Analiza danych panelowych – podejścia uproszczone

Rozważamy *N* jednostek w *T* okresach

1. Regresje liniowe dla każdego okresu osobno (zbiór danych tworzą jednostki obserwowane w ustalonym okresie – **dane przekrojowe**)
2. Regresja liniowa dla jednostek (**dane przekrojowe**) i dla średnich wartości zmiennych w *T* okresach
3. Regresje liniowe dla każdej jednostki osobno (**szereg czasowy**), gdy *T* jest duże
4. „Zwykła" regresja dla wszystkich obserwacji (*N* jednostek w *T* okresach) – z tzw. wspólnym wyrazem wolnym

12

Liniowe modele dla danych panelowych

Założenia – prawie identyczne z KM(N)RL*

1. Liniowość y = „X·beta" + epsilon („" - różnice)
2. X – nielosowe (albo losowe i cov(X, epsilon)=0)
3. Rząd(X) = liczbie kolumn
4. E(epsilon) = 0
5. V(epsilon) = „sigma^2 · I"

Albo

5. UM(N)RL** → V(epsilon) = „Ω"

*Klasyczny model (normalnej) regresji liniowej, U** - Uogólniony ...

13

Model ze wspólnym wyrazem wolnym (ang. pooled regression)

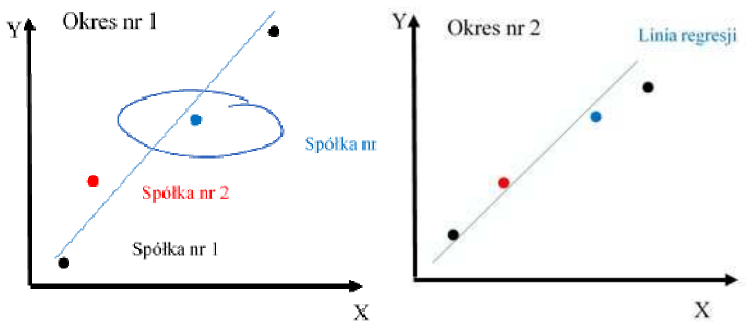
- Postać (i – nr obiektu, t – okres)

$$\underset{1\times1}{y_{it}} = \underset{1\times1}{\beta_0} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + \underset{1\times1}{v_{it}} \quad \text{dla} \quad i = 1,\ldots,N \quad t = 1,\ldots,T$$

- Stosujemy estymator MNK (dla K+1 parametrów).
- Nie korzysta z charakteru danych, nie uwzględnia heterogeniczności.
- Punkt „odniesienia" dla budowy właściwych modeli danych panelowych.

14

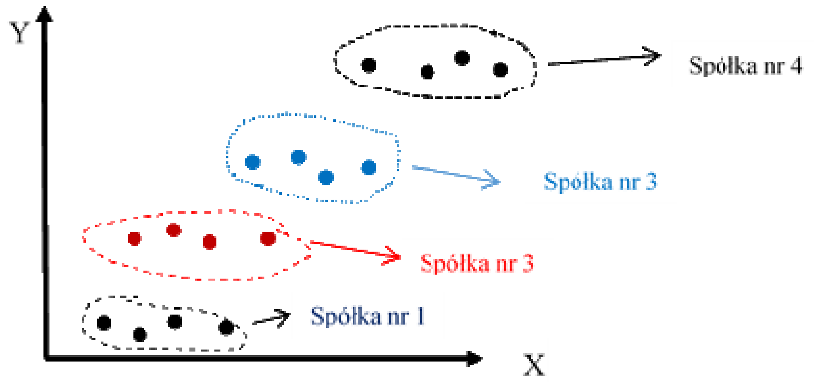
Gdy nie stosujemy modele dla danych panelowych, to …



- Pytanie: czy na poziomie podmiotów (a podejmują one decyzje niezależnie) zmiana czynnika X wpływa na zmiany Y?

