Codebook

Unless otherwise stated, column definitions are as follows

1. Treated + all available controls
2. Treated + 1:1 matched controls based on pre-treatment consumption
3. Treated + 1:1 matched controls based on building characteristics
4. Treated + 1:1 matched controls based on pre-treatment consumption and building characteristics
5. Treated only

TWFE estimates for total energy; monthly data

Dependent Variable:	log(energy)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
treated_postTRUE	-0.1571***	-0.1453***	-0.1512***	-0.1396***	-0.1608***
	(0.0063)	(0.0072)	(0.0070)	(0.0074)	(0.0109)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
cons_date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	2,882,662	429,578	429,833	429,866	210,503
R^2	0.79219	0.80557	0.81004	0.81181	0.81159
Within R^2	0.00267	0.01011	0.01134	0.00962	0.00701

Clustered (id & cons_date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

TWFE estimates for gas; monthly data

Dependent Variable:	log(gas)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
treated_postTRUE	-0.2081***	-0.1944***	-0.1993***	-0.1901***	-0.2260***
	(0.0062)	(0.0076)	(0.0074)	(0.0078)	(0.0093)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
cons_date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	2,874,433	428,688	429,313	429,103	210,050
R^2	0.85408	0.85549	0.85894	0.86033	0.86047
Within R^2	0.00384	0.01381	0.01480	0.01352	0.01040

Clustered (id & cons_date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

TWFE estimates for electricity; monthly data

Dependent Variable:	log(elec)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
treated_postTRUE	-0.0387***	-0.0138	-0.0084	0.0015	0.0390
	(0.0099)	(0.0121)	(0.0123)	(0.0117)	(0.015
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
cons_date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	2,908,550	432,238	432,263	432,424	211,76
R^2	0.46709	0.50410	0.49463	0.49444	0.5078
Within R^2	8.72×10^{-5}	5.66×10^{-5}	2.07×10^{-5}	7.29×10^{-7}	0.0002

Clustered (id & cons_date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sun + Abraham estimates

The Sun and Abraham correction is estimated by interacting a cohort dummy with a time-to-treatment dummy. In our monthly data, we observe 49 different "cohorts" (i.e., households retrofit in 49 different months). Our data has 250 different "time to treatments" (i.e., months before and after retrofit). This implies estimating a model with about 10,000 dummy variables. This is not feasible. Instead, to estimate the Sun and Abraham model, I convert to annual data first. The results first show the TWFE estimates based on annual data.

TWFE estimates for energy; annual data

Dependent Variable:	log(energy)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
treated_postTRUE	-0.1682***	-0.1573***	-0.1605***	-0.1489***	-0.1910***
	(0.0060)	(0.0079)	(0.0077)	(0.0090)	(0.0086)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
consyear	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	246,541	35,175	35,165	35,178	16,419
R^2	0.80826	0.84582	0.85472	0.85053	0.85274
Within R^2	0.01388	0.06664	0.07291	0.06048	0.03305

Clustered (id & consyear) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sun+Abraham estimates for energy; annual data

Dependent Variable:	log(energy)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
ATT	-0.1668***	-0.1544***	-0.1577***	-0.1468***	-0.0570***
	(0.0037)	(0.0049)	(0.0047)	(0.0048)	(0.0092)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
consyear	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	246,541	35,175	35,165	35,178	16,419
R^2	0.80831	0.84640	0.85521	0.85139	0.85298
Within R^2	0.01414	0.07016	0.07603	0.06588	0.03463

Clustered (id & consyear) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sun+Abraham estimates for gas; annual data

Dependent Variable:	log(gas)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
ATT	-0.1958***	-0.1848***	-0.1890***	-0.1781***	-0.0390***
	(0.0042)	(0.0055)	(0.0053)	(0.0054)	(0.0055)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
consyear	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	247,154	35,234	35,244	35,248	16,460
R^2	0.82162	0.84958	0.86034	0.85334	0.85702
Within R^2	0.01857	0.08732	0.09513	0.08423	0.04636

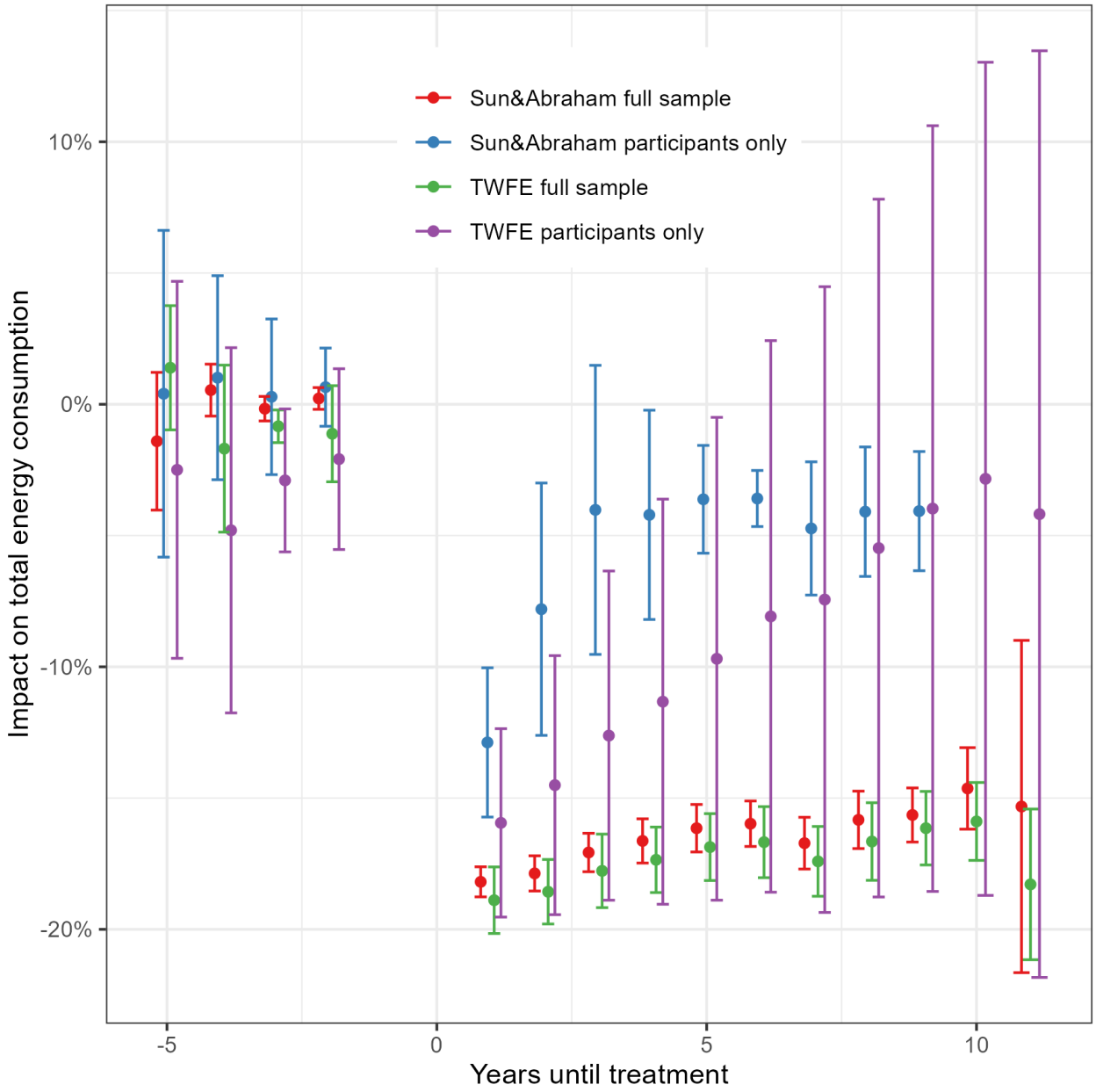
Clustered (id & consyear) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sun+Abraham estimates for electricity; annual data

