Difference in difference : Obesity

Isaac Meza

November 5, 2021

---

1 Methodology

Data consists of an unbalanced panel of 128056 households for which we have data (at least) for the period spanning 2013 and the first 3 months of 2014. Our main variables are

1. Kcal of SD
2. Kcal of non-SD
3. Kcal of HCF
4. Kcal of non-HCF
5. Total taxable expenditure

Data is at week-individual level but we collapse it (by mean) at the monthly-individual level. We smooth all series using a MA with 3 lags, 2 forward terms and current observation; so that the smoother applied (by individual) is

$$(1/6)[x_{t-3} + x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}]$$

The first task is to define a pure treatment/control group. We will do this by choosing an *optimal* partition of the total taxable expenditure distribution, as high spenders will be more likely to be more sensitive to a price change in SD and HCF, therefore we will define this as the treatment group. Total taxable expenditure is defined as total expenditure in SD and HCF.

The *optimal* partition is found by solving the following problem

$$
\begin{aligned}
\min_{H,L} \quad & \sum_{t=-12}^{-2} |\beta_t| \\
\text{s.t} \quad & \\
& (\beta_t)_{-12\leq t\leq 12} = \operatorname{argmin}\Bigg\{\Bigg(y_{it} - \sum_{k=-12}^{12} \alpha_k \mathbb{1}(t=k) - \\
& \qquad \sum_{k=-12}^{12} \beta_k \mathbb{1}(i=T, k=t) + \gamma \mathbb{1}(i=T) - \lambda_i\Bigg)^2\Bigg\} \\
& T = \mathbb{1}(x_i \geq H) \\
& C = \mathbb{1}(x_i \leq L) \\
& \min(x_i) \leq L \leq H \leq \max(x_i)
\end{aligned}
$$

Note that the coefficients β solve for a fixed effects regression including time calendar dummies and leads and lags in treatment effect[1] Moreover $(\beta_t)_{t=-12}^{-2}$ gives a 'test' on parallel trends, so what we are looking for is to find the optimal partition of Treatment/Control group so that a parallel trend is preserved. We do this in order to try to capture 'true' treatment effects.

Also note that in principle the partition is allowed to be non-symmetrical or to not span the whole distribution.

We use as a dependent variable (y) total calories of SD and HCF and total calories of SD. Variable (x_i) corresponds to total taxable expenditure of individual i. Finally, H and L are the respective cuts on the distribution to define Treatment/Control groups. Note that event study corresponds to $t = 0$.

[1] As recommended by Borusyak and Jaravel (2016), but unlike McCrary (2007) and most event study papers, include all relative time dummies in the regression rather than "binning" periods below a or above b. Then we can just graph the periods from a to b if we want. But binning can cause bias if the trend isn't flat for periods less than a or greater than b (Borusyak and Jaravel, 2016). Note that when there is no pure control group, binning periods less than a or greater than b (i.e. imposing flat trend for those periods) is needed to pin down calendar time fixed effects, which is why Borusyak and Jaravel (2016) recommend having a pure control, which pin down the calendar time fixed effects without having to make these additional assumptions.

2 Results

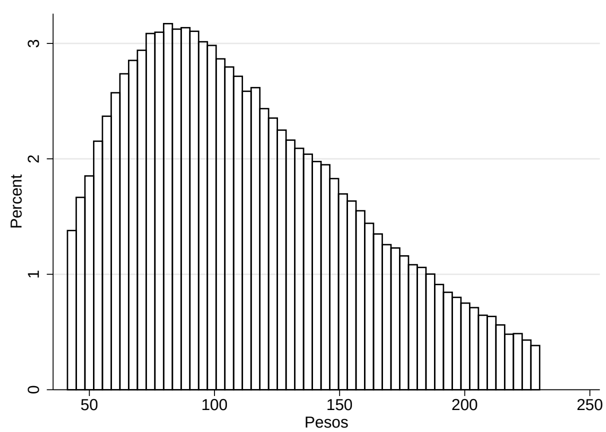
Once we find the optimal "cuts" on the distribution of total taxable expenditure and define our pure Treatment/Control group, we graph

(a) The average calorie consumption throughout time by treatment group

(b) The coefficients of the fixed effect regression with leads and lags

The following graph shows the distribution (cut at the 95th percentile) of the total taxable expenditure.