16

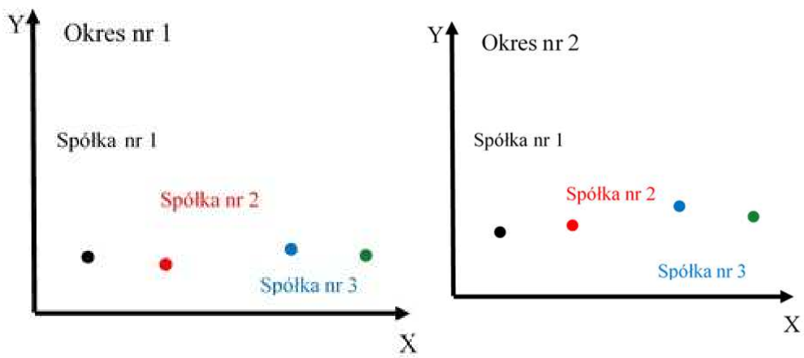
Gdy nie stosujemy modele dla danych panelowych, to …



- Ponawiamy pytanie, gdy mamy wszystkie obserwacje razem. Czy zmiana X wywołuje zmiany Y?

17

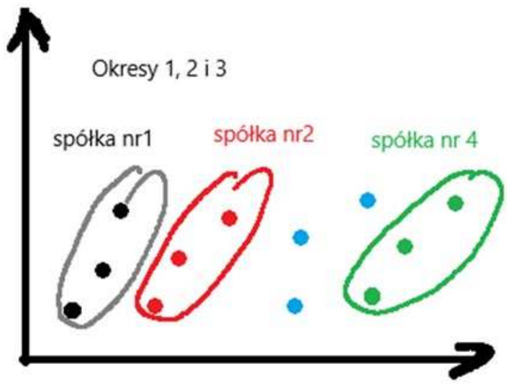
Przypadek nr 2



- Pytanie: czy na poziomie podmiotów (a podejmują one decyzje niezależnie) zmiana czynnika X wpływa na zmiany Y?

18

Przypadek nr 2 …



- Ponawiamy pytanie, gdy mamy wszystkie obserwacje razem. Czy zmiana X wywołuje zmiany Y?

19

MDP – ogólna postać modelu (z dwoma efektami)

$$\underset{1\times1}{y_{it}} = \underset{1\times1}{\alpha_i} + \underset{1\times1}{\lambda_t} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + v_{it} \quad \text{dla } i=1,\dots,N \quad t=1,\dots,T$$

- α_i - odzwierciedla nieobserwowany i nieuwzględniony w równaniu efekt indywidualny,
- λ_t - odzwierciedla nieobserwowany i nieuwzględniony efekt swoisty dla danego okresu; tzw. efekt czasowy,
- v_{it} – składnik „czysto” losowy; zmienna losowa,
- W ogólnym przypadku składniki te odzwierciedlają efekt wynikający z przynależności do danej grupy (tzw. efekt grupowy); przykład dane przekrojowe o przedsiębiorstwach z różnych branż. Czasem (ale rzadko) w równaniu występują trzy efekty albo więcej.

20

Rola efektu indywidualnego (α_i)

Odzwierciedla (kontroluje) wpływ czynników ukrytych:

- bonitacji gleby, doświadczenia lub organizacji pracy rolnika na wielkość plonów,
- aktywności (zdolności) samorządów (powiatów, gmina) na pozyskiwanie środków zewnętrznych (np. z UE),
- zarzadzania w przedsiębiorstwie na wielkość produkcji, kosztu i zysku.
- polityki spółek na zakres ujawnianych informacji w sprawozdaniach finansowych,
- w międzynarodowych analizach porównawczych krajów: wpływ ich specyficzności na wybrane wskaźniki ekonomiczne,
- zatrudnienie kobiet: wpływ wynagrodzenia męża i dochodu niezwiązanego z pracą na aktywność zawodową żony,
- indywidualnych zdolności pracownika na zróżnicowanie płacy,
- przywiązanie klienta do marki.

21

Rola efektów czasowych (λ_t)

Nieobserwowalny efekt czasowy reprezentuje np.:

- wpływ strajków na spadek produkcji,
- efekt niedostatecznej podaży surowców (np. energetycznych) na ich ceny i PKB,
- wpływ polityki rządu (np. podatku ekologicznego, stawek celnych, wysokości opłat rejestracyjnych itp. na ceny i podaż aut nowych i importowanych),
- wpływ referencyjnej stopy procentowej (NBP) na „politykę” kredytową banków,
- warunki pogodowe $\rightarrow$ wpływ na plony upraw polowych,
- zmiany organizacji pracy lub postępu technologicznego na wielkość produkcji,
- zmiany przyzwyczajeń konsumentów.