Dependent Variable:	log(elec)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
ATT	-0.0495***	-0.0265**	-0.0195*	-0.0146	-0.0172
	(0.0071)	(0.0093)	(0.0101)	(0.0093)	(0.0229)
Fixed-effects					
id	Yes	Yes	Yes	Yes	Yes
consyear	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	246,819	35,183	35,170	35,180	16,420
R^2	0.64327	0.70877	0.70024	0.70385	0.72055
Within R^2	0.00060	0.00172	0.00150	0.00112	0.00172

Clustered (id & consyear) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Event study plot



Manual cohort-by-cohort analysis

The Sun and Abraham estimator (like other new TWFE estimators) is simply as weighted aggregate of DID estimators for all treated cohorts. To better understand what's going on, I am going to manually estimate treatment effects for all cohorts. In doing so, I can also decompose the treatment effect into three types of comparisons (a la Goodman-Bacon): treated vs. never treated; later treated vs. earlier treated; and earlier treated vs. later treated. Using earlier treatments as controls is the potentially problematic case.

Dynamic cohort effects

In the following slide, I use the participants only sample, and only make "good" comparisons – i.e., comparing earlier treated (the treatment group) with not-yet-treated (the control group) households. I run the regression for the 2008, 2009, and 2010 cohorts. I drop years between the pre- and post-retrofits. The treatment effects are very similar to what we observed earlier under the TWFE estimator. They are NOT the same as under the Sun and Abraham estimator.

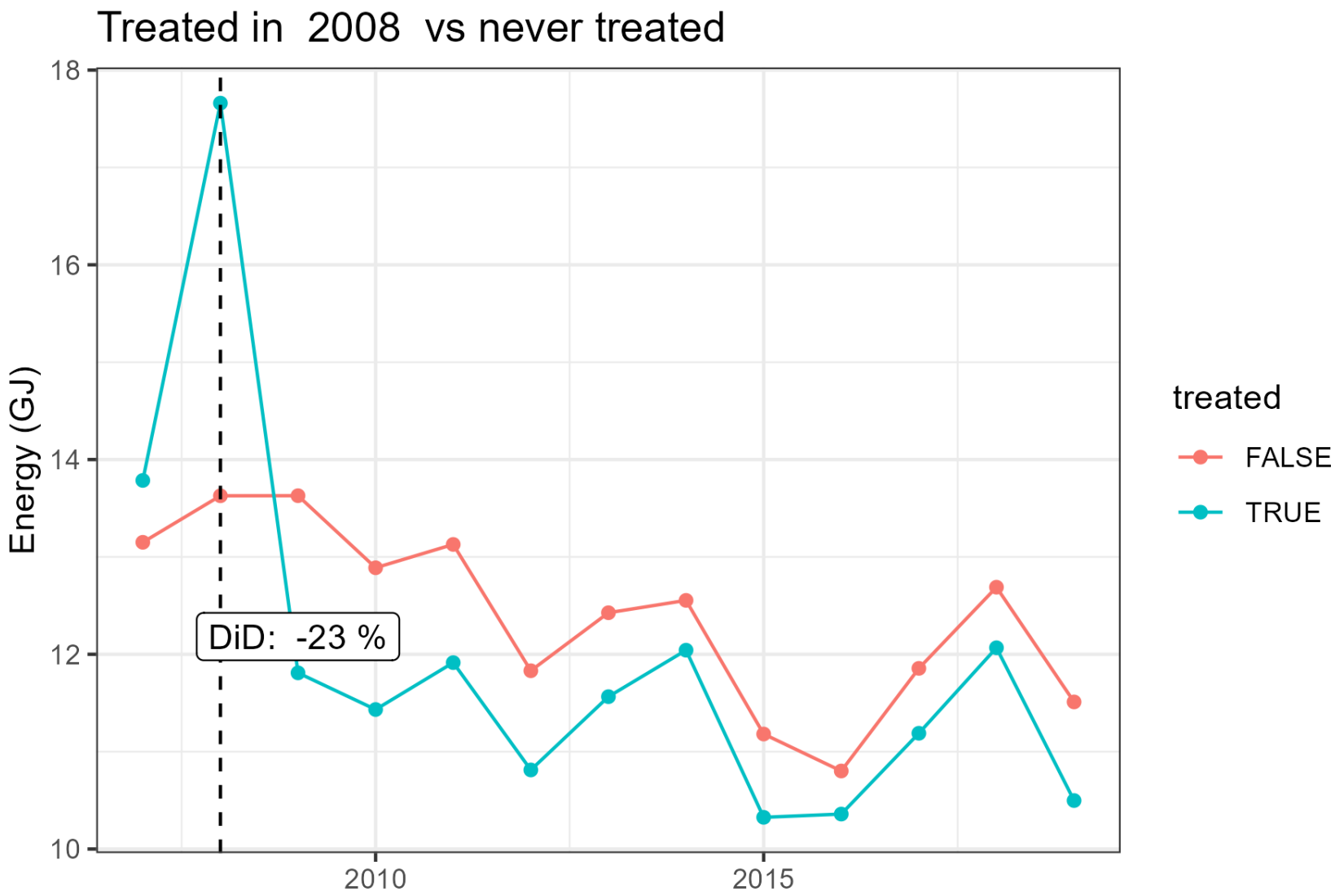
Dynamic cohort effects

Dependent Variable:	log(energy)		
Model:	(1)	(2)	(3)
Variables			
years_to_treatment = -4	-0.0051	-0.0201	-0.0057
	(0.0120)	(0.0124)	(0.0135)
years_to_treatment = -3	-0.0072	-0.0163*	-0.0068
	(0.0087)	(0.0088)	(0.0104)
years_to_treatment = -2	-0.0038	-0.0126**	-0.0054
	(0.0051)	(0.0054)	(0.0053)
years_to_treatment = 1	-0.1969***	-0.1769***	-0.1625***
	(0.0585)	(0.0140)	(0.0324)
years_to_treatment = 2	-0.1720***	-0.1494***	
	(0.0527)	(0.0340)	
years_to_treatment = 3	-0.1322**		
	(0.0591)		
Fixed-effects			
id	Yes	Yes	Yes
consyear	Yes	Yes	Yes
Fit statistics			
Observations	4,107	6,932	8,616
R^2	0.92203	0.91032	0.89377
Within R^2	0.02146	0.06381	0.01327

