E-Appendix for "Power calculation for the trend-in-trend design"

1. Notation

We first introduce the notation required for the power calculation.

- N: number of individuals.
- T: number of time points.
- $\boldsymbol{X}_i$: vector of covariates associated with individual i, which represents intrinsic characteristics that might influence a particular exposure and/or outcome.
- Z_{it}: exposure indicator for individual i at time t.
- Y_{it}: outcome for individual i at time t.
- G: the index for the subject's CEP group.

2. Assumptions on the data generating law

Trend-in-trend analysis relies on the following assumptions:

A1. Covariates and time have multiplicative effects on being exposed, i.e., $p(Z_{it} = 1|\boldsymbol{X}_i) = h_1(\boldsymbol{X}_i)h_2(t)$.

A2. Covariates for all individuals in any subgroup G are random variables from an unknown distribution. i.e., $p(\boldsymbol{X}_i|G) = f_G$.

A3. The outcome is a rare event that follows a logit model.

3. Monte Carlo power calculation

The power calculation requires the following parameters: 1) the type-1 error rate; 2) the probability of a study subject experiencing the study outcome during any study interval; 3) the c-statistic of the CPE model; 4) the number of CPE strata into which the population is divided; 5) the shape of the exposure trend, expressed as a linear or quadratic function of time on log scale; and 6) the desired statistical power or minimum detectable causal odds ratio.

Given these parameters, we can perform the proposed Monte Carlo power analysis using the following steps:

a) Construct a treatment assignment model as $p(Z_{it} = 1|\boldsymbol{X}_i = x) = \frac{e^{\alpha_x X + \alpha_t w(t)}}{1+e^{\alpha_x X + \alpha_t w(t)}}$. The parameters α_x, and α_t are specified using the c-statistic and the exposure prevalence, respectively. The function $w(t)$ denotes the shape of the exposure trend.
b) Construct an outcome model as $p(Y_t = 1|Z_t,) = \frac{e^{\beta_0+\beta_1 Z_t}}{1+e^{\beta_0+\beta_1 Z_t}} \approx e^{\beta_0+\beta_1 Z_t}$. The parameters β_0 and β_1 are specified using the probability of a study subject experiencing the outcome and the odds ratio of interest, respectively.
c) Using models in (a) and (b), generate M datasets of sizes N and calculate the proportion of times that the null hypothesis H_0: $\beta_1 = 0$ is rejected. The test is performed by constructing a confidence interval for the estimated β_1 using TT analysis with g number of CPE strata. If the confidence interval does not contain zero, the null is rejected.

d) To find a detectable difference with certain power, increase the odds ratio β_1 and repeat steps (b) and (c) until the desired power is achieved.