22

Modele danych panelowych

- Jak traktować α_i i λ_t ?
  - **Nieznane parametry** - Model z **efektami stałymi**, *ang. Fixed Effects Model*
  - **Zmienne losowe** - Model z **efektami losowymi**, *Random Effects Model*
- Czy uwzględniać oba jednocześnie w modelu?
  - **Jeden** - Model jednoczynnikowy (jednokierunkowy); *The One-way Error Component Regression Model*
  - **Oba** - Model dwuczynnikowy (dwukierunkowy); *The Two-way Error Component Regression Model*
- Testy statystyczne – wybór specyfikacji

23

Modele danych panelowych

- Modele jednoczynnikowe
  - Model ze stałymi (losowymi) efektami indywidualnymi

$$y_{it} = \underset{1\times 1}{\alpha_i} + \underset{1\times K}{x_{it}} \cdot \underset{K\times 1}{\beta} + v_{it} \qquad \text{dla } i = 1,\ldots,N \quad t = 1,\ldots,T$$

  - Model ze stałymi (losowych) efektami czasowymi

$$y_{it} = \underset{1\times 1}{\lambda_t} + \underset{1\times K}{x_{it}} \cdot \underset{K\times 1}{\beta} + v_{it} \qquad \text{dla } i = 1,\ldots,N \quad t = 1,\ldots,T$$

gdzie α_i i λ_t - traktowane jako nieznane stałe, czyli parametry (zmienne losowe), które podlegają estymacji (opisujące rozkład skł. losowego).

24

Modele danych panelowych

- Modele dwuczynnikowe
  - Model z indywidualnymi i czasowymi efektami stałymi (ustalonymi)

$$y_{it} = \underset{1\times1}{\alpha_i} + \underset{1\times1}{\lambda_t} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + v_{it} \quad \text{dla } i = 1,\ldots,N \quad t = 1,\ldots,T$$

  - Model z indywidualnymi i czasowymi efektami losowych

$$y_{it} = \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + \varepsilon_{it} \quad \text{gdzie} \quad \varepsilon_{it} = \alpha_i + \lambda_t + v_{it}$$

gdzie α_i i λ_t - traktowane jako zmienne losowe.

25

Modele z efektami czasowymi

$$y_{it} = \underset{1\times1}{\lambda_t} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + v_{it} \qquad (i = 1,\ldots,N,\ t = 1,\ldots,T)$$

Dlaczego model z trendem nie jest alternatywą?

$$y_{it} = \underset{1\times1}{\beta_0} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + \gamma \cdot t + v_{it}$$

- Parametr γ jest ustalonego znaku (trend rosnący albo malejący), interpretacja: wpływ przeciętny.
- Natomiast efekty czasowe uwzględniają specyficzne dla każdego okresu działanie czynników zewnętrznych (pogoda, koniunktura, polityka państwa itp.), λ_t mogą mieć znaki na przemian +/-.

26

Model z indywidualnymi efektami stałymi – konstrukcja

- Różnice między obiektami (ze względu na y_{it}) są modelowane przez różnice w wyrazie wolnym (pojawiają się „swoiste" wyrazy wolne)

$$\underset{1\times 1}{y_{it}} = \underset{1\times 1}{\alpha_i} + \underset{1\times K}{x_{it}} \cdot \underset{K\times 1}{\beta} + \underset{1\times 1}{v_{it}} \qquad \text{dla } i = 1,\ldots,N \quad t = 1,\ldots,T$$