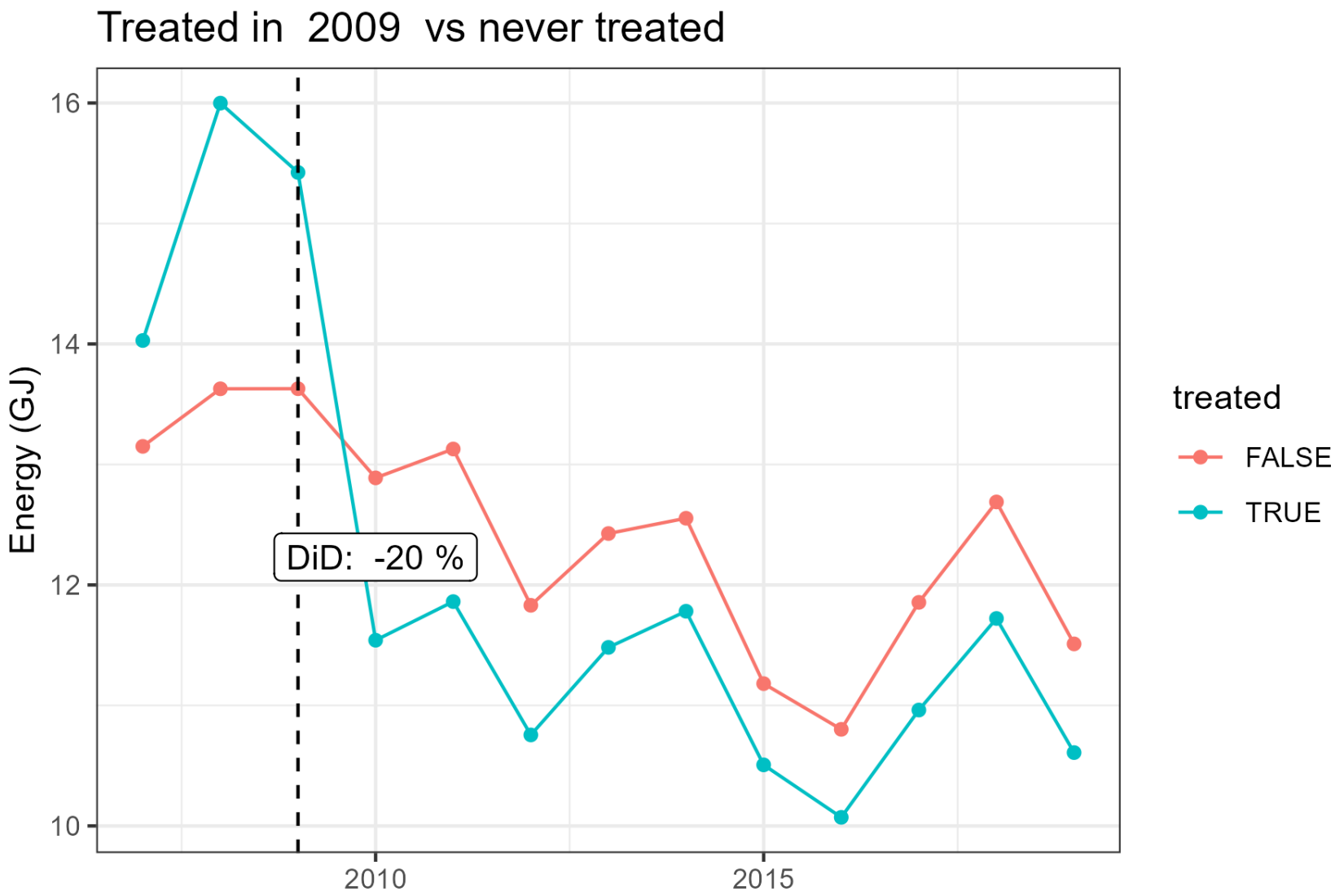
Clustered (id) standard-errors in parentheses

Now, I use the data WITH untreated households, and illustrate DID results from all possible comparisons. I do this graphically, but also show a regression coefficient. I didn't drop the between-retrofit period in producing these figures (although I could).

