

1 Wielowskaźnikowe modele regresji liniowej

G -równaniowy ekonometryczny model regresji wielorakiej - postać uogólniona.

1.1 G -równaniowy model regresji wielorakiej (bez restrykcji zerowych):

$$\begin{aligned} t_{t1} &= \gamma_{12}t_{t2} + \dots + \gamma_{1G}y_{tG} + \beta_{10} + \beta_{11}x_{t1} + \beta_{12}x_{t2} + \dots + \beta_{1k}x_{tk} + \varepsilon_{t1} \\ t_{t1} &= \gamma_{21}t_{t1} + \gamma_{23}t_{t3} + \dots + \gamma_{2G}y_{tG} + \beta_{20} + \beta_{21}x_{t1} + \beta_{22}x_{t2} + \dots + \beta_{2k}x_{tk} + \varepsilon_{t2} \\ &\vdots \\ t_{tG} &= \gamma_{G1}t_{t1} + \gamma_{G2}t_{t2} + \dots + \beta_{G0} + \beta_{G1}x_{t1} + \beta_{G2}x_{t2} + \dots + \beta_{Gk}x_{tk} + \varepsilon_{tG} \end{aligned}$$

1.1.1 SYNTETYCZNY ZAPIS MACIERZOWY MODELU

$$y_t \cdot \Gamma + x_t \cdot B = \varepsilon_t$$

gdzie:

$y_t = \begin{bmatrix} y_{t1} & y_{t2} & \dots & y_{tG} \end{bmatrix}$ - wektor zmiennych endogenicznych modelu

$\Gamma = \begin{bmatrix} \gamma_{G \times G} \end{bmatrix}$ - macierz parametrów strukturalnych zmiennych endogenicznych

$x_t = \begin{bmatrix} 1 & x_{t1} & x_{t2} & \dots & x_{tk} \end{bmatrix}$ - wektor zmiennych egzogenicznych modelu

$B = \begin{bmatrix} \beta_{(k+1) \times G} \end{bmatrix}$ - macierz parametrów strukturalnych zmiennych egzogenicznych

$\varepsilon_t = \begin{bmatrix} \varepsilon_{t1} & \varepsilon_{t2} & \dots & \varepsilon_{tG} \end{bmatrix}$ - wektor składników zakłócających modelu

$$Y \cdot \Gamma + X \cdot B = E$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1G} \\ y_{21} & y_{22} & \dots & y_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{nG} \end{bmatrix} = Y_{n \times G} \text{ - macierz obserwacji endogenicznych modelu}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} = X_{n \times (k+1)} \text{ - macierz obserwacji zmiennych egzogenicznych modelu}$$

$$E = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \dots & \varepsilon_{1G} \\ \varepsilon_{21} & \varepsilon_{22} & \dots & \varepsilon_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \dots & \varepsilon_{nG} \end{bmatrix} = E_{n \times G} - \text{macierz składników zakłócających } G\text{-równaniowego modelu}$$

1.2 Wstępne założenia dotyczące struktury stochastycznej modelu wielorównaniowego.

$$\begin{aligned} \mathbb{E}\varepsilon_{tj} &= 0 & t &= 1, 2, \dots, n & j &= 1, 2, \dots, G \\ \mathbb{E}\varepsilon &= 0 \end{aligned}$$

co oznacza, że składnik zakłócający jest zmienną losową o wartości oczekiwanej równe 0, w każdym okresie t i w każdym równaniu j -tym.

$$\mathbb{E}\varepsilon_{tj}^2 = \sigma_{\varepsilon j}^2 = \text{const.} \quad j = 1, 2, \dots, G$$

co oznacza, że w każdym równaniu j -tym wariancja składnika losowego jest stała i niezależna od okresu obserwacji.

$$\mathbb{E}\varepsilon_{tj}\varepsilon_{(t-s)j} = 0 \quad , (s \neq 0)$$

co oznacza, że kowariancja między składnikami losowymi w poszczególnych równaniach dla różnych okresów jest równa 0. Klasyfikacja modeli z punktu widzenia powiązań pomiędzy zmiennymi endogenicznymi.

1.2.1 MODEL PROSTY

Model prosty jest to model, w którym nieopóźnione zmienne endogeniczne nie oddziałują na siebie. Z uwagi na brak powiązań pomiędzy nieopóźnionymi zmiennymi endogenicznymi macierz Γ jest macierz diagonalną, będąc macierzą jednostkową.

Przykład trójrównaniowego modelu prostego:

$$\begin{aligned} y_{t1} &= \beta_{10} + \beta_{11}x_{t1} + \beta_{12}x_{t2} - \varepsilon_{t1} \\ y_{t2} &= \beta_{20} + \beta_{21}x_{t2} + \beta_{23}x_{t3} - \varepsilon_{t2} \\ y_{t3} &= \beta_{30} + \beta_{32}x_{t1} + \beta_{33}x_{t3} - \varepsilon_{t3} \end{aligned}$$

Zapis w postaci macierzowej

$$Y \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + X \cdot \begin{bmatrix} -\beta_{10} & -\beta_{20} & -\beta_{30} \\ -\beta_{11} & -\beta_{21} & 0 \\ -\beta_{12} & 0 & -\beta_{32} \\ 0 & -\beta_{23} & -\beta_{33} \end{bmatrix} = E$$

Wstępne założenia stochastyczne dla modelu prostego:
Zakładając nielosowość zmiennych z góry ustalonych (egzogenicznych) musimy uznać, że zmienne egzogeniczne są niezależne od składników losowych, tzn.

$$\mathbb{E}(X'_j \cdot \varepsilon_j) = (\mathbb{E}X'_j) \cdot (\mathbb{E}\varepsilon_j) = X'_j \cdot 0 = 0$$

1.2.2 MODEL REKURENCYJNY

Model rekurencyjny jest to model, w którym występują jednokierunkowe powiązania pomiędzy zmiennymi endogenicznymi nieopóźnionymi w czasie. Oznacza to, że macierz Γ - parametrów występujących przy zmiennych endogenicznych - jest macierzą trójkątną.

Przykład trójrównaniowego modelu rekurencyjnego:

$$Y \cdot \begin{bmatrix} 1 & 0 & 0 \\ -y_{12} & 1 & 0 \\ -y_{13} & -y_{23} & 1 \end{bmatrix} + X \cdot \begin{bmatrix} -\beta_{10} & -\beta_{20} & -\beta_{30} \\ -\beta_{11} & -\beta_{21} & 0 \\ -\beta_{12} & 0 & -\beta_{32} \\ 0 & -\beta_{23} & -\beta_{33} \end{bmatrix} = E$$

Wniosek

W modelach rekurencyjnych istnieje zawsze takie równanie, w którym występują jedynie zmienne z góry ustalone (egzogeniczne).