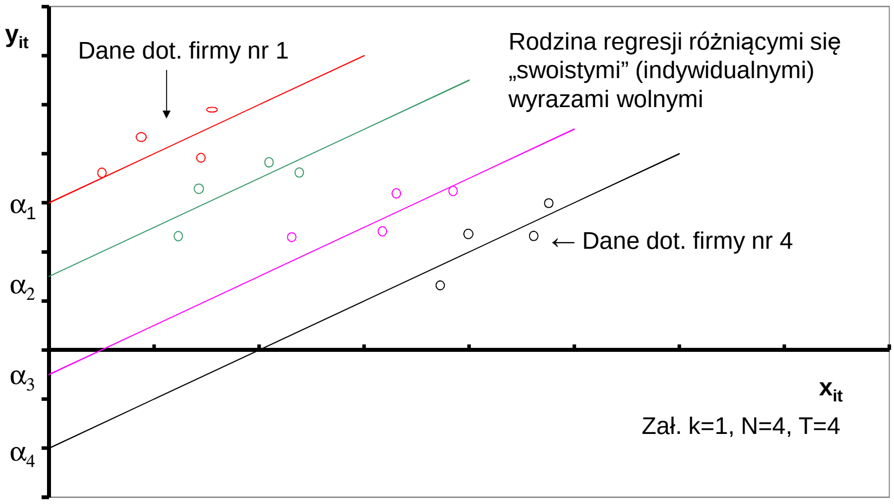
- Zapis dla *i*-tego obiektu

$$\iota_T = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{T\times 1}$$

$$\underset{T\times 1}{y_i} = \underset{T\times 1}{\iota_T} \cdot \alpha_i + \underset{T\times K}{X_i} \cdot \beta + \underset{T\times 1}{v_i} \qquad \text{dla} \quad i = 1,\ldots,N$$

28

Model z efektami stałymi



- Źródło: Arellano M. [2003]

29

Model z efektami stałymi – założenia

$$\underset{T\times 1}{y_i} = \underset{T\times 1}{\iota} \cdot \underset{1\times 1}{\alpha_i} + \underset{T\times K}{X_i} \cdot \underset{K\times 1}{\beta} + \underset{T\times 1}{v_i} \quad \text{dla} \quad i = 1,\ldots,N$$

- Założenia stochastyczne

$$A1: \quad E[v_i | X_i, \alpha_i] = \underset{T\times 1}{0}$$

$$A2: \quad V[v_i | X_i, \alpha_i] = E[v_i \cdot v_i' | X_i, \alpha_i] = \sigma_v^2 \cdot I_T$$

$$A3: \quad E[v_i \cdot v_j' | X_i, \alpha_i] = \underset{T\times T}{0} \ \text{dla} \ j \neq i$$

A4: X_i nielosowe, rząd (X_i)=K

30

Model z efektami stałymi – konstrukcja

- Jak uwzględnić wyraz wolny β_0 ?

1. Ukryć go w parametrach α_i

$$y_{it} = \alpha_i + x_{it} \cdot \beta + v_{it} \quad \Rightarrow y_{it} = \underbrace{\alpha_i^* + \beta_0}_{\alpha_i} + x_{it} \cdot \beta + v_{it}$$

2. Uwzględnić go w wektorze β i dodać „1” do X

$$y_{it} = \alpha_i^* + [1 \quad x_{it}] \cdot \begin{bmatrix} \beta_0 \\ \beta \end{bmatrix} + v_{it}$$

$$\text{gdzie} \quad \alpha_i^* = \alpha_i - \beta_0$$

$$\text{przy warunku} \sum\nolimits_i^N \alpha_i^* = 0$$

$$\beta_0 = \frac{1}{N} \sum_{i=1}^{N} \alpha_i$$

α_i^* to **tzw. kontrast**

Identyfikowalność

31

Model z efektami stałymi – estymacja

- Regresja liniowa z N+K parametrami, wymiar macierzy [Z X] = NT×(N+K)

$$\underset{T\times 1}{y_i} = \underset{T\times 1}{\iota_T} \cdot \underset{1\times 1}{\alpha_i} + \underset{T\times K}{X_i} \cdot \underset{K\times 1}{\beta} + \underset{T\times 1}{v_i} \quad \text{dla} \quad i = 1,\ldots,N$$

$$\underset{NT\times 1}{y} = \underset{NT\times N}{Z} \cdot \underset{N\times 1}{\alpha} + \underset{NT\times K}{X} \cdot \beta + \underset{NT\times 1}{v} = \begin{bmatrix} Z & X \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \underset{NT\times 1}{v}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \iota_T & 0 & \ldots & 0 \\ 0 & \iota_T & & \\ \vdots & & & \\ 0 & \ldots & & \iota_T \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \cdot \beta + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}$$

sztuczne zmienne 0-1

32

Model efektów stałych – estymacja

- Własności macierzy Z

$$\underset{NT\times N}{Z} = I_N \otimes \iota_T \qquad (Z' \cdot Z)^{-1} = T^{-1} \cdot I_N$$

$$Z \cdot Z' = I_N \otimes J_T \qquad Z \cdot \underset{NT\times NT}{(Z' \cdot Z)^{-1} \cdot Z'} = I_N \otimes \bar{J}_T$$

$$J_T = \iota_T \cdot \iota_T' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & & \\ \vdots & & \ddots & \\ 1 & & & 1 \end{bmatrix}_{T\times T} \qquad \bar{J}_T = \frac{1}{T} J_T$$