Figure 1: Distribution taxable expenditure

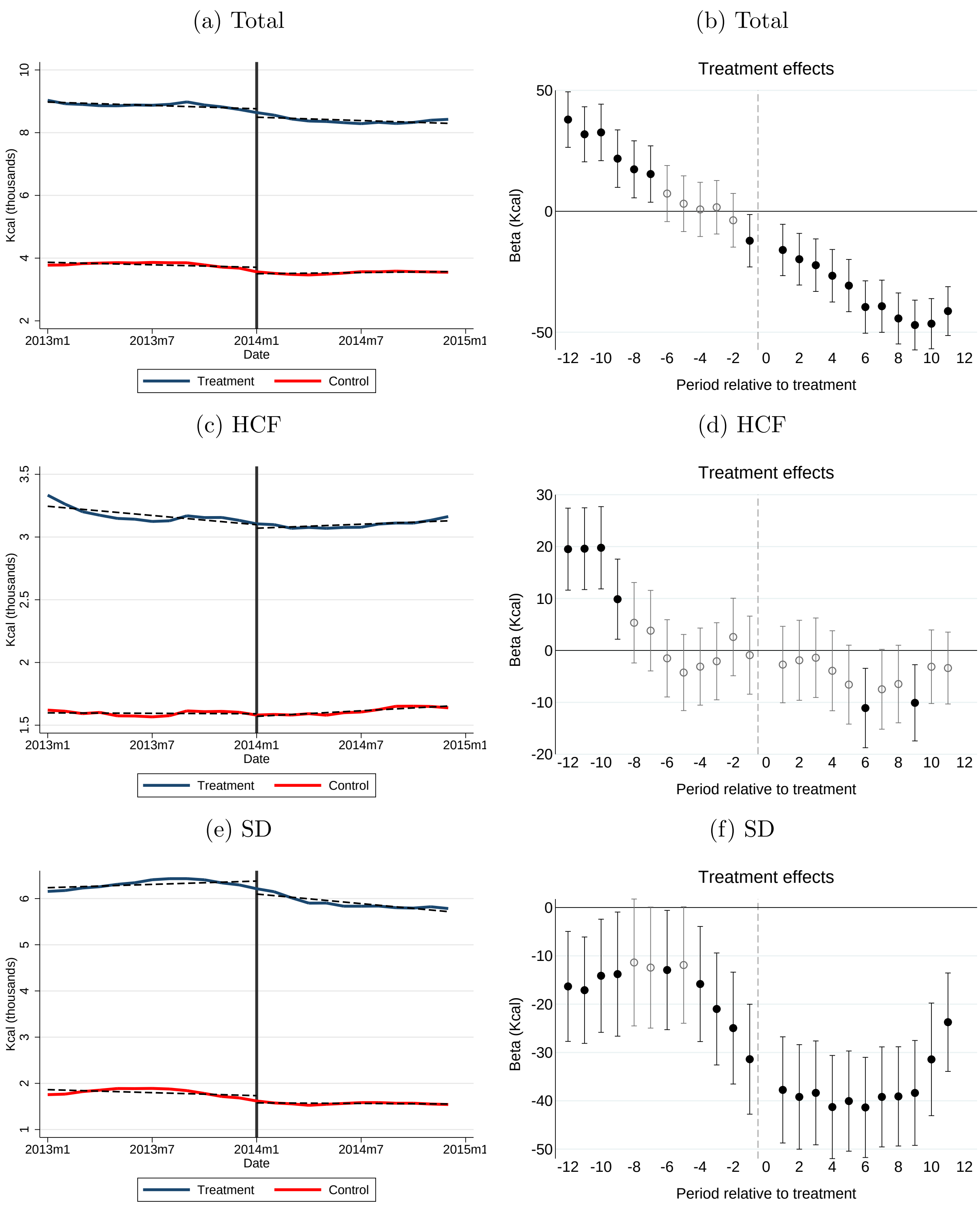


Notes: Do file: `dist_totalexp.do`

The DiD specification is the following:

$$y_{it} = \sum_{k=-12}^{12} \alpha_k \mathbb{1}(t = k) + \sum_{k=-12}^{12} \beta_k \mathbb{1}(i = T, k = t) + \gamma \mathbb{1}(i = T) - \lambda_i + \epsilon_{it}$$