W przypadku pozostałych równań musimy uznać, że w zbiorze zmiennych objaśniających występują zmienne losowe, którymi są zmienne endogeniczne występujące w charakterze zmiennych objaśniających. Zmienne endogeniczne występujące w charakterze zmiennych objaśniających mogą być uznane za:

- nieskorelowane ze składnikiem zakłócającym danego równania, jeśli konstrukcja modelu nie wymusza odrzucenia takiego założenia
- skorelowane ze składnikiem zakłócającym danego równania, jeśli wynika to z założeń tkwiących u podstaw konstrukcji modelu.

1.3 Model o równaniach współzależnych

Model o równaniach współzależnych jest to model, w którym występują nieopóźnione w czasie sprzężenia zwrotne pomiędzy zmiennymi endogenicznymi. Oznacza to, że macierz Γ nie jest macierzą ani diagonalną, ani trójkątną.

Przykład trójrównaniowego modelu regresji wielorakiej:

$$\begin{aligned} y_{t1} &= \gamma_{12}y_{t2} + \gamma_{13}y_{t3} + \beta_{10} + \beta_{11}x_{t1} + \beta_{12}x_{t2} + \varepsilon_{t1} \\ y_{t2} &= \gamma_{23}y_{t3} + \beta_{20} + \beta_{21}x_{t1} + \beta_{23}x_{t3} + \varepsilon_{t2} \\ y_{t3} &= \gamma_{31}y_{t1} + \beta_{30} + \beta_{33}x_{t3} + \varepsilon_{t3} \end{aligned}$$

$$Y \cdot \begin{bmatrix} 1 & 0 & 0 \\ -y_{12} & 1 & 0 \\ -y_{13} & -y_{23} & 1 \end{bmatrix} + X \cdot \begin{bmatrix} -\beta_{10} & -\beta_{20} & -\beta_{30} \\ -\beta_{11} & -\beta_{21} & 0 \\ -\beta_{12} & 0 & -\beta_{32} \\ 0 & -\beta_{23} & -\beta_{33} \end{bmatrix} = E$$

Uporządkowany całościowy zapis modelu z restrykcjami zerowymi

$$y_{t1} - \gamma_{12}$$

$$\begin{aligned} & \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -\gamma_{31} \\ -y_{12} & 1 & 0 \\ -y_{13} & -y_{23} & 1 \end{bmatrix} + \\ & + \begin{bmatrix} 1 & x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} -\beta_{10} & -\beta_{20} & -\beta_{30} \\ -\beta_{11} & -\beta_{21} & 0 \\ -\beta_{12} & 0 & 0 \\ 0 & -\beta_{23} & -\beta_{33} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \end{bmatrix} \end{aligned}$$

Z uwagi na to, że macierz Γ nie jest ani diagonalna, ani trójkątna, model uznajemy za model o równaniach współzależnych. Oznacza to, że między zmiennymi endogenicznymi nieopóźnionymi w czasie występują sprzężenia zwrotne.

Zauważmy, że zmienne endogeniczne, będąc między innymi funkcjami zmiennych losowych (ε), jednocześnie nawzajem się wyjaśniają. Tym samym musimy uznać, że zmienne te - występując w poszczególnych równaniach w charakterze zmiennych objaśniających - są skorelowane ze składnikami losowymi tychże równań.

1.3.1 REDUKCJA MODELI O RÓWNANIACH WSPÓŁZALEŻNYCH

Procedura sprowadzania modelu o równaniach współzależnych do wielorównaniowego modelu prostego nazywamy redukcją. Przykładowa procedura redukcji naszego modelu

$$y_t \cdot \Gamma + x_t \cdot \beta = \varepsilon_t$$

Przekształcona postać modelu

$$y_t \cdot \Gamma = x_t \cdot (-\beta) + \varepsilon_t$$

Zapis macierz w postaci zredukowanej

$$y_t = x_t \cdot \pi + v_t$$

gdzie

$$\pi = -B\Gamma = \begin{bmatrix} \pi_{10} & \pi_{20} & \pi_{30} \\ \pi_{11} & \pi_{21} & \pi_{31} \\ \pi_{12} & \pi_{22} & \pi_{32} \\ \pi_{13} & \pi_{23} & \pi_{33} \end{bmatrix}$$

π - macierz parametrów strukturalnych postaci zredukowanej o wymiarach $(k+1) \times G$ (gdzie $k+1=4$, $G=3$)

Wektor składników zakłócających postaci zredukowanej o wymiarach $G \times G$, $G=3$

$$v_t = \varepsilon_t \Gamma^{-1} =$$

Tym samym model w postaci macierzowej zapiszemy następująco

$$\begin{bmatrix} v_{t1} & v_{t2} & v_{t3} \end{bmatrix} = \begin{bmatrix} 1 & x_{t1} & x_{t2} & x_{t3} \end{bmatrix} \cdot \begin{bmatrix} \pi_{10} & \pi_{20} & \pi_{30} \\ \pi_{11} & \pi_{21} & \pi_{31} \\ \pi_{12} & \pi_{22} & \pi_{32} \\ \pi_{13} & \pi_{23} & \pi_{33} \end{bmatrix} + \begin{bmatrix} v_{t1} & v_{t2} & v_{t3} \end{bmatrix}$$

Co po rozpisaniu przybierze następującą postać

$$\begin{aligned} y_{t1} &= \pi_{10} + x_{t1}\pi_{11} + x_{t2}\pi_{12} + x_{t3}\pi_{13} + v_{t1} \\ y_{t2} &= \pi_{20} + x_{t1}\pi_{21} + x_{t2}\pi_{22} + x_{t3}\pi_{23} + v_{t2} \\ y_{t3} &= \pi_{30} + x_{t1}\pi_{31} + x_{t2}\pi_{32} + x_{t3}\pi_{33} + v_{t3} \end{aligned}$$

Rozpisany zapis macierzowy modelu dla n obserwacji przedstawia się następująco

$$Y = X\Pi + V$$

Zauważmy, że w każdym z równań postaci zredukowanej występują jedynie zmienne z góry ustalone (egzogeniczne) modelu. Jeśli uznamy, że są one nielosowe, mamy prawo wykluczyć ich ewentualną zależność ze składnikami losowymi poszczególnych równań. Wynika z tego, że każde z równań możemy oszacować stosując metodę najmniejszych kwadratów.