33

Model efektów stałych – estymacja (podejścia równoważne)

- Efektywny sposób zastosowania MNK?

$$\min_{\alpha,\beta} \sum_{i=1}^{N} v_i' \cdot v_i = \min_{\alpha,\beta} \sum_{i=1}^{N} \left(y_i - \iota \cdot \alpha_i - X_i \cdot \beta\right)' \left(y_i - \iota \cdot \alpha - X_i \cdot \beta\right)$$

- Twierdzenie Frischa-Waugha

$$y = X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + v \qquad y - X_2 \cdot \beta_2 = X_1 \cdot \beta_1 + v$$

- Transformacja zmiennych; Q = I_{NT} - Z(Z'Z)$^{-1}$Z'

$$Q \cdot \underset{NT\times 1}{y} = Q \cdot \underset{NT\times N}{Z} \cdot \underset{N\times 1}{\alpha} + Q \cdot \underset{NT\times K}{X} \cdot \underset{K\times 1}{\beta} + Q \cdot \underset{NT\times 1}{v}$$

34

Uwagi o Uogólnionej MNK

Rozważamy uogólniony model regresji liniowej

$$y = X\beta + \varepsilon \qquad gdzie \quad V(\varepsilon) = \Omega$$

Ω jest macierzą symetryczną, dodatnio określoną.
Zatem istnieje taka macierzy $\Omega^{-1/2}$, że $\Omega^{-1/2} \cdot \Omega^{-1/2} = \Omega^{-1}$

Estymator UMNK ma postać (Ω - znana)

$$\hat{\beta}_{UMNK} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}y, \quad V\left(\hat{\beta}_{UMNK}\right) = \left(X'\Omega^{-1}X\right)^{-1}$$

lub równoważnie (dla transformowanych danych)

$$\hat{\beta}_{UMNK} = \left(X^{*}{}'X^{*}\right)^{-1} X^{*}{}'y^{*}, \quad V\left(\hat{\beta}_{UMNK}\right) = \left(X^{*}{}'X^{*}\right)^{-1}, \quad gdzie$$

$$X^{*} = \Omega^{-\frac{1}{2}} \cdot X \qquad y^{*} = \Omega^{-\frac{1}{2}} \cdot y$$

35

Model efektów stałych – estymacja

- Twierdzenie Frischa-Waugha

$$y = X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + v$$

$$\begin{cases} b_2 = \left(X_2' \cdot Q \cdot X_2\right)^{-1} X_2' \cdot Q \cdot y \\ b_1 = \left(X_1' \cdot X_1\right)^{-1} X_1' \cdot \left(y - X_2 \cdot b_2\right) \end{cases}$$

gdzie $Q = I - X_1\left(X_1' \cdot X_1\right)^{-1} X_1' = I - P$

$$Q = Q'$$

$$Q \cdot Q = Q$$

$$Q \cdot X_1 = 0 \qquad P \cdot Q = 0$$

36

Model efektów stałych – zastosowanie twierdzenia Frischa-Waugha

- Q, Q_0 – macierze symetryczne i idempotentne
- det(Q)=0 rząd(Q)=N (liczbie kolumn Z)

$$\underset{NT \times NT}{Q} = I_{NT} - Z \cdot \left(Z' \cdot Z\right)^{-1} \cdot Z' = I_N \otimes \underset{T \times T}{Q_0}$$

$$Q_0 = I_T - T^{-1} \cdot J_T \quad \det\left(Q_0\right) = 0 \quad rzad\left(Q_0\right) = ślad\left(Q_0\right) = T - 1$$

$$Q = \begin{bmatrix} Q_0 & 0 & \cdots & \\ 0 & Q_0 & & \\ \vdots & & \ddots & \\ 0 & & & Q_0 \end{bmatrix}$$

