DiD Estimator Moment Conditions

Deriving a Generalized Maximum Entropy (GME) Difference-in-Differences (DiD) estimator involves combining the principles of maximum entropy estimation with the DiD framework. The idea behind GME is to estimate the parameters of interest by maximizing the entropy of the probability distribution subject to the moment conditions derived from the data.

Step 1: Set Up the Moment Condition

From the previous discussion, the moment condition for the DiD estimator is:

$$\mathbb{E}[(Y_{i1} - Y_{i0}) - \delta D_i \mid X_i] = 0$$

This condition can be written as:

$$\mathbb{E}[Z_i \cdot (Y_{i1} - Y_{i0} - \delta D_i)] = 0$$

where Z_i is an instrument or a set of instruments (typically, $Z_i = 1$ for a simple case, but it could include covariates X_i).

Step 2: Generalized Maximum Entropy Formulation

To implement the GME estimator, we express the unknown parameter δ as a discrete probability distribution over a finite support. Let the parameter δ be expressed as:

$$\delta = \sum_{k=1}^{K} p_k \delta_k$$

where:

- δ_k are the support points, which are fixed.
- p_k are the probabilities associated with each support point, satisfying $p_k \geq 0$ and $\sum_{k=1}^{K} p_k = 1$.

The GME objective is to maximize the entropy of the distribution of δ subject to the moment conditions.

Step 3: Entropy Maximization Problem

The entropy of the distribution $p = \{p_1, p_2, \ldots, p_K\}$ is given by:

$$H(p) = -\sum_{k=1}^{K} p_k \log(p_k)$$

The goal is to maximize $H(p)$ subject to the moment conditions:

$$\sum_{i=1}^{N} Z_i \cdot \left[Y_{i1} - Y_{i0} - \Big(\sum_{k=1}^{K} p_k \delta_k \Big) D_i \right] = 0$$

and the probability constraints:

$$\sum_{k=1}^{K} p_k = 1, \quad p_k \geq 0 \text{ for all } k.$$

Step 4: Set Up the Lagrangian

The Lagrangian for this problem is:

$$\mathcal{L}(p, \lambda, \mu) = -\sum_{k=1}^{K} p_k \log(p_k) + \lambda \Big(\sum_{k=1}^{K} p_k - 1 \Big) + \mu \sum_{i=1}^{N} Z_i \cdot \left[Y_{i1} - Y_{i0} - \Big(\sum_{k=1}^{K} p_k \delta_k \Big) D_i \right]$$

where λ and μ are Lagrange multipliers.

Step 5: First-Order Conditions

Taking the derivative of the Lagrangian with respect to p_k and setting it to zero gives:

$$-\log(p_k) - 1 + \lambda - \mu \sum_{i=1}^{N} Z_i D_i \delta_k = 0$$

This simplifies to:

$$p_k = \exp\Big(-1 + \lambda - \mu \sum_{i=1}^{N} Z_i D_i \delta_k \Big)$$

Given the constraint $\sum_{k=1}^{K} p_k = 1$, we normalize the probabilities:

$$p_k = \frac{\exp\left(-\mu \sum_{i=1}^{N} Z_i D_i \delta_k \right)}{\sum_{j=1}^{K} \exp\left(-\mu \sum_{i=1}^{N} Z_i D_i \delta_j \right)}$$

Step 6: Solve for the GME Estimator

The solution to the GME estimator involves the following steps:

1. **Choose the support points** δ_k: These are typically chosen based on the range of plausible values for δ.

2. **Estimate** p_k using the expressions derived above by solving the entropy maximization problem.

3. **Compute the estimator** δ as a weighted sum of the support points:

$$\hat{\delta} = \sum_{k=1}^{K} \hat{p}_k \delta_k$$

where $\hat{p}_k$ are the estimated probabilities.

Summary

The Generalized Maximum Entropy DiD estimator provides a nonparametric way to estimate the treatment effect by maximizing the entropy subject to the moment condition derived from the DiD framework. This approach can be particularly useful in settings where the traditional DiD assumptions may not hold perfectly or where the distribution of treatment effects is of interest.