1.3.2 IDENTYFIKACJA

Identyfikacja jest ot proces rozpoznawania (identyfikowania) parametrów strukturalnych modelu (elementów macierzy Γ i B) na podstawie parametrów

postaci zredukowanej, a więc na podstawie elementów macierzy Π . Ze zdefiniowania macierzy Π wynikają następujące konsekwencje

$$\Pi = -B\Gamma^{-1} \Leftrightarrow \Pi\Gamma = -B$$

Wykorzystując badany model oraz jego postać zredukowaną i odpowiednio zdefiniowane dla tych modeli macierze π, Γ oraz B drugi człon powyższego wyrażenia zapiszemy następująco

dużo rozpisanych macierzy

Powiemy, że j -ty układ równań:

1. Zawiera więcej parametrów γ i β aniżeli równań, to równanie M_j -te postaci strukturalnej uznajemy za **nieidentyfikowalne**.
2. Zawiera taką samą ilość parametrów γ i β aniżeli równań, to równanie M_j -te postaci strukturalnej uznajemy za **jednoznacznie identyfikowalne**.
3. Zawiera mniej parametrów γ i β aniżeli równań, to równanie M_j -te postaci strukturalnej uznajemy za **niejednoznacznie identyfikowalne**.

W rozpatrywanym przypadku otrzymujemy następujące trzy układy równań. Pierwszy układ równań (a) dla $j = 1$ równania postaci strukturalnej modelu

$$\begin{aligned}\pi_{10} - \pi_{20}\gamma_{12} - \pi_{30}\gamma_{13} &= \beta_{10} \\ \pi_{11} - \pi_{21}\gamma_{12} - \pi_{31}\gamma_{13} &= \beta_{11} \\ \pi_{12} - \pi_{22}\gamma_{12} - \pi_{32}\gamma_{13} &= \beta_{12} \\ \pi_{13} - \pi_{23}\gamma_{12} - \pi_{33}\gamma_{13} &= 0\end{aligned}$$

Z uwagi na fakt, iż liczba poszukiwanych parametrów γ i β wynosi 5 (2 parametry γ oraz 3 parametry β) jest większa od liczby równań, powyższy układ nie ma rozwiązania. Tym samym powiemy, że równanie pierwsze postaci strukturalnej jest nieidentyfikowalne.

Drugi układ równań (b) dla $j = 2$ równania postaci strukturalnej modelu

$$\begin{aligned}\pi_{20} - \pi_{30}\gamma_{23} &= \beta_{20} \\ \pi_{21} - \pi_{31}\gamma_{23} &= \beta_{21} \\ \pi_{22} - \pi_{32}\gamma_{23} &= 0 \\ \pi_{23} - \pi_{33}\gamma_{23} &= \beta_{23}\end{aligned}$$

Z uwagi na fakt, iż liczba poszukiwanych parametrów γ i β wynosi 4 (1 parametr γ i 3 parametry β) jest równa liczbie równań, powyższy układ ma

jednoznaczne rozwiązanie. Tym samym powiemy, że równanie drugie postaci strukturalnej jest jednoznacznie identyfikowalne. Zauważmy, że z równania trzeciego tego układu równań wynika, że

$$\pi_{22} - \gamma_{23}\pi_{32} = 0 \Rightarrow \gamma_{23} = \left(\frac{\pi_{21}}{\pi_{32}} \right)$$

Trzeci układ równań (c) dla $j = 3$ równania postaci strukturalnej modelu

$$\pi_{30} - \pi_{10}\gamma_{31} = \beta_{30}$$

$$\pi_{31} - \pi_{11}\gamma_{31} = 0$$

$$\pi_{32} - \pi_{12}\gamma_{31} = 0$$

$$\pi_{33} - \pi_{13}\gamma_{31} = \beta_{33}$$

Z uwagi na fakt, iż liczba poszukiwanych parametrów γ i β wynosi 3 (1 parametr γ i 2 parametry β) jest mniejsza od liczby równań, powyższy układ nie ma jednoznacznego rozwiązania. Tym samym powiemy, że równanie trzecie postaci strukturalnej jest niejednoznacznie identyfikowalne.

Aby sformułować użyteczne twierdzenia dotyczące identyfikacji modelu o równaniach współzależnych przyjmujemy następujący system oznaczeń:

- k_j - liczba zmiennych z góry ustalonych występujących w j -tym równaniu modelu łącznie z wyrazem wolnym
- K - liczba zmiennych z góry ustalonych występujących w całym modelu łącznie z wyrazem wolnym
- G_j - liczba zmiennych łącznie współzależnych (endogenicznych) występujących w j -tym równaniu
- G_{j-1} - liczba zmiennych łącznie współzależnych (endogenicznych) występujących w j -tym równaniu w charakterze zmiennych objaśniających (bez zmiennej objaśnianej w tym równaniu)
- G - liczba zmiennych łącznie współzależnych zależnych (endogenicznych) występujących w całym modelu (równa liczbie równań modelu)

$$K_j^* = K - k_j$$

$$G_j^* = G - G_j$$

Twierdzenie 1

Warunkiem koniecznym na to, aby równanie j -te było identyfikowalne, jest aby liczba zmiennych z góry ustalonych oraz łącznie współzależnych występujących w równaniu j -tym w charakterze zmiennych objaśniających była mniejsza od liczby zmiennych z góry ustalonych występujących w całym modelu

$$K \geq K_j + G_j - 1$$

Równanie możemy przekształcić w następujący sposób

$$K - k_j \geq G_j - 1$$

Wprowadzając do siebie wyrażenia otrzymujemy

$$K - k_j \geq G - G_j^* - 1$$

a stąd ostatecznie otrzymujemy

$$K_j^* + G_j^* \geq G - 1$$

Wykorzystując powyższe wyrażenie formułujemy następujące twierdzenia

Twierdzenie 2

Warunkiem koniecznym na to, aby równanie j -te postaci strukturalnej modelu było **identyfikowalne**, jest by liczba zmiennych nie występujących w j -tym równaniu była nie mniejsza od liczby równań modelu pomniejszonej o jeden.

Twierdzenie 3

Warunkiem koniecznym na to, aby równanie j -te postaci strukturalnej modelu było **jednoznacznie identyfikowalne**, jest by liczba zmiennych nie występujących w j -tym równaniu była równa liczbie równań modelu pomniejszonej o jeden.

Twierdzenie 4

Warunkiem koniecznym na to, aby równanie j -te postaci strukturalnej modelu było **niejednoznacznie identyfikowalne**, jest by liczba zmiennych nie występujących w j -tym równaniu była większa od liczby równań modelu pomniejszonej o jeden.

Twierdzenie 5

Warunkiem dostatecznym na to, aby równanie j -te postaci strukturalnej modelu było **nieidentyfikowalne**, jest by liczba zmiennych nie występujących w j -tym równaniu była mniejsza od liczby równań modelu pomniejszonej o jeden.

Zadanie

Rozpatrz następujący model o równaniach współzależnych

$$\begin{aligned}Y_t &= q_0 + q_1 L_t + a_2 k_t + a_3 H_t + a_4 ER_t + e_{t1} \\L_t &= b_0 + b_1 Y_t + b_2 K_t + b_3 W_t e_{t2} \\K_t &= c_0 + c_1 Y_t + c_2 R_t + e_{t3}\end{aligned}$$

1.3.3 METODA ZMIENNYCH INSTRUMENTALNYCH

W dotychczas przedstawianych modelach przyjmowaliśmy założenie o braku korelacji między zmiennymi uwzględnionymi w specyfikacji modelu a składnikiem losowym. Jednak w wielu ważnych z punktu widzenia teorii ekonomicznej zastosowaniach takie założenie nie jest spełnione. W takim przypadku nie można udowodnić zgodności estymatora MNK.