37

Model efektów stałych – estymacja

- Estymator WewnątrzGrupowy, kowariancyjny, efektów stałych, (within-group estimator, covariance estimator, LSDV estimator – least squares dummy-variable estimator)

$$b_{WG} = (X' \cdot Q \cdot X)^{-1}[X' \cdot Q \cdot y]$$

$$b_{WG} = \left(\sum_{i=1}^{N} X_i' \cdot Q_0 \cdot X_i\right)^{-1} \cdot \left[\sum_{i=1}^{N} X_i' \cdot Q_0 \cdot y_i\right]$$

$$Q_0 = I_T - \frac{1}{T} \cdot \iota \cdot \iota' = I_T - \frac{1}{T} J_T \qquad Q_0 = Q_0 \cdot Q_0$$

  - Q_0 – Operator wewnątrzgrupowy; operator odchyleń od średniej po czasie, macierz idempotentna

38

Model efektów stałych – estymacja

- Q_0 – Operator wewnątrzgrupowy; operator odchyleń od średniej po czasie

$$\underset{T\times T}{Q_0} \cdot \underset{T\times 1}{y_i} = \left(I_T - \frac{1}{T} J_T\right) \cdot y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix} - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} y_{it} \\ \vdots \\ \frac{1}{T}\sum_{t=1}^{T} y_{it} \end{bmatrix} = y_i - \iota_T \cdot \bar{y}_i$$

$$\underset{T\times T}{Q_0} \cdot \underset{T\times K}{X_i} = X_i - \underset{T\times 1}{\iota_T} \cdot \underset{1\times K}{\bar{x}_i} \quad \text{gdzie} \quad \bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it} \qquad Q_0 \cdot \iota_T = \underset{T\times 1}{0}$$

39

Model efektów stałych – równoważny sposób estymacji

- Regresja dla odchyleń zmiennych od średnich po t:

$$y_{it}^{*} = x_{it}^{*} \cdot \beta + v_{it}^{*} \qquad \text{gdzie}$$

$$y_{it}^{*} = y_{it} - \bar{y}_i \quad x_{it}^{*} = x_{it} - \bar{x}_i \quad \bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it} \quad \bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$$

$$b_{WG} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T} (x_{it} - \bar{x}_i)' \cdot (x_{it} - \bar{x}_i) \right)^{-1} \cdot \sum_{i=1}^{N}\sum_{t=1}^{T} (x_{it} - \bar{x}_i)' \cdot (y_{it} - \bar{y}_i)$$

$$b_{WG} = \left(X^{*'} X^{*} \right)^{-1} X^{*'} y^{*} \quad gdzie \quad X^{*} = QX$$

40

Model efektów stałych – estymacja

- Estymator parametrów α_i

$$a = (Z' \cdot Z)^{-1} \cdot Z' \cdot (y - X \cdot b)$$

$$a_i = \bar{y}_i - \bar{x}_i \cdot b_{WG} \quad \text{dla} \quad i = 1, \ldots, N$$

- Estymator β_0 (wspólnego wyrazu wolnego)

$$b_0 = \frac{1}{N}\sum_{i=1}^{N} a_i = \bar{y} - \bar{x} \cdot b_{WG}$$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} \bar{x}_i = \frac{1}{N \cdot T}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it} \qquad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} \bar{y}_i = \frac{1}{N \cdot T}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$$

41

Model efektów stałych – estymacja

- Estymator dla σ^2_v

$$s_v^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - a_i - x_{it}\cdot b_{WG}\right)^2}{N\cdot T - N - K} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - \bar{y}_i - \left(x_{it} - \bar{x}_i\right)\cdot b_{WG}\right)^2}{N\cdot T - N - K}$$

- Macierz kowariancji dla b_{WG}

$$\hat{V}\left(b_{WG}\right) = s_v^2\cdot\left(X'\cdot Q\cdot X\right)^{-1} = s_v^2\cdot\left(\sum_{i=1}^{N} X_i'\cdot Q_0\cdot X_i\right)^{-1}$$

42

Model efektów stałych – estymacja

- Estymator wariancji dla α_i

$$\hat{V}\left(a_i\right) = \frac{s_v^2}{T} + \underset{1\times K}{\bar{x}_i}\cdot\underset{K\times K}{\hat{V}\left(b_{WG}\right)}\cdot\left(\bar{x}_i\right)' \quad \text{dla} \quad i = 1,\ldots,N$$

- Wariancja estymatora parametru β_0

$$\hat{V}\left(b_0\right) = \frac{s_v^2}{N\cdot T} + \underset{1\times K}{\bar{x}}\cdot\underset{K\times K}{\hat{V}\left(b_{WG}\right)}\cdot\left(\bar{x}\right)'$$

43

Testowanie hipotez

Model – układ złożeń odzwierciedlających potencjalne zależności między zmiennymi ekonomicznymi.