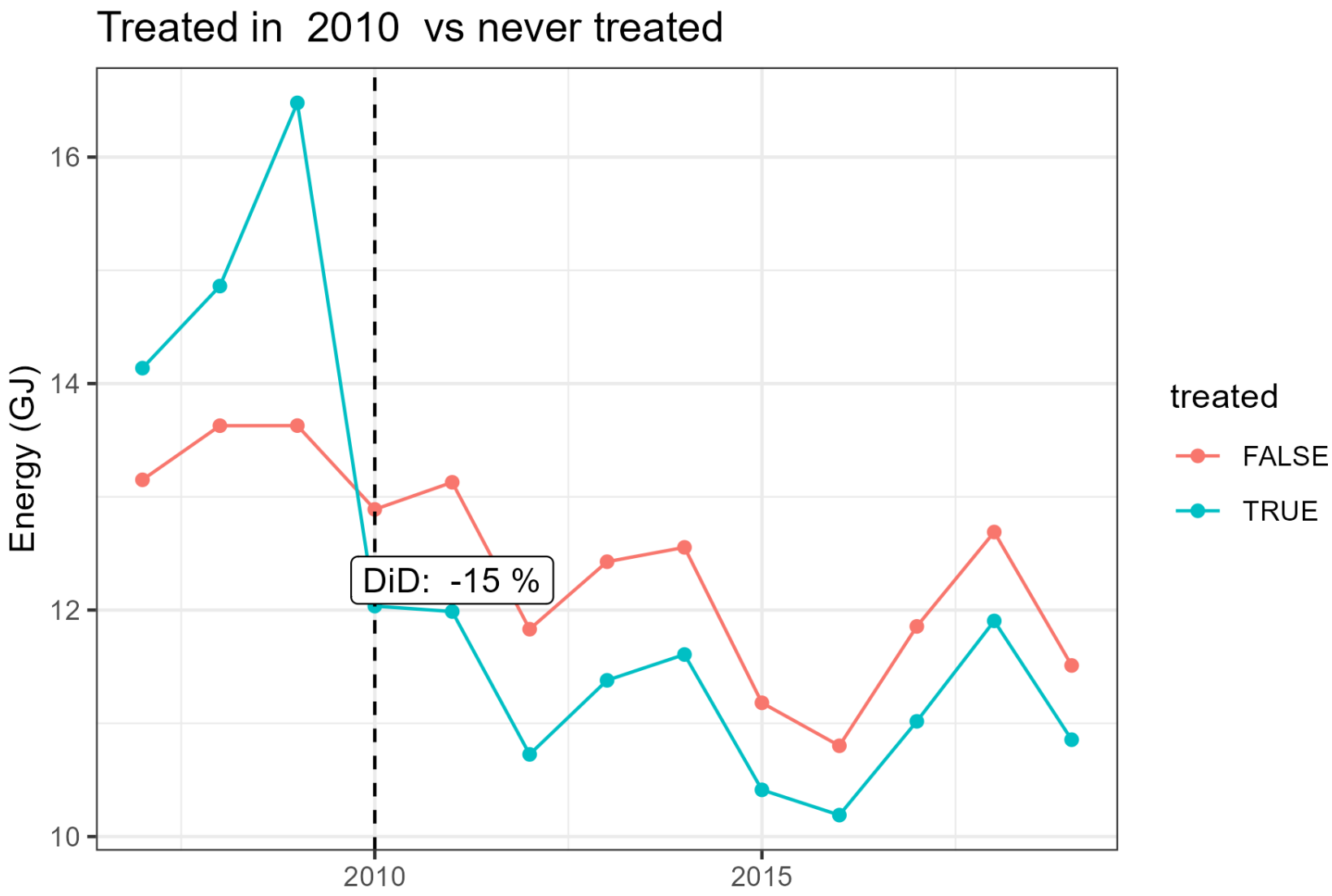
Treated vs. never treated



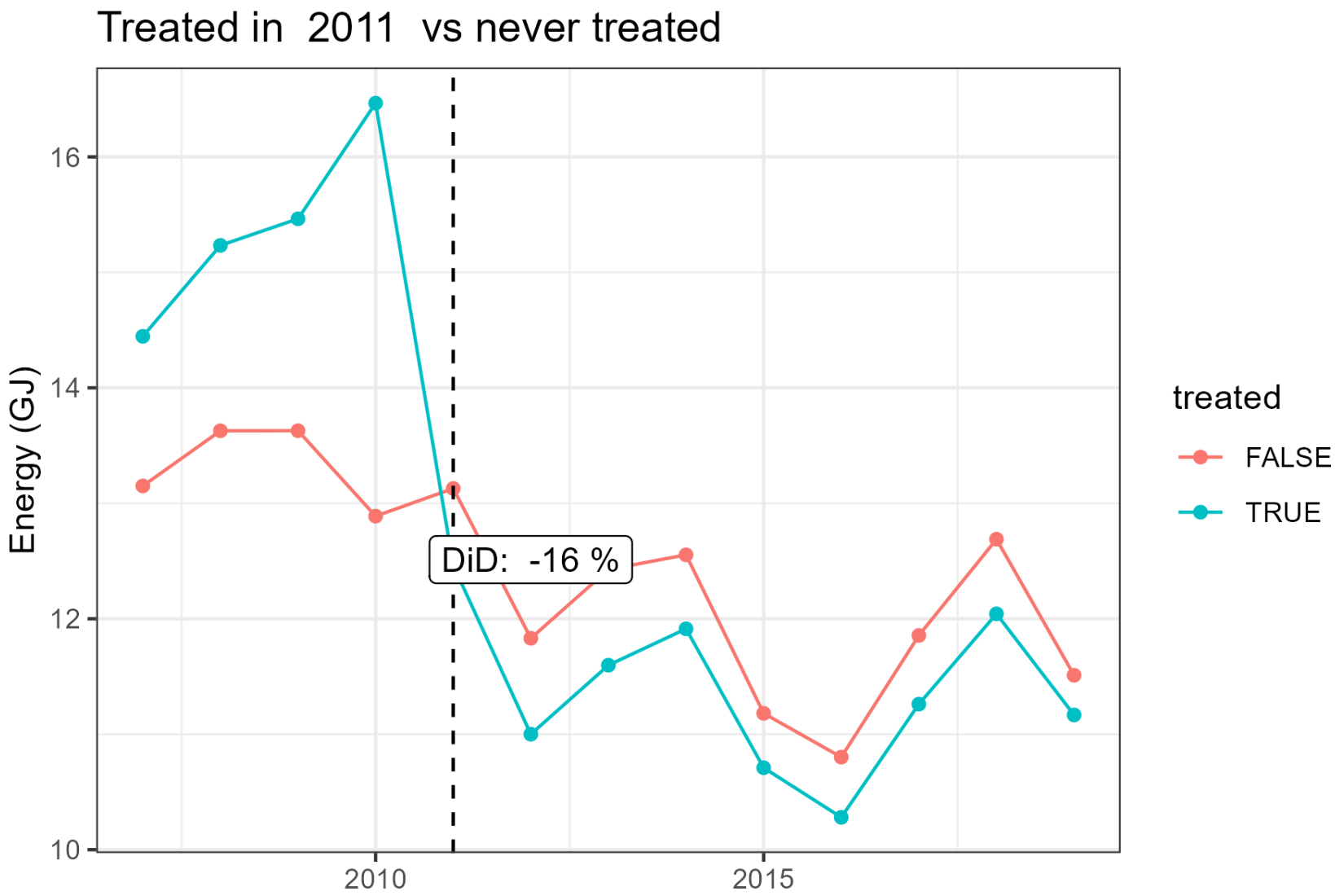
Treated vs. never treated



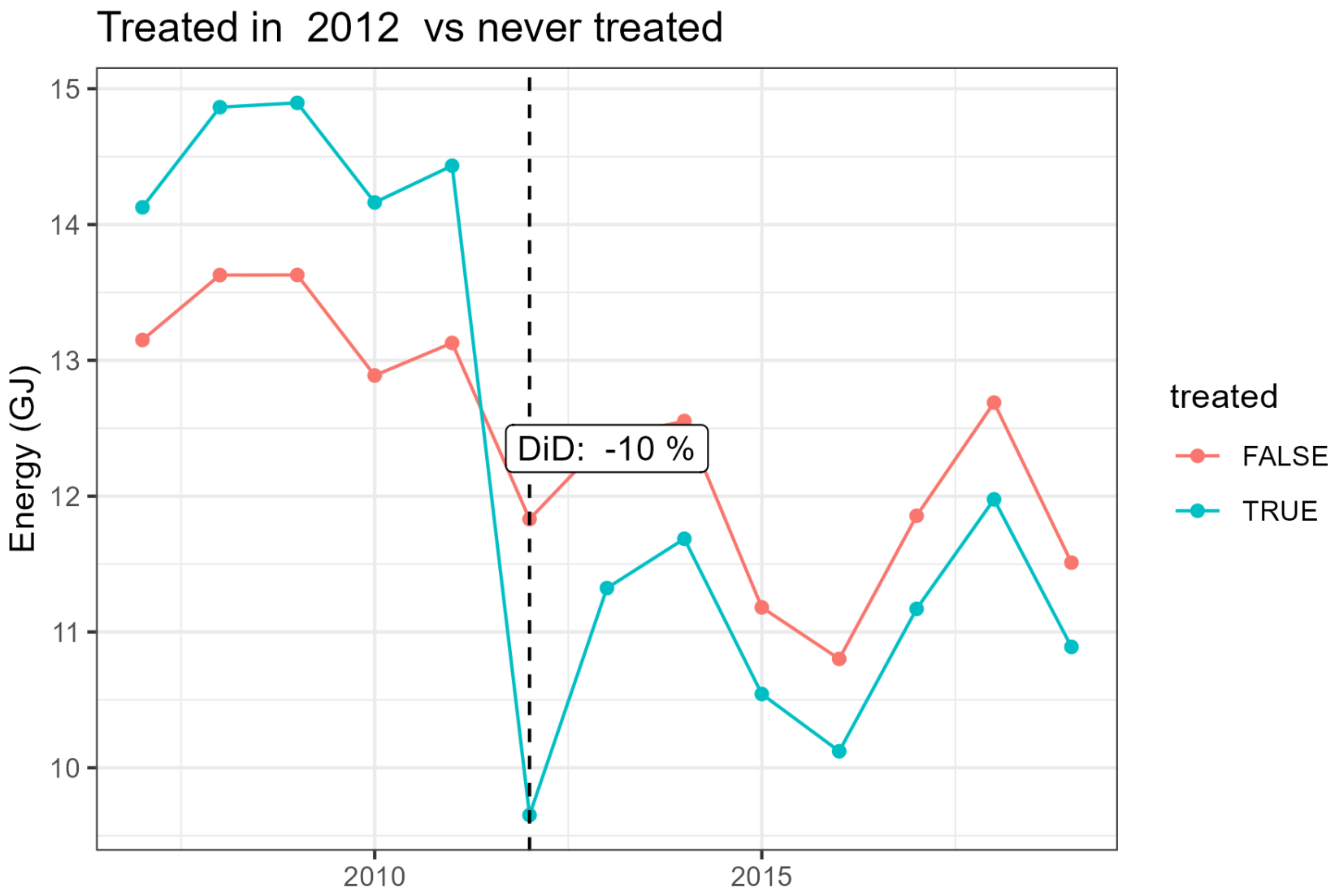
Treated vs. never treated



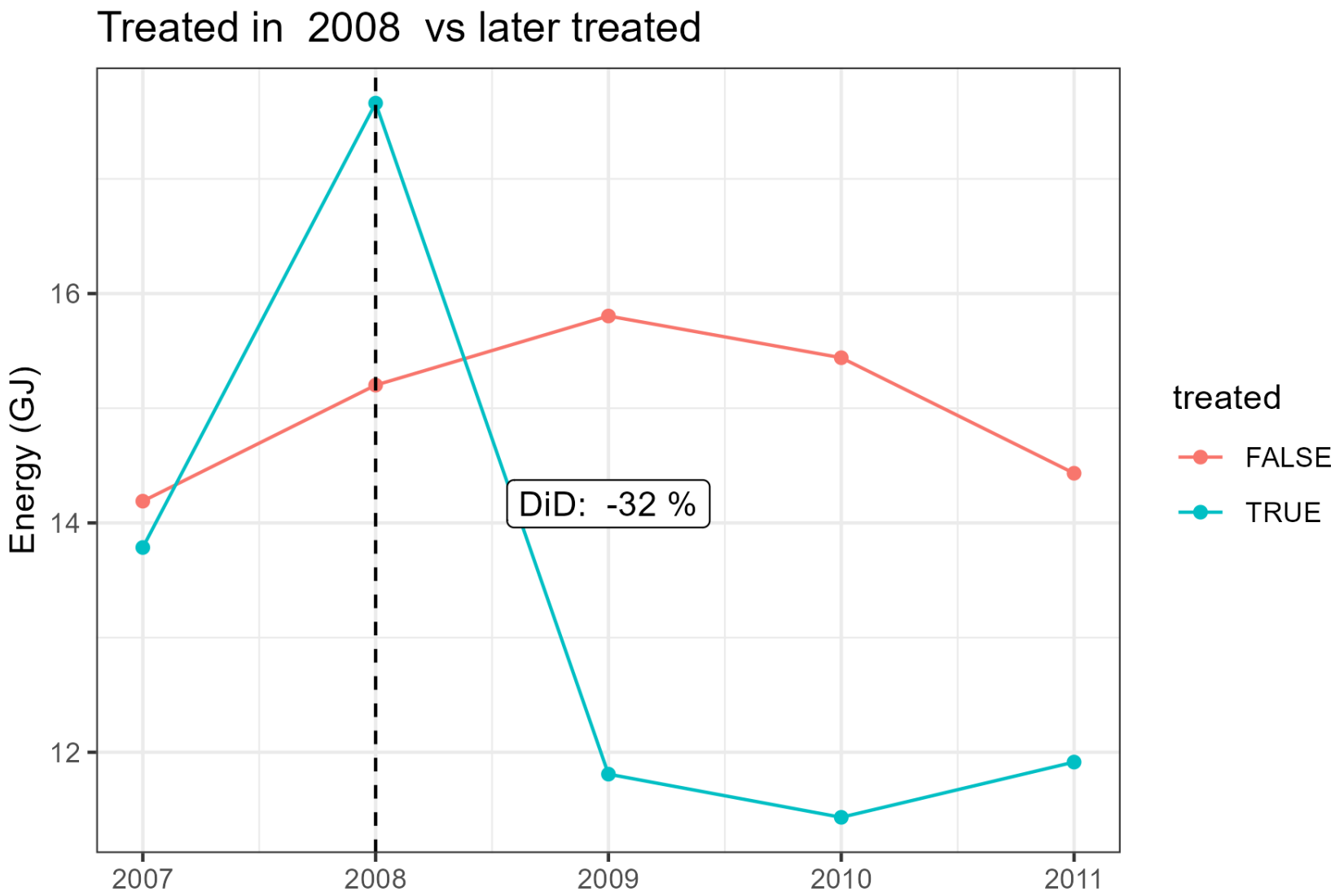
Treated vs. never treated