W modelu $Y = X\beta + \varepsilon$

- zmiennymi egzogenicznymi nazywamy zmienne, które nie są skorelowane ze składnikami losowymi
- zmiennymi endogenicznymi nazywamy zmienne, które są skorelowane ze składnikami losowymi

2 Równoczesność

O problemie równoczesności mówimy, gdy występuje niezerowa korelacja pomiędzy zmienną objaśniającą x_i a równoczesnym błędem losowym ε_j . Gdy $\mathbb{E}(\varepsilon|X) \neq 0$ to

$$\begin{aligned}\mathbb{E}(b) &= \mathbb{E}\left((X'X)^{-1} X'y|X\right) \\ \mathbb{E}(b) &= \beta + (X'X)^{-1} X'\mathbb{E}(\varepsilon|X) \neq \beta\end{aligned}$$

Więc estymator wektora parametrów jest obciążony.

2.1 Przykłady

Model Keynesowski gospodarki zakłada, że

$$\text{PKB} = \text{konsumpcja} + \text{inwestycje} + \text{export netto}$$

Z drugiej strony Keynesowska funkcja konsumpcji zakłada, że $C_t = f(Y_t)$. Szacując jej parametry

$$C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t$$

nie są spełnione założenia modelu, gdyż $\text{cov}(X_t, \varepsilon_t) \neq 0$
Szacujemy model autoregresji postaci

$$y_t = f(y_{t-1}, y_{t-2}, \dots) + \varepsilon_t$$

ale

$$y_{t+1} = f(y_{t-2}, y_{t-3}, \dots) + \varepsilon_{t+1}$$

zatem $\text{cov}(y_{t-1}, \varepsilon_{t-1}) \neq 0$. Wobec tego w modelu (1) zmienne objaśniające są skorelowane z błędem losowym. Zazwyczaj w zbiorach danych mikroekonomicznych brakuje informacji o zdolnościach respondentów. Mimo wszystko szacuje się równanie płacy typu Mincera

$$\ln(placa) = \beta_0 + \beta_1 plec + \beta_2 wiek + \beta_3 wiek^2 + \beta_4 wyksz + u$$

Ponieważ w modelu pominięto zmienną niezależną zdolności to składnik losowy ma postać

$$u = \gamma_0 + \gamma_1 zdolnosci + \phi$$

Z drugiej strony poziom wykształcenia jest determinowany przez zdolności respondenta. Zatem

$$\begin{aligned} \text{cov}(u, wyksz) &= \text{cov}(zdolnosci + \phi, wyksz) = \\ &= \text{cov}(zdolnosci, wyksz) + \text{cov}(\phi, wyksz) \neq 0 \end{aligned}$$

Ponieważ zmienna pominięta jest dodatnio skorelowana z uzyskanym wykształceniem i wpływa dodatnio na zmienną zależną to parametr przy zmiennej będzie dodatnio obciążony.

Metoda zmiennych instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym. Pozwala również uzyskać zgodne oszacowania parametrów w przypadku występowania problemu równoczesności. Polega ona na zastępowaniu oryginalnych zmiennych instrumentalami. Instrumenty powinny być skorelowane ze zmiennymi objaśniającymi, ale nie powinny być skorelowane z błędem losowym. Znalezienie właściwych instrumentów jest najciekawszym, ale również najtrudniejszym etapem badania.

Oznaczmy przez Z macierz zmiennych instrumentalnych (instrumentów). Estymator MZI jest zgodny, gdy spełnione są następujące warunki

$$\begin{aligned} \text{plim} \left(\frac{1}{n} Z' \varepsilon \right) &= \text{plim} \left(\frac{1}{n} \sum_i z'_i \varepsilon_i \right) = \mathbb{E}(z'_i \varepsilon_i) = 0 \\ \text{plim} \left(\frac{1}{n} Z' X \right) &= \text{plim} \left(\frac{1}{n} \sum_i z'_i x_i \right) = \mathbb{E}(z'_i x_i) \neq 0 \end{aligned}$$

Dodatkowo $r(\mathbb{E}(z'_i x_i)) = k$

$$plim\left(\frac{1}{n}Z'Z\right) = plim\left(\frac{1}{n}\sum_i z'_i z_i\right) = \mathbb{E}(z'_i z_i) \neq 0$$

Metoda polega na zastępowaniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne.

- macierz instrumentów Z musi zawierać co najmniej tyle zmiennych, ile oryginalna macierz X
- Ale nie w każdym przypadku konieczne jest posiadanie k nowych zmiennych
- Zmienne z macierzy X , które są nieskorelowane ze składnikiem losowym mogą same dla siebie stanowić instrumenty
- W rezultacie potrzeba co najmniej tylu dodatkowych zmiennych instrumentalnych ile jest zmiennych skorelowanych ze składnikiem losowym.

Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1} Z'X}_{\beta} = Z\beta = \hat{P}_Z X$$

Dysponując macierzą instrumentów \hat{X} budujemy estymator

$$\begin{aligned}\beta_{MZI} &= (\hat{X}'X)^{-1} \hat{X}'y = (X'P_Z X)^{-1} X'P_Z y = \\ &= (X'Z(Z'Z)^{-1}Z')^{-1} X'Z(Z'Z)^{-1}Z'y\end{aligned}$$

jeżeli $r(2) = r(X)$, czyli tyle nowych zmiennych, ile zmiennych w macierzy X skorelowanych ze składnikiem.

Estymator MZI jest również nazywany estymatorem dwustopniowej metody najmniejszych kwadratów.

- W pierwszym kroku przeprowadzana jest regresja zmiennych endogenicznych na zmienne instrumentalne
- W drugim, oryginalne wartości zmiennych są zastępowane przez zwartości dopasowane z pierwszego kroku i oblicze są oszacowania poszczególnych parametrów

- W kolejnym kroku pokażemy, że przy warunkach zdefiniowanych wcześniej estymator MZI jest zgodny

Estymator MZI jest dany wzorem

$$\beta_{MZI} = \left(X'Z (Z'Z)^{-1} Z' \right)^{-1} X'Z (Z'Z)^{-1} Z'y$$

więc

$$plim \beta_{MZI} = plim \left(X'Z (Z'Z)^{-1} Z' \right)^{-1} X'Z (Z'Z)^{-1} Z'y$$

Wnioski

- Estymator MZI jest zgodny nawet, gdy w modelu występują zmienne endogeniczne
- Ale, gdy nie ma takich zmiennych, a spełnione są założenia MNK to estymator MZI nie jest efektywny
- MZI można traktować jako uogólnienie MNK
- Wobec tego wszystkie testy stosowane przy MNK mogą być stosowane przy MZI
- Testy specyficzne dla MZI weryfikują założenie o egzogeniczności zmiennych oraz poprawności wykorzystanych instrumentów.