Koncepcja budowy modelu – od ogółu do szczegółu (*general to specific modelling*).

Hipotezy:

- restrykcje zerowe dla parametrów (uproszczenie modelu),
- porównanie siły wpływu zm. objaśniających na objaśnianą.

44

Testowanie istotności układu parametrów regresji - koncepcja

Założenie: $y = X\beta + \varepsilon \rightarrow y = X_{(0)}\beta^{(0)} + X_{(1)}\beta^{(1)} + \varepsilon$

$$\underset{[NT\times k]}{X} = \begin{bmatrix} \underset{[NT\times k_0]}{X_{(0)}} & \underset{[NT\times k_1]}{X_{(1)}} \end{bmatrix} \qquad \underset{[k\times 1]}{\beta} = \begin{bmatrix} \beta^{(0)} \\ \beta^{(1)} \end{bmatrix} \begin{matrix} \textit{wektor } [k_0\times 1] \\ \textit{wektor } [k_1\times 1] \end{matrix} \qquad k = k_0 + k_1$$

$$H_0\colon \beta^{(1)} = 0_{[k_1\times 1]} \quad \Rightarrow \quad y = X_{(0)}\beta^{(0)} + \varepsilon$$

$$H_1\colon \beta^{(1)} \neq 0_{[k_1\times 1]} \quad \Rightarrow \quad y = X\beta + \varepsilon$$

W dalszej części k_1 (liczbę restrykcji narzucanych na parametry w H_0) oznaczmy jako J.

45

Testowanie istotności układu parametrów regresji

- Test Chowa: statystyka testowa

$$F_{emp} = \frac{\left(SKR_{H0} - SKR_{H1}\right)/J}{SKR_{H1}/df} \sim F(J, df)$$

- Wartość krytyczna testu $F_{\alpha}(J, df)$

$$\text{Pr-o}\left\{F_{emp} > F_{\alpha}(J, df)\right\} = \alpha$$

- Decyzja: Jeżeli $F_{emp} > F_{\alpha}(J, df)$, to odrzucamy H_0 na rzecz H_1, a w przeciwnym razie nie ma podstaw do odrzucenia H_0 (na poziomie α).

Wyjaśnienie: SKR_{H0} = suma kwadratów reszt w modelu dla H0,
SKR_{H1} = suma kwadratów reszt w modelu dla H1.
df = stopnie swobody w modelu dla H1 *(np. NT–k–N)*

46

Testy alternatywne

- Testowanie restrykcji na parametry – test Walda w modelu regresji y = Xβ + ε

$$H_0 : \underset{J\times K}{R} \cdot \underset{K\times 1}{\beta} = \underset{J\times 1}{r}$$

$$W = \left(R\hat{\beta} - r\right)'\left(R \cdot V\left(\hat{\beta}\right) \cdot R'\right)^{-1}\left(R\hat{\beta} - r\right) \sim \chi^2(J)$$

Dla małej próby (nieznane σ^2 zastępujemy s^2)

$$\frac{W}{J} \sim F(J, df)$$

47

Model z efektami stałymi – testowanie efektów stałych kontra „zwykła regresja”

- Test Chowa

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_{N-1} = 0 \qquad y_{it} = \underbrace{\beta_0}_{\alpha_N} + x_{it} \cdot \beta + v_{it}$$

H1 – model panelowy z efektami stałymi.

Statystyka testowa:

$$F_{emp} = \frac{(SKR_{H0} - SKR_{H1})/(N-1)}{SKR_{H1}/(NT-N-K)} \overset{H_0}{\sim} F(N-1, NT-N-K)$$

gdzie SKR_{H0} (SKR_{H1}) = suma kwadratów w modelu z restrykcjami (bez restrykcji); zob. licznik we wzorze dla s^2_v.

48

Model efektów stałych – dopuszczalność MNK

- Jeżeli odrzucamy H_0, to estymator MNK w

$$y_{it} = \underbrace{\beta_0}_{\alpha_N} + x_{it} \cdot \beta + v_{it}$$

jest estymatorem niezgodnym i obciążonym.