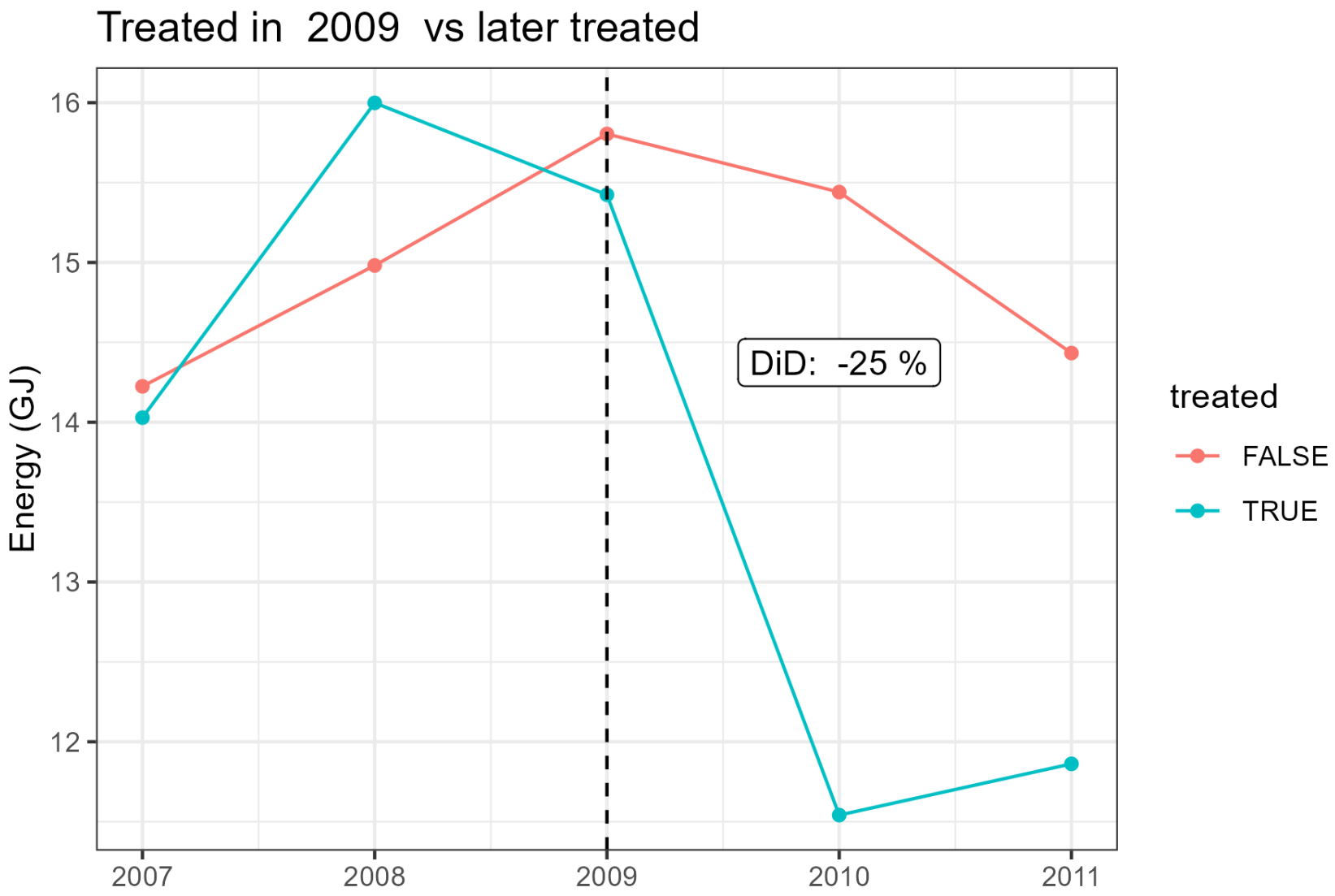
Treated vs. never treated



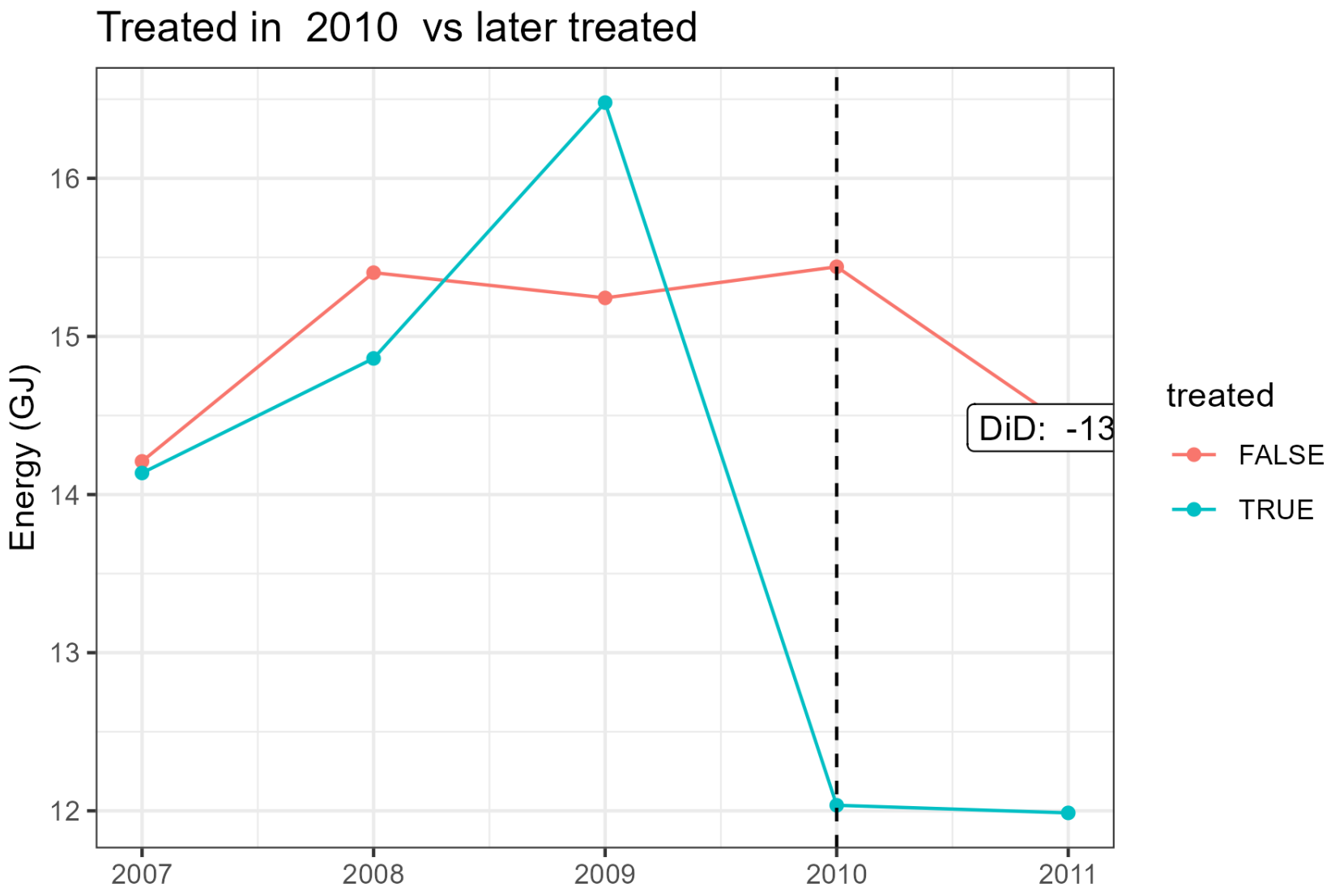
Earlier treated vs. later treated



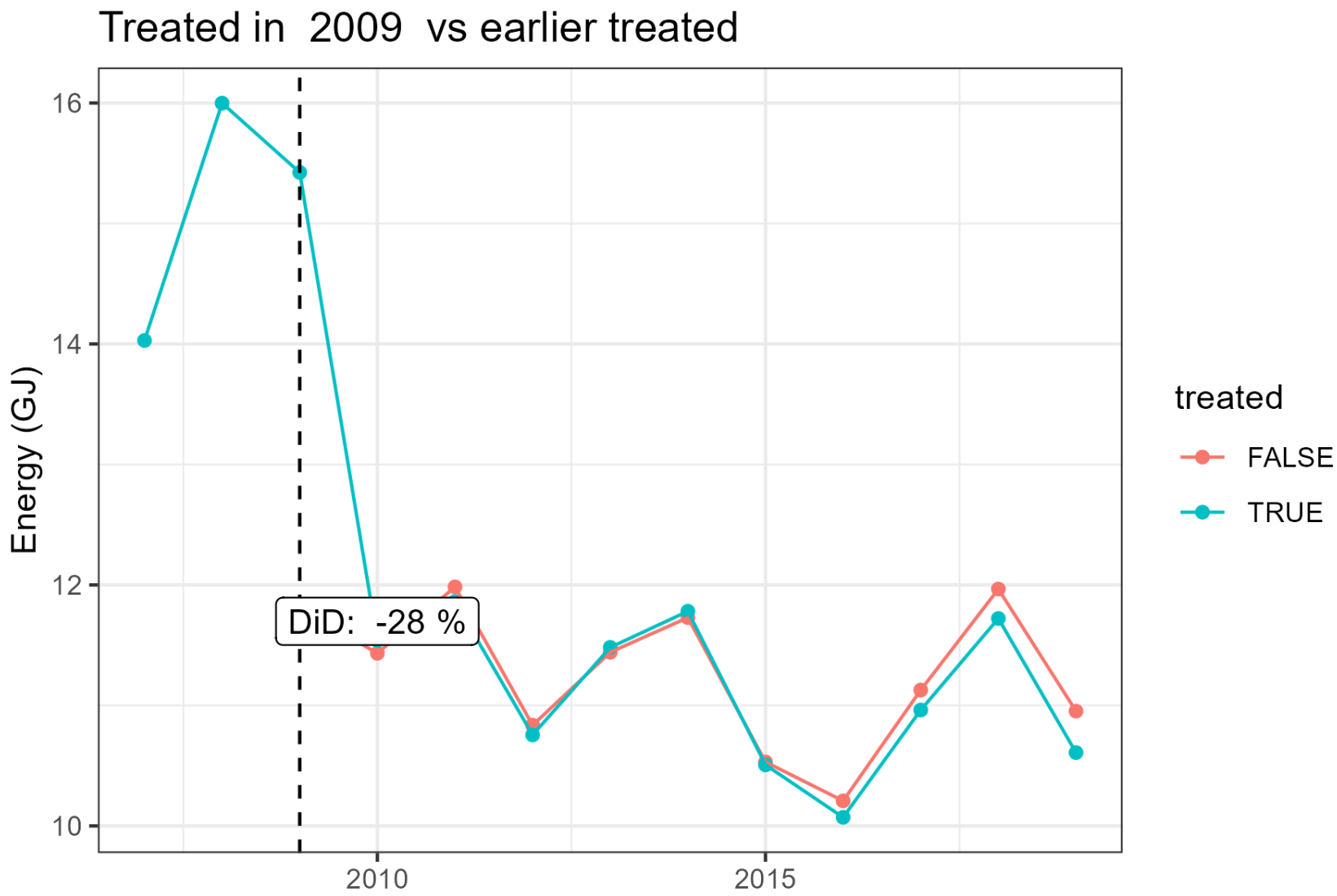
Earlier treated vs. later treated



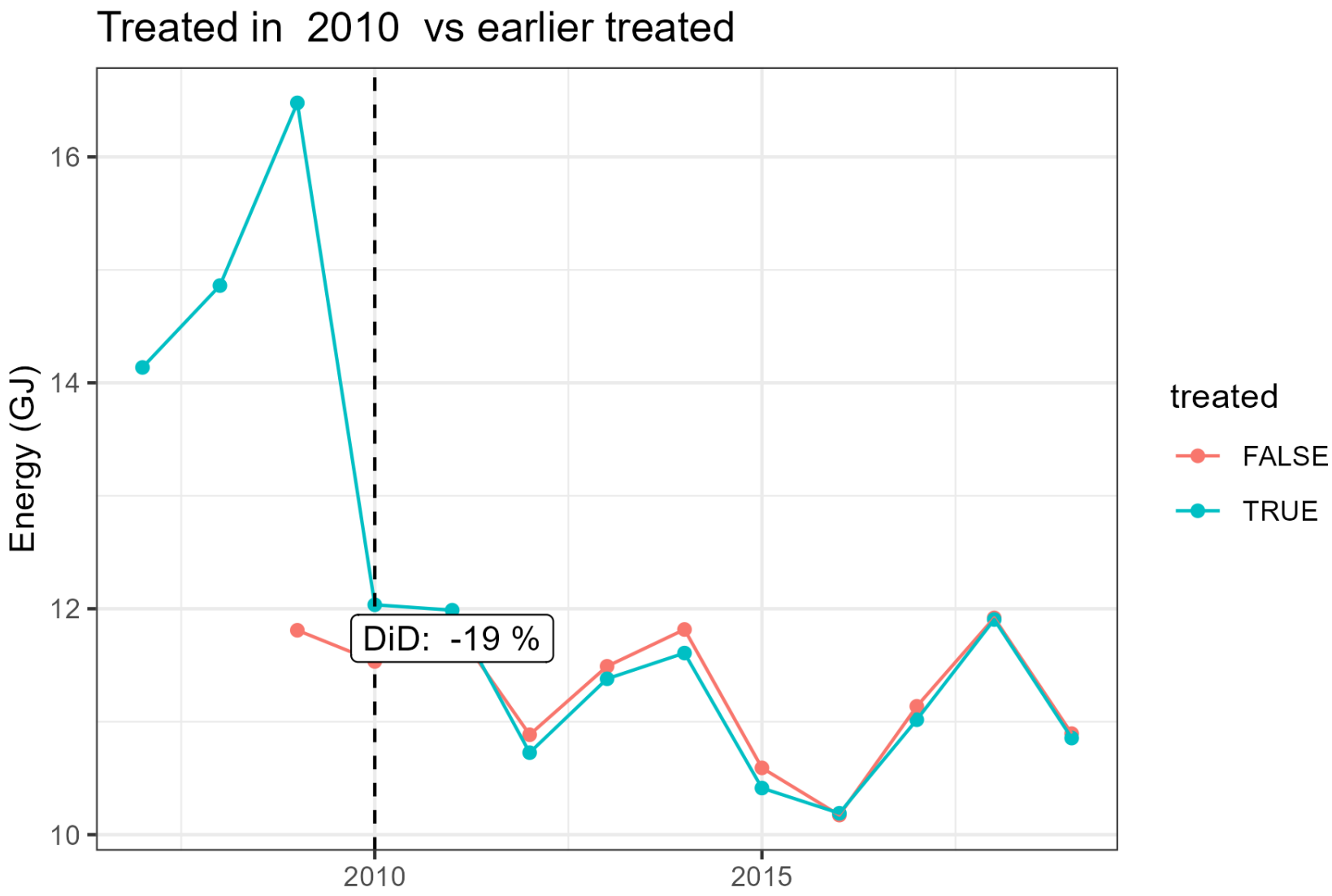