2.2 Test Hausmana

Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych. Estymator MZI jest estymatorem zawsze zgodnym.

Przy prawdziwej H_0 o egzogeniczności zmiennych X oba estymatory są zgodne, ale estymator MZI ma większą wariancję niż estymator MNK.

Przy fałszywej H_0 estymator MZI jest zgodny i efektywny, a estymator MNK nie jest zgodny.

Statystyka testowa jest dana przez formę kwadratową

$$(\beta_{MZI} - \beta_{MNK})' \Sigma_{\beta_{MZI} - \beta_{MNK}}^{-1} (\beta_{MZI} - \beta_{MNK}) \xrightarrow{D} \chi^2(r(\Sigma))$$

Gdy różnica jest duża sugeruje to wykorzystanie MNK.

2.3 Test Surgana

- Test Surgana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = \left(I - X (X' P_Z X)^{-1} X' P_Z \right) \varepsilon$$

- Oczywiście M_Z jest macierzą idempotentną, rzędu $p - k$
- Przy prawdziwości H_0 jest brak korelacji instrumentów z błędami losowymi

$$\frac{e' P_Z e}{\sigma^2} \sim \chi_{p-k}^2$$

- Niestety przeprowadzenie testu jest wyłącznie możliwe, gdy liczba instrumentów przekracza liczbę zmiennych objaśniających

3 Uogólniona metoda zminnych instrumentalnych

Zakładamy, że mamy więcej instrumentów niż zminnych objaśniających

wymiar macierzy $X \neq$ wymiar macierzy Z

Założyliśmy, że $\mathbb{E}(Z'\varepsilon) = 0$, zatem $\mathbb{E}(Z'(y - X\beta)) = 0$

Skoro średnia z próby jest oceną wartości oczekiwanej, to

$$\frac{1}{n} (Z'(y - X\beta)) = 0$$

3.1 Estymator uogólnionej MZI (UMZI, GIV)

Przy powyższym założeniu, UMZI polega na minimalizacji następującego wyrażenia

$$Q_n(\beta) = \left[\frac{1}{n} Z'(y - X\beta) \right]' W_n \left[\frac{1}{n} Z'(y - X\beta) \right]$$

gdzie W_n jest symetryczną, kwadratową macierzą wag, o wymiarze równym liczbie instrumentów w macierzy Z . Macierz ta wskazuje jakie wagi należy przypisać poszczególnym równaniom w układzie równań $\frac{1}{n} (Z'(y - X\beta)) =$

0, czyli określa które instrumenty są bardziej, a które mniej ważne. Minimalizując Q_n pierwszą pochodną po β przyrównujemy do zera

$$-ZX'ZW_nZ'y + ZX'ZW_nZ'X\hat{\beta} = 0$$

co oznacza, że

$$X'ZW_nZ'y = X'ZW_nZ'X\hat{\beta}$$

Zakładając, że

$$\det(X'ZW_nZ'X) \neq 0$$

uzyskujemy estymator UMZI

$$\hat{\beta}_{GIV} = (X'ZW_nZ'X)^{-1} X'ZW_nZ'y$$

W jaki sposób dobieramy macierz wag W_n ?

Różne macierze wag prowadzą do estymatorów różnej postaci, ale przy spełnieniu wymaganych założeń wszystkie uzyskane estymatory są nieobciążone i zgodne. Optymalna macierz W_n jest proporcjonalna do odwrotności macierzy wariancji i kowariancji parametrów. W szczególnym przypadku, jeśli przyjmujemy, że składnik losowy jest sferyczny, uzyskujemy optymalną postać W_n

$$W_n^{opt} = \left(\frac{1}{n}Z'Z\right)^{-1}$$

Podstawiając optymalną macierz wag do estymatora, uzyskujemy estymator

$$\hat{\beta}_{GIV} = \left(X'Z \left(\frac{1}{n}Z'Z\right)^{-1} Z'X\right)^{-1} X'Z \left(\frac{1}{n}Z'Z\right)^{-1} Z'y$$

Estymator ten, przyjmuje różną postać, w zależności od wybranej macierzy wag.

4 Estymator potrójnej metody najmniejszy kwadratów (3MNK)

- Estymacja parametrów równania za pomocą KMNK abstrahuje od endogeniczności niektórych zmiennych objaśniających
- 2MNK uwzględnia endogeniczność, lecz za jej pomocą estymujemy parametry każdego równania osobno

- Korelacje między składnikami losowymi poszczególnych równań nie zostają uwzględnione w procesie estymacji
- Wady te w modelu nie będzie, gdy przeprowadzimy łączną estymację parametrów wszystkich równań, uwzględniając korelacje składników losowych poszczególnych równań

Potraktujemy model wielorównaniowy jako macierzowy model jednorównaniowy. Zapisujemy nasz n -wymiarowy model jako

$$\begin{aligned} y_1 &= \tilde{Z}_1 \cdot \tilde{F}_1 + \varepsilon_1 \\ y_2 &= \tilde{Z}_2 \cdot \tilde{F}_2 + \varepsilon_2 \\ &\vdots \\ y_m &= \tilde{Z}_m \cdot \tilde{F}_m + \varepsilon_m \end{aligned}$$

gdzie $\tilde{Z}_i = [\tilde{Y}_i \quad \tilde{Z}_i]$ - wektor zmiennych objaśniających (endo- i egzogenicznych) w n -tym równaniu modelu.

Niech \tilde{F}_j^{2MNK} - wektor oszacowań parametrów j -tego równania za pomocą 2MNK.

Znając go, możemy oszacować macierz wariancji-kowariancji wektora składników losowych modelu

$$\begin{aligned} \text{Var}(\varepsilon_t) &= \mathbb{E}(\varepsilon_t \varepsilon_t^T) \\ \hat{\Sigma} &= [\hat{\sigma}_{ij}] \\ \hat{\sigma}_{ij} &= (y_i - Z_i \tilde{F}_i^{2MNK})(y_j - Z_j \tilde{F}_j^{2MNK})^T / T \end{aligned}$$

gdzie T - liczba obserwacji.