49

Model efektów stałych – testowanie efektów stałych kontra „zwykła regresja”

Inny zapis statystyki testowej:

$$\varphi^2 = RSS \Big/ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(y_{it} - \bar{y}\right)^2 = 1 - R^2$$

$$(RRSS - URSS)/URSS = \left(\varphi^2_{RRSS} - \varphi^2_{URSS}\right) / \varphi^2_{URSS} =$$

$$\left(1 - R^2_{RRSS} - 1 + R^2_{URSS}\right) / \left(1 - R^2_{URSS}\right) = \left(R^2_{URSS} - R^2_{RRSS}\right) / \left(1 - R^2_{URSS}\right)$$

$$F_{emp} = \frac{\left(R^2_{URSS} - R^2_{RRSS}\right)/(N-1)}{\left(1 - R^2_{URSS}\right)/(NT - N - K)}$$

gdzie R^2 (ϕ^2) współczynnik determinacji (zbieżności) oraz

(URSS) RRSS (*Un)Restricted Residual Sum of Squares* = suma kwadratów w modelu H0 z restrykcjami (H1→ bez restrykcji).

50

Przypomnienie: R^2 w modelu regresji

- R^2 – względny udziału roli zmiennych objaśniających w wyjaśnieniu zmienności Y:

$$\underset{1\times1}{y_i} = \underset{1\times1}{\beta_0} + \underset{1\times K}{x_i} \cdot \underset{K\times1}{\beta} + \underset{1\times1}{v_i} \quad \text{dla } i = 1, \ldots, N$$

w relacji do modelu referencyjnego:

$$\underset{1\times1}{y_i} = \underset{1\times1}{\beta_0} + \underset{1\times1}{v_i} \quad \text{dla } i = 1, \ldots, N$$

Inaczej: względny przyrost dopasowania do danych modelu (z X-sami) w odniesieniu do modelu referencyjnego (bez X-sów).

Opis dla danych przekrojowych (pominięto indeks *t*).

51

Model z efektami stałymi – mierniki R^2

- **R^2_{LSDV}** - miernik w modelu $\underset{1\times1}{y_{it}} = \underset{1\times1}{\alpha_i} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + \underset{1\times1}{v_{it}}$

  w relacji do modelu $\underset{1\times1}{y_{it}} = \underset{1\times1}{\beta_0} + \underset{1\times1}{v_{it}}$ (np. $\beta_0 = \alpha_N$)
- R^2_{LSDV} - przyjmuje bardzo wysokie wartości (np. ponad 0,9), liczne sztuczne zmienne zero-jedynkowe odpowiadające efektom stałym α_i.
- **R^2_{Within}** - miernik w modelu $\underset{1\times1}{y_{it}} = \underset{1\times1}{\alpha_i} + \underset{1\times K}{x_{it}} \cdot \underset{K\times1}{\beta} + \underset{1\times1}{v_{it}}$

  w relacji do modelu $\underset{1\times1}{y_{it}} = \underset{1\times1}{\alpha_i} + \underset{1\times1}{v_{it}}$
- R^2_{Within} - przyjmuje wartości < R^2_{LSDV}

52

Model efektów stałych – własności estymatorów

Własności estymatorów b_{WG} i a (spełnione są założenia modelu Gaussa-Markowa)

- dla β: zgodny, gdy T→∞ lub N→∞
- dla α: zgodny, gdy T→∞
- b_{WG} ma asymptotyczny rozkład normalny

53

Modele efektów stałych – wady i zalety

Wady

- Niemożność (dokładna współliniowość) estymowania
  - efektów indywidualnych, gdy występują zmienne egzogeniczne stałe po czasie (płeć, rasa, wyznanie itp.)
  - efektów czasowych, gdy występują zmienne egzogeniczne stałe po obiektach (ceny, oprocentowanie itp.)
- Zwykle T jest niewielkie, więc estymator dla efektów indywidualnych („a”) nie jest estymatorem zgodnym
- Wymaga wprowadzenia N-1 dodatkowych parametrów, co zmniejsza liczbę stopni swobody

Zalety

- Dopuszcza się założenie o skorelowaniu x_{it} z α_i (wtedy b_{WG} wciąż jest zgodny i nieobciążony)

54