Earlier treated vs. later treated



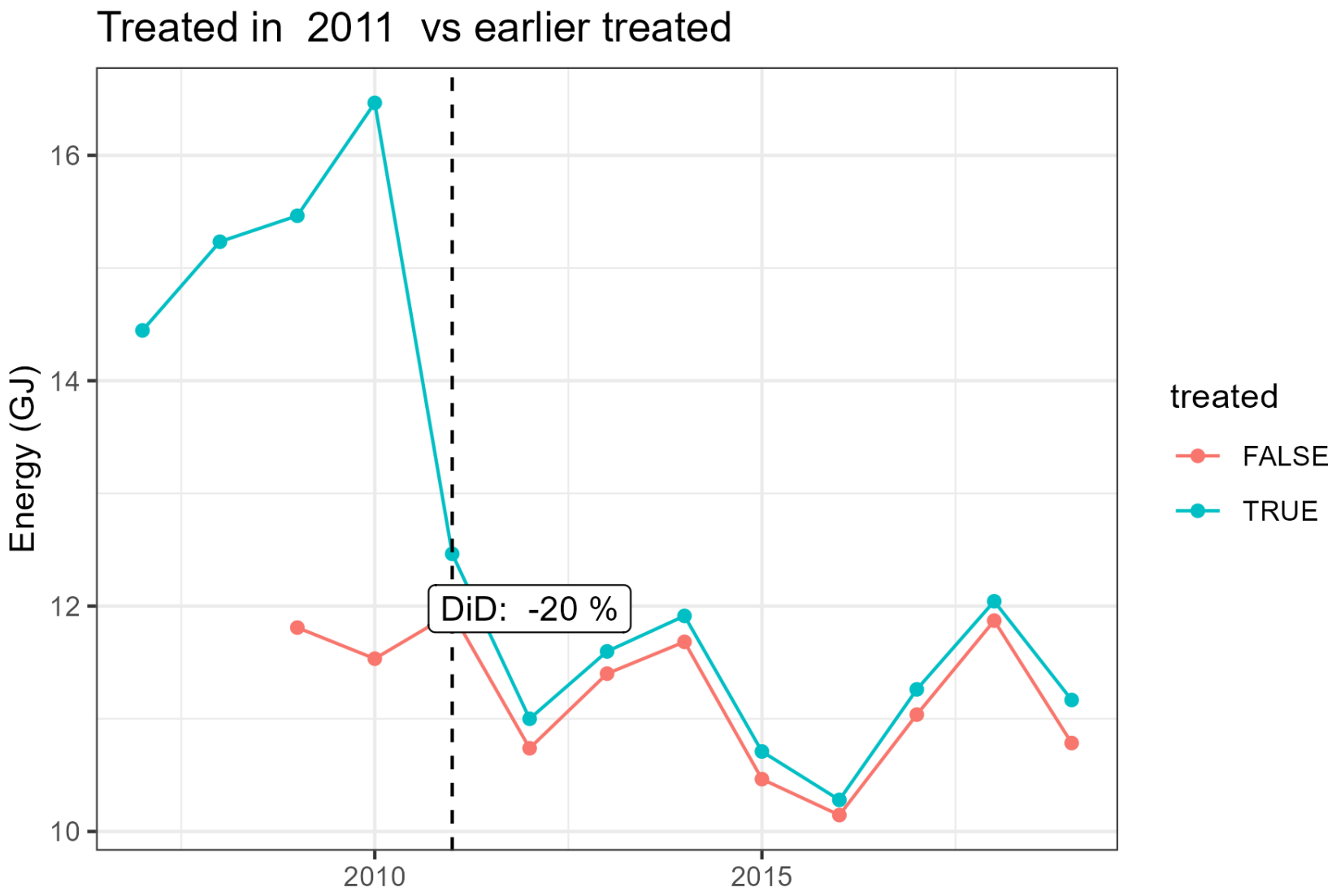
Later treated vs. earlier treated



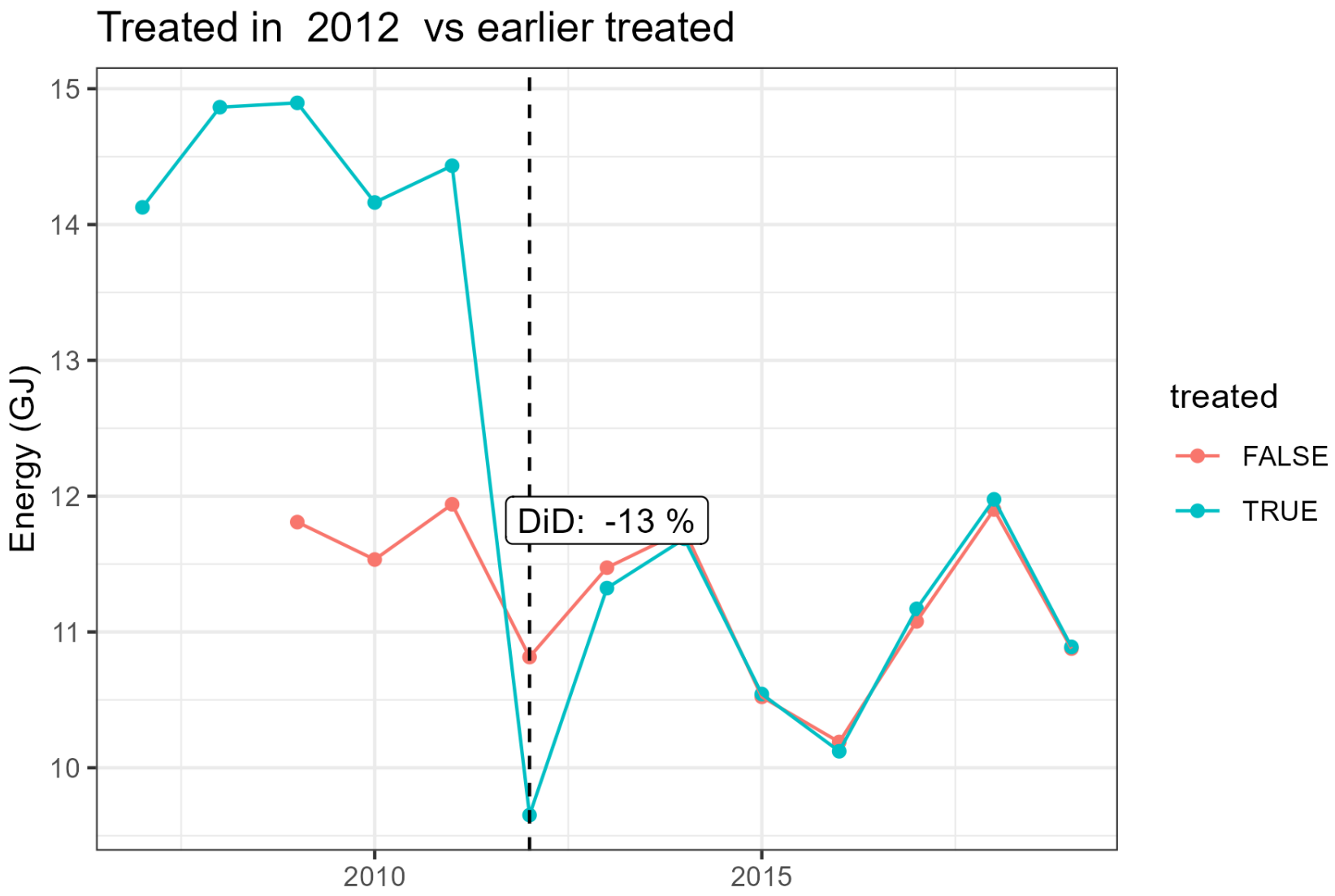
Later treated vs. earlier treated



Later treated vs. earlier treated



Later treated vs. earlier treated



Summary

All of these comparisons reveal a coefficient of around 16%, just like our main estimates in the current version of the paper. This leads me to think that I have not estimated the Sun and Abraham model properly, and that the effect of retrofits really is an energy saving of approximately 16%. Moreover, the event study plot in slide 9 looks "wrong": with the particpants only sample, it should not be possible to estimate coefficients on treatment+4 years or more, since we don't observe an appropriate control group.