Alternatywnie zamiast T , można podzielić sumę iloczynów w liczniku przez średnią geometryczną liczbę stopni swobody obu równań i oraz j

$$\sqrt{(T - \tilde{M}_i - \tilde{K}_i)(T - \tilde{M}_j - \tilde{K}_j)}$$

Zapisujemy cały model w jednym równaniu macierzowym

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_y = \underbrace{\begin{bmatrix} \tilde{Z}_1 & 0 & \dots & 0 \\ 0 & \tilde{Z}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{Z}_m \end{bmatrix}}_Z \underbrace{\begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_m \end{bmatrix}}_F + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}}_\varepsilon$$

Przy czym $\varepsilon - 1, \varepsilon - 2, \dots$ są wektorami pionowymi o wymiarach $T \times 1$

4.1 Kolejne kroki w estymacji 3MNK

Krok 1. Estymacja w postaci zredukowanej (MNK) i obliczenie wartości teoretycznych dla równania j

$$\hat{\tilde{Z}}_j = X\Pi_j = X(X^T X)^{-1} X^T \tilde{Z}_j$$

Macierzowo dla całego systemu

$$\begin{bmatrix} X(X^T X)^{-1} X^T \tilde{Z}_1 & 0 & \dots & 0 \\ 0 & X(X^T X)^{-1} X^T \tilde{Z}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X(X^T X)^{-1} X^T \tilde{Z}_m \end{bmatrix}$$

lub inaczej

$$\left\{ I_m \otimes \left[X(X^T X)^{-1} X^T \right] \right\} Z$$

Krok 2. Estymacja w postaci strukturalnej parametrów pojedynczych równań (2MNK). Wartości z próby w KMNK zamienione na wartości teoretyczne

$$\hat{F}^{2MNK} = (\hat{Z}^T \hat{Z}) \hat{Z}^T y$$

Krok 3. Uwzględnienie jednoczesnych korelacji składników losowych w procesie estymacji. Jeżeli poszczególne składniki losowe są sferyczne, to każdy taki wektor ma macierz wariancji-kowariancji postaci

$$\begin{bmatrix} \sigma_{jj}^2 & 0 & \dots & 0 \\ 0 & \sigma_{jj}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{jj}^2 \end{bmatrix}$$

$$\begin{pmatrix} \sigma_{11}^2 & 0 & \dots & 0 & \sigma_{12}^2 & 0 & \dots & 0 & \dots & \sigma_{1m}^2 & 0 & \dots & 0 \\ 0 & \sigma_{11}^2 & \dots & 0 & 0 & \sigma_{12}^2 & \dots & 0 & \dots & 0 & \sigma_{1m}^2 & \dots & 0 \\ \vdots & \vdots & \sigma_{11}^2 & \vdots & \vdots & \vdots & \sigma_{12}^2 & \vdots & \dots & \vdots & \vdots & \sigma_{1m}^2 & \vdots \\ 0 & 0 & \dots & \sigma_{11}^2 & 0 & 0 & \dots & \sigma_{12}^2 & \dots & 0 & 0 & \dots & \sigma_{1m}^2 \\ \sigma_{21}^2 & 0 & \dots & 0 & \sigma_{22}^2 & 0 & \dots & 0 & \dots & \sigma_{2m}^2 & 0 & \dots & 0 \\ 0 & \sigma_{21}^2 & \dots & 0 & 0 & \sigma_{22}^2 & \dots & 0 & \dots & 0 & \sigma_{2m}^2 & \dots & 0 \\ \vdots & \vdots & \sigma_{21}^2 & \vdots & \vdots & \vdots & \sigma_{22}^2 & \vdots & \dots & \vdots & \vdots & \sigma_{2m}^2 & \vdots \\ 0 & 0 & \dots & \sigma_{21}^2 & 0 & 0 & \dots & \sigma_{22}^2 & \dots & 0 & 0 & \dots & \sigma_{2m}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{m1}^2 & 0 & \dots & 0 & \sigma_{m2}^2 & 0 & \dots & 0 & \dots & \sigma_{mm}^2 & 0 & \dots & 0 \\ 0 & \sigma_{m1}^2 & \dots & 0 & 0 & \sigma_{m2}^2 & \dots & 0 & \dots & 0 & \sigma_{mm}^2 & \dots & 0 \\ \vdots & \vdots & \sigma_{m1}^2 & \vdots & \vdots & \vdots & \sigma_{m2}^2 & \vdots & \dots & \vdots & \vdots & \sigma_{mm}^2 & \vdots \\ 0 & 0 & \dots & \sigma_{m1}^2 & 0 & 0 & \dots & \sigma_{m2}^2 & \dots & 0 & 0 & \dots & \sigma_{mm}^2 \end{pmatrix}$$

Wykorzystanie iloczynu Kroneckera w 3MNK

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix}$$

Macierz wariancji-kowariancji składników reszt z 2MNK

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \dots & \sigma_{1m}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}^2 & \sigma_{m2}^2 & \dots & \sigma_{mm}^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{T \times T} = \Sigma \otimes I_T$$

W praktyce korzystamy z 2MNK i otrzymujemy wektor wartości teoretycznych i reszt dla każdego z równań osobno

$$\forall_{j=1,\dots,m} \hat{Y}_j \hat{\varepsilon}_j$$

Obliczamy kowariancje między resztami losowymi poszczególnych równań

$$\hat{\sigma}_{ij} = \frac{\hat{\varepsilon}_i^T \hat{\varepsilon}_j}{T}$$

$$\hat{\Sigma} = \begin{bmatrix} \hat{\sigma}_{11}^2 & \hat{\sigma}_{12}^2 & \dots & \hat{\sigma}_{1m}^2 \\ \hat{\sigma}_{21}^2 & \hat{\sigma}_{22}^2 & \dots & \hat{\sigma}_{2m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{m1}^2 & \hat{\sigma}_{m2}^2 & \dots & \hat{\sigma}_{mm}^2 \end{bmatrix}$$

Jeżeli znamy macierz wariancji-koowariancji całego wektora składników losowych po oszacowaniu $\hat{\Sigma}$, to możemy zastosować UMNK w odniesieniu do modelu $yZF + \varepsilon, \varepsilon \sim (o, \Omega)$

$$\tilde{F}^{UMNK} = \left(Z^T \Omega^{-1} Z \right)^{-1} Z^T \Omega^{-1} y$$

Przy jednoczesnym zastosowaniu 2MNK zastępujemy w powyższym wzorze Z wartościami teoretycznymi z kroku 1.

$$\hat{F}^{UMNK} = \left(\hat{Z}^T \Omega^{-1} \hat{Z} \right)^{-1} \hat{Z}^T \Omega^{-1} y$$

Skoro $\hat{\Omega} = \hat{\Sigma} \otimes I_T$ oraz $\hat{Z} = \left\{ I_m \otimes \left[X (X^T X)^{-1} X^T \right] \right\} Z$, to

$$\begin{aligned} \tilde{F}^{UMNK} &= \\ &= \left[Z^T \left\{ I_m \otimes \left[X (X^T X)^{-1} X^T \right] \right\} (\hat{\Sigma} \otimes I_T)^{-1} \left\{ I_m \otimes \left[X (X^T X)^{-1} X^T \right] \right\} Z \right]^{-1} \cdot \\ &\quad \cdot Z^T \left\{ I_m \otimes \left[X (X^T X)^{-1} X^T \right] \right\}^T (\hat{\Sigma} \otimes I_T)^{-1} y \end{aligned}$$

