Регрессионный анализ: панельные данные и каузальность

Лекция 2
Модель с фиксированными эффектами на временные периоды
Модель с разными наклонами

FE-модель на временные периоды

Представим спецификацию модели с фиксированными эффектами в формате модели с набором дамми-переменных (LSDV)

FE-модель на временные периоды

Представим спецификацию модели с фиксированными эффектами в формате модели с набором дамми-переменных (LSDV)

$$\hat{y}_{it} = \hat{\beta}_0 + \hat{\gamma}_1 D_{1t} + \ldots \hat{\gamma}_{T-1} D_{(T-1)t} + \hat{\beta}_1 x_{it}$$

FE-модель на временные периоды

Представим спецификацию модели с фиксированными эффектами в формате модели с набором дамми-переменных (LSDV)

$\hat{y}_{it} = \hat{\beta}_0 + \hat{\gamma}_1 D_{1t} + ...\hat{\gamma}_{T-1} D_{(T-1)t} + \hat{\beta}_1 x_{it}$

- $\hat{\beta}_0$ – чему в среднем равно значение зависимой переменной в базовом временном периоде при равенстве предикторов нулю

FE-модель на временные периоды

Представим спецификацию модели с фиксированными эффектами в формате модели с набором дамми-переменных (LSDV)

$\hat{y}_{it} = \hat{\beta}_0 + \hat{\gamma}_1 D_{1t} + ...\hat{\gamma}_{T-1} D_{(T-1)t} + \hat{\beta}_1 x_{it}$

- $\hat{\beta}_0$ – чему в среднем равно значение зависимой переменной в базовом временном периоде при равенстве предикторов нулю
- $\hat{\gamma}_i$ – на сколько в среднем отличается значение зависимой переменной в t-ом временном периоде от базового при прочих равных
- $\hat{\beta}_0 + \hat{\gamma}_t$ – индивидуальная константа (фиксированный эффект)

FE на временные периоды

Содержательно охватывают весь набор неизменяющихся в пространстве характеристик. Эти факторы влияют на все пространственные единицы анализа (к примеру, страны, регионы, компании, индивиды) одновременно

FE на временные периоды

Содержательно охватывают весь набор неизменяющихся в пространстве характеристик. Эти факторы влияют на все пространственные единицы анализа (к примеру, страны, регионы, компании, индивиды) одновременно

Примеры:

- природные катаклизмы
- COVID-19
- финансовые кризисы
- военные конфликты

FE на временные периоды

Содержательно охватывают весь набор неизменяющихся в пространстве характеристик. Эти факторы влияют на все пространственные единицы анализа (к примеру, страны, регионы, компании, индивиды) одновременно

Примеры:

- природные катаклизмы
- COVID-19
- финансовые кризисы
- военные конфликты

То есть, включение фиксированных эффектов не позволяет полностью избавиться от эндогенности, так как мы можем пропустить существенные переменные, характеризующие различия между пространственными единицами

Модель с внутригрупповым преобразованием

Мы можем представить FE-модель в более экономном виде (то есть, с меньшим количеством параметров), при этом сохранив скорректированную на панельную структуру данных оценку коэффициента наклона

Алгоритм:

1. рассчитываем центрированный y, при этом считаем средние по временным периодам: $y^*_{it} = y_{it} - \overline{y}_{\bullet t}$
2. аналогичным образом преобразуем предикторы: $x^*_{it} = x_{it} - \overline{x}_{\bullet t}$
3. оцениваем регрессию y^*_{it} на x^*_{it}:

$$\hat{y}^*_{it} = \hat{\beta}_1 x^*_{it}$$

Как рассчитывается $\hat{\beta}_1^{time}$?

$\hat{\beta}_1^{time}$ в такой FE-модели можно получить на основании соответствующих коэффициентов регрессий, оцененных на отдельных T подвыборках

Нас интересует оценка коэффициента при предикторе

$$\hat{y}^*_{it} = \hat{\beta}_1^{time} x^*_{it}$$

1. Для каждого из T временных периодов оценим регрессию $\hat{y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 x_i$ и сохраним $\hat{\alpha}_1$ для каждого t-го периода
2. Суммируем взвешенные значения $\hat{\alpha}_1$:

$$\hat{\beta}_1^{time} = \frac{\sum_{t=1}^{T} w_t \hat{\alpha}_{1t}}{\sum_{t=1}^{T} w_t}$$

$$w_t = \sum_{i=1}^{N_t} (x_{it} - \bar{x}_{\cdot t})^2$$

Выводы по формуле $\hat{\beta}_1^{time}$ в FE-модели:

- Временные периоды с большей межгрупповой изменчивостью предиктора (то есть, изменчивостью предиктора между пространственными единицами в один и тот же период) вносят больший вклад в оценку $\hat{\beta}_1^{time}$
- Те временные периоды, в которых предиктор одинаков для всех пространственных единиц, не участвуют в формировании оценки коэффициента $\hat{\beta}_1^{time}$ в FE-модели
- $\hat{\beta}_1^{time}$ формируется на основании межгрупповой изменчивости предиктора внутри каждого периода. Поэтому интерпретация $\hat{\beta}_1^{time}$ будет в общем виде следующей:

  В один и тот же период времени, при различии предиктора на одну единицу между пространственными единицами, зависимая переменная в среднем отличается на $\hat{\beta}_1^{time}$ при прочих равных условиях

Модель с разными наклонами

Не всегда правдоподобно, что взаимосвязь предиктора и зависимой переменной остается постоянной для пространственных единиц / временных периодов. Смоделируем разные взаимосвязи, показав это на примере разных наклонов для разных пространственных единиц:

$$\hat{y}_{it} = \hat{\beta}_0 + \hat{\gamma}_1 D_{1i} + \ldots \hat{\gamma}_{n-1} D_{(n-1)i} + \hat{\beta}_1 x_{it} + \hat{\mu}_1 D_{1i} x_{it} + \ldots \hat{\mu}_{n-1} D_{(n-1)i} x_{it}$$

Интерпретация оценок коэффициентов модели с разными наклонами

1. $\hat{b}_0$ – то, чему в среднем равен y_{it} в стране – базовой категории – при всех предикторах равных 0
2. $\hat{\gamma}_i$ – отклонение y_{it} в среднем в i-ой стране от y_{it} в базовой категории при всех предикторах равных 0
3. $\hat{b}_1$ – насколько в среднем при прочих равных при увеличении x_{it} на 1 изменяется отклик в стране – базовой категории
4. $\hat{\mu}_i$ – насколько в среднем отличается взаимосвязь x_{it} и y_{it} в i-ой стране в отличие от базовой категории при прочих равных