 18

still need to mention the results

Text Comment. Good project! Key features of the data were well characterized and you did a good job in producing and evaluating econometric models, especially the use of the double selection method and checking for the effect of the audit variable. However, it is important to discuss the choices of controls in your models.

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41

PAGE 42

PAGE 43

PAGE 44

PAGE 45

PAGE 46

PAGE 47

PAGE 48

PAGE 49

PAGE 50

PAGE 51

PAGE 52

PAGE 53

PAGE 54

PAGE 55

PAGE 56

PAGE 57

PAGE 58

PAGE 59

ANALYSIS 1 3 / 3

Motivation for proposed econometric approach and models.

MISSING (0)	
POOR (1)	Not clear that the key research questions are understood and/or little or no evidence of an appropriate plan for the modelling undertaken given the available data.
ADEQUATE (2)	Understands the key research questions and provides an appropriate plan for addressing these questions given the available data.
GOOD (2.50)	Good understanding of the key research questions and a thoughtful plan appropriate for addressing these questions given the available data. Recognition of some of the strengths and weaknesses of the approach taken.
EXCELLENT (3)	Comprehensive understanding of the key research questions and an insightful plan appropriate for addressing these questions given the available data. Recognition and understanding of the strengths and weaknesses of the approach taken.

ANALYSIS 2 2 / 2

Characterizing key features of the data.

MISSING (0)	
POOR (0.50)	Only superficial attention to the data and the key features that provide context relevant for the subsequent analysis of the research questions.
ADEQUATE (1)	Sufficient description of the data and the key features to provide some context relevant for the subsequent analysis of the research questions.
GOOD (1.50)	Comprehensive description of the data and the key features that provides the necessary context relevant for the subsequent analysis of the research questions and interpretation of the subsequent analysis. Good use of basic analyses beyond the basic summary statistics.
EXCELLENT (2)	Comprehensive and insightful description of the data and the key features that provides the necessary context relevant for the subsequent analysis of the research questions and interpretation of the subsequent analysis. Very good use of basic analyses beyond the basic summary statistics.

ANALYSIS 3 4 / 5

Production and evaluation of econometric results.

MISSING (0)	

POOR (2)	Modelling is superficial and inadequate to support analysis of the key research questions. No systematic evidence of an understanding of how to appropriately interpret the econometric results.
ADEQUATE (3)	Implements the proposed modelling in a competent manner and provides appropriate interpretation of the econometric results.
GOOD (4)	Implements the proposed modelling and demonstrates a good understanding of the methods being used. Explores some potential threats and extensions to the basic analysis and provides clear and appropriate interpretation of the econometric results.
EXCELLENT (5)	Implements the proposed modelling and demonstrates a very good understanding of the methods being used. Explores some potential threats and extensions to the basic analysis and provides clear and appropriate interpretation of the econometric results. Analysis is comprehensive, often showing good insights into the modelling problem and the methods used.

DISCUSSION 1 1.50 / 2

Justification of main conclusions.

MISSING (0)	
POOR (0.50)	A mismatch between the analysis and the conclusions drawn from that analysis either because there is a lack of understanding of how to interpret the results and/or conclusions are inappropriate given the analysis that was performed and/or the key questions have not been addressed.
ADEQUATE (1)	Reasonable conclusions are drawn from the analysis that was conducted that address the key research questions.
GOOD (1.50)	Reasonable and appropriate conclusions are drawn from the analysis that was conducted. Demonstrates a good understanding of how the results are translated into conclusions that address the key research questions.
EXCELLENT (2)	Reasonable and appropriate conclusions are drawn from the analysis that was conducted. Demonstrates a very good understanding of how the results are translated into conclusions that address the key research questions.

DISCUSSION 2 2 / 2

Executive summary

MISSING (0)	
POOR (0.50)	There is an over-use of technical jargon and/or the discussion of the core research questions is inadequate because of key omissions and/or the inclusion of too much non-essential material.
ADEQUATE (1)	The core results have been presented but there is room for improvement in terms of the over-use of technical jargon and/or the inclusion of non-essential material.

GOOD (1.50)	The core results have been presented in an accessible and logical fashion that avoids the use of technical jargon.
EXCELLENT (2)	The presentation of the core results demonstrates an excellent understanding of how to present relatively technical material in a clear, accessible and logical fashion.

PRESENTATION 3.50 / 4

MISSING (0)	
POOR (2)	No obvious well-defined structure for the report and relationships between ideas and arguments are unclear. May contain material, such as tables, that are not presented carefully or accurately.
ADEQUATE (3)	Structure of the report is acceptable but could have been enhanced by more clarity in the flow of material and extra clarity in presenting the relationships between analysis and conclusions. Tables are presented adequately but potentially with room for improvement.
GOOD (3.50)	An effective organization of component parts enhances the presentation of the required material and the links between analysis and conclusions are clear and logical. Appropriate tables of results are presented in an effective manner that enhance the presentation.
EXCELLENT (4)	An effective organization of component parts enhances the presentation of the required material and the links between analysis and conclusions are clear, logical and insightful. Appropriate tables of results are of high quality consistent with a very professional presentation.