Z własności iloczynu Kroneckera

$$\begin{aligned} \tilde{F}^{UMNK} &= \\ &= \left[Z^T \left\{ \hat{\Sigma}^{-1} \otimes \left[X (X^T X)^{-1} X^T \right] \right\} Z \right]^{-1} Z^T \left\{ \hat{\Sigma}^{-1} \otimes \left[X (X^T X)^{-1} X^T \right] \right\} y \end{aligned}$$

4.2 Własności estymatora 3MNK

1. Jest zgodny
2. Jest asymptotycznie najlepszy w pewnej klasie estymatorów zmiennych instrumentalnych (do której należy 2MNK), więc jest bardziej efektywny od 2MNK.
3. Jeżeli założymy normalność ε , estymator jest asymptotycznie równoważny estymatorowi MNK z pełną informacją (przedstawionej dalej): 3MNK ma ten sam rozkład asymptotyczny, ale nie oznacza to numerycznej równoważności oszacowań w skończonej próbie.

4.3 Szczególne przypadki 3MNK

- W sytuacji, gdy wiem, *a priori*, że Σ jest macierzą diagonalną, czyli równoczesne kowariancje składników losowych są zerowe, metoda 3MNK

oczywiście nie poprawia efektywności. W takim przypadku 3MNK sprowadza się do 2MNK. Taki sąd *a priori* jest jednak dość mocny i w praktyce zakładamy, raczej niezerowe równoczesne kowariancje.

- Kiedy wszystkie równania są jednoznacznie identyfikowalne, 3MNK sprowadza się do 2MNK.

5 Estymator metody największej wiarygodności z pełną informacją (MNWPO, FIML)

Układ równań w postaci strukturalnej $Ay + BX = \varepsilon$. Składnik losowy postaci zredukowanej ma macierz wariancji-kowariancji $\hat{\Sigma}$. Jeżeli poszczególne wektory składnika losowego z różnych okresów (i) są niezależne i (ii) mają wielowymiarowy rozkład normalny, to ich łączna funkcja gęstości

$$L_\varepsilon = (2\pi)^{-\frac{mT}{2}} |\hat{\Sigma} \otimes I_T|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} \varepsilon^T (\hat{\Sigma} \otimes I_T)^{-1} \varepsilon \right)$$

Funkcja gęstości obserwacji i składnika losowego zależą od siebie $L_y = L_\varepsilon \cdot |\det A|$

$$(2\pi)^{-\frac{mT}{2}} |\hat{\Sigma} \otimes I_T|^{-\frac{1}{2}} |\det A| \exp \left(-\frac{1}{2} (y - \tilde{Z}\tilde{F})^T (\hat{\Sigma} \otimes I_T)^{-1} (y - \tilde{Z}\tilde{F}) \right)$$

Po zlogarytmowaniu

$$L_y = -\frac{mT}{2} \ln(2\pi) - \frac{T}{2} \ln |\Sigma| + T \cdot \ln |\det A| - \frac{1}{2} \left((y - \tilde{Z}\tilde{F})^T (\hat{\Sigma} \otimes I_T)^{-1} (y - \tilde{Z}\tilde{F}) \right)$$

Oszacowanie FIML otrzymamy maksymalizując powyższe wyrażenie ze względu na niezerowe wartości parametrów modelu.

5.1 Estymacja parametrów modelu liniowego estymatorem uogólnionym Aitkena

W sytuacji, w której niesferyczność składników zakłócających, przejawiająca się zmiennością wariancji, bądź kowariancji w czasie składników zakłócających, nie jest wynikiem błędów konstrukcji modelu estymator KMNK nie jest efektywny. Zdefiniujemy strukturę stochastyczną modelu, w którym składniki zakłócające nie są sferyczne

$$\begin{aligned} \mathbb{E}\varepsilon &= 0 \\ \mathbb{E}\varepsilon\varepsilon' &= \sigma_\varepsilon^2 \Omega, \end{aligned}$$

gdzie Ω jest $T \times T$ wymiarową macierzą, określoną dodatnio

$$\mathbb{E}X'\varepsilon = 0$$

Jeżeli składniki zakłócające są nieskorelowane w czasie, ale mają zmienne w czasie wariancje, wtedy macierz Ω ma postać

$$\Omega = \begin{bmatrix} \omega_{11} & 0 & \dots & 0 \\ 0 & \omega_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{TT} \end{bmatrix}$$

gdzie $\sigma_\varepsilon^2 \omega_{tt}$ jest wariancją jednoczesnych składników zakłócających.

Jeśli składniki zakłócające mają stałe wariancje, ale są skorelowane w czasie, wtedy macierz Ω może być zapisana następująco

$$\Omega = \begin{bmatrix} 1 & \omega_{12} & \dots & \omega_{1T} \\ \omega_{21} & 1 & \dots & \omega_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{T1} & \omega_{T2} & \dots & 1 \end{bmatrix}$$

gdzie $\sigma_\varepsilon^2 \omega_{ij}$ jest kowariancją niejednoczesnych składników zakłócających modelu.

W przypadku, gdy występuje zarówno skorelowanie w czasie, jak również zmienność wariancji, wtedy

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} & \dots & \omega_{1T} \\ \omega_{21} & \omega_{22} & \dots & \omega_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{T1} & \omega_{T2} & \dots & \omega_{TT} \end{bmatrix}$$

Rozważmy estymator parametrów strukturalnych $\hat{\beta}$ modelu $y = X\beta + \varepsilon$, w którym składniki zakłócające są niesferyczne i który minimalizuje uogólnioną sumę kwadratów reszt

$$\tilde{S} = \tilde{\varepsilon}'\Omega^{-1}\tilde{\varepsilon}$$

gdzie $\tilde{\varepsilon} = y - X\tilde{\beta}$ jest wektorem reszt. Rozwijając otrzymujemy

$$\tilde{S} = y'\Omega^{-1}y - 2\tilde{\beta}'X'\Omega^{-1}y + \tilde{\beta}'X'\Omega^{-1}X\tilde{\beta}.$$

Szukać będziemy takiego $\tilde{\beta}$, dla którego zachodzi

$$\tilde{S} \xrightarrow{\tilde{\beta}} \min$$

Obliczamy wektor pochodnych cząstkowych \tilde{S} względem $\tilde{\beta}$ i przyrównujemy go od zera. Otrzymamy w ten sposób

$$\frac{\partial \tilde{S}}{\partial \tilde{\beta}} = -2X'\Omega^{-1}y + 2X'\Omega^{-1}X\tilde{\beta} = 0$$

W konsekwencji możemy zapisać tzw. uogólniony układ równań normalnych

$$X'\Omega^{-1}X\tilde{\beta} = X'\Omega^{-1}y$$

którego rozwiązaniem jest, jeśli X jest macierzą o pełnym rzędzie kolumnowym, tak, że $|X'\Omega^{-1}X| > 0$, estymator uogólnionej metody najmniejszych kwadratów, zwany estymatorem Aitkena i mającym postać

$$\tilde{\beta} = (X'\Omega X)^{-1} X'\Omega^{-1}y$$

Ponieważ macierz drugich pochodnych \tilde{S} względem $\tilde{\beta}$, tj.

$$\frac{\partial^2 \tilde{S}}{\partial \tilde{\beta} \partial \tilde{\beta}'} = 2X'\Omega^{-1}X$$

jest macierzą określoną dodatnio, zatem estymator Aitkena rzeczywiście minimalizuje uogólnioną sumę kwadratów reszt.