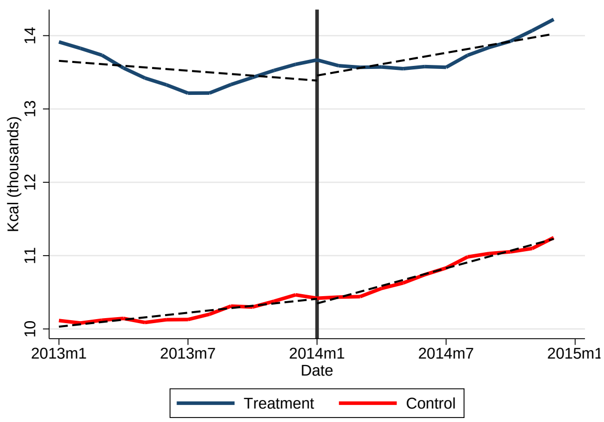
Figure 2: Treatment effects



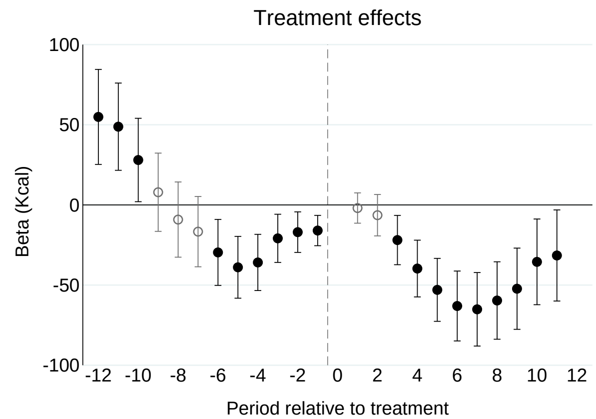
Notes: Do file: `did.do` , `beta_coef_did.do`

Figure 3: Treatment effects - placebo

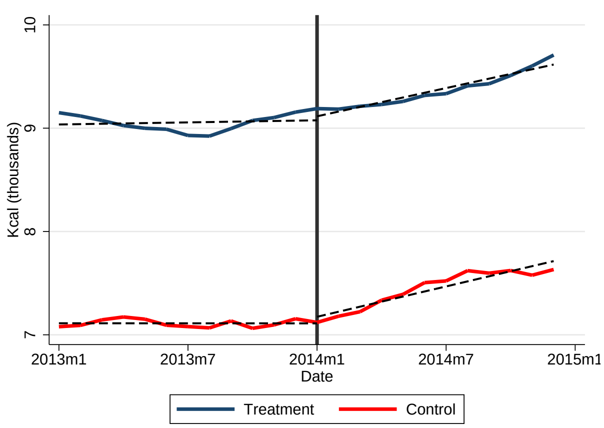
(a) Total calories placebo



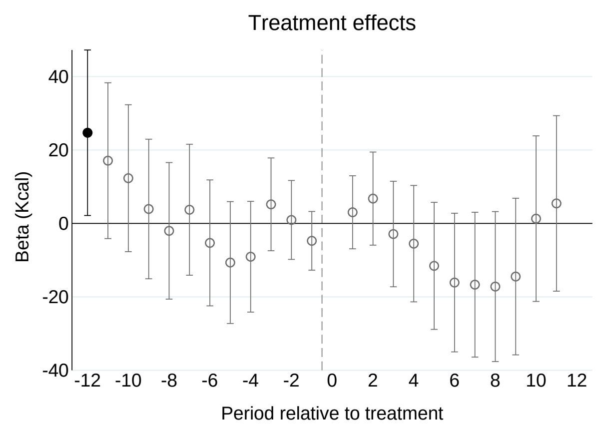
(b) Total calories placebo



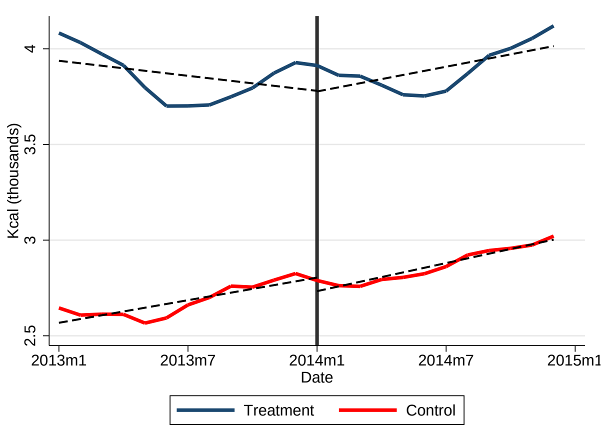
(c) non-HCF



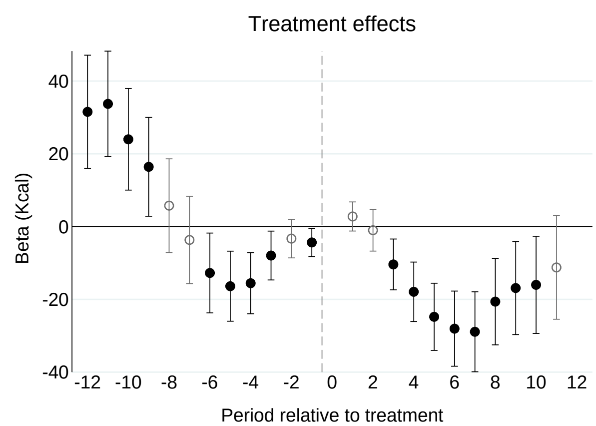
(d) non-HCF



(e) non-SD



(f) non-SD



Notes: Do file: `did.do` , `beta_coef_did.do`