Estymator $\tilde{\beta}$ należy do klasy estymatorów liniowych, który można zapisać jako

$$\tilde{\beta} = \tilde{C}'y$$

gdzie $\tilde{C}' = (X'\Omega X)^{-1} X'\Omega^{-1}$ jest macierzą $(K+1) \times T$ wymiarową. Jeśli elementy macierzy X są nielosowe lub losowe, ale nieskorelowane ze składnikami zakłócającymi modelu, estymator Aitkena jest nieobciążony. Możemy zapisać, że

$$\tilde{\beta} = \beta + \tilde{C}'\varepsilon$$

dlatego, że spełniony jest warunek konieczny nieobciążoności, tj.

$$\tilde{C}'X = I_{K+1}$$

Zatem możemy zapisać, że

$$\mathbb{E}\tilde{\beta} = \beta$$

Można udowodnić, zakładając nielosowość macierzy X , że macierz wariancji-kowariancji błędów estymacji w tym przypadku jest równa

$$\Sigma_{\tilde{\beta}}^2 = \mathbb{E} \left(\tilde{\beta} - \beta \right) \left(\tilde{\beta} - \beta \right)' = \mathbb{E} \tilde{C}' \varepsilon \left(\tilde{C}' \varepsilon \right)' = \tilde{C}' \sigma_{\varepsilon}^2 \Omega \tilde{C} = \sigma_{\varepsilon}^2 \tilde{C}' \Omega \tilde{C}$$

Znając postać macierzy \tilde{C}' możemy wykazać, że

$$\Sigma_{\tilde{\beta}} = \sigma_{\varepsilon}^2 \left(X' \Omega^{-1} X \right)^{-1}$$

Estymator Aitkena jest najlepszym liniowym, nieobciążonym estymatorem w klasie liniowych estymatorów dla modelu z niesferycznymi składnikami zakłócającymi. Zatem dowolny estymator nieobciążony typu

$$\tilde{\beta}_D = \tilde{C}_D' y$$

dla którego zachodzi $\tilde{C}_D X = I_{K+1}$, charakteryzuje się macierzą wariancji-kowariancji

$$\Sigma_{\tilde{\beta}_D}^2 = \mathbb{E} \left(\tilde{\beta}_D - \beta \right) \left(\tilde{\beta}_D - \beta \right)' = \sigma_{\varepsilon}^2 \tilde{C}_D' \Omega \tilde{C}_D$$

której elementy są nie mniejsze niż macierz wariancji-kowariancji estymatora Aitkena. Jeśli macierz Ω jest określona dodatnio, wtedy istnieje $T \times T$ wymiarowa macierz P spełniająca następujące własności

$$\begin{aligned} P' P &= \Omega^{-1} \\ P \Omega \Omega' &= I_T \end{aligned}$$

gdzie I_T jest macierzą jednostkową stopnia T . Macierz P nazywać będziemy macierzą transformacji. Macierz ta umożliwia ukazanie estymatora Aitkena jako specyficznego estymatora MNK, dla transformowanych obserwacji zmiennych modelu. Niech pierwotny model

$$y = X\beta + \varepsilon$$

zostanie przemnożony przez macierz P tak, że

$$Py = PX\beta + P\varepsilon$$

Ponieważ macierz P ma nielosowe elementy, zatem jest prawdą, że

$$\begin{aligned} \mathbb{E} P \varepsilon &= 0 \\ \mathbb{E} P \varepsilon \left(P \varepsilon \right)' &= \sigma_{\varepsilon}^2 P \Omega \Omega' = \sigma_{\varepsilon}^2 I_T \end{aligned}$$

Widzimy zatem, że składniki zakłócające modelu transformowanego mają sferyczne składniki zakłócające i parametry tego modelu można szacować metodą najmniejszych kwadratów. Niech

$$Y^* = X^*\beta + \varepsilon^*$$

gdzie

$$\begin{aligned} Y^* &= Py \\ X^* &= PX \\ \varepsilon^* &= P\varepsilon \end{aligned}$$

Stosując MNK modelu, w którym pierwotne obserwacje zostały zważone elementami macierzy Ω jest równoważne estymacji uogólnioną metodą najmniejszych kwadratów.

5.2 Przypadek zmienności wariancji

Dla przypadku składników zakłócających o zmiennych wariancjach z macierzą Ω o postaci

$$\Omega = \begin{bmatrix} \omega_{11} & 0 & \dots & 0 \\ 0 & \omega_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{TT} \end{bmatrix}$$

można wykazać, że

$$\Omega^{-1} = \begin{bmatrix} \frac{1}{\omega_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{\omega_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\omega_{TT}} \end{bmatrix} \quad P = \begin{bmatrix} \frac{1}{\sqrt{\omega_{11}}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\omega_{22}}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\omega_{TT}}} \end{bmatrix}$$

Macierze te spełniają równania

$$P'P = \Omega^{-1} \quad P\Omega P' = I_T$$

5.3 Przypadek autokorelacji

W przypadku, gdy składniki zakłócające są skorelowane w czasie i gdy skorelowanie nie jest wynikiem niepoprawnej specyfikacji modelu, staramy się wykorzystać estymator Aitkena w celu eliminacji skutków jej występowania.

W takim przypadku elementy macierzy Ω i P są funkcjami parametrów występujących w procesie generującym skorelowane w czasie składniki losowe. Rozważmy przykład, w którym składniki zakłócające w liniowym modelu.

$$y_t\beta_0 + \beta_1x_{1t} + \cdots + \beta_kx_{kt} + \varepsilon_t$$

są generowane przez stacjonarny proces autoregresyjny rzędu pierwszego.

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \eta_t$$

gdzie

- $|\rho_1| < 1$ jest współczynnikiem autokorelacji rzędu pierwszego
- η_t jest czysto losowym (sferycznym) składnikiem zakłócającym, spełniającym następujące warunki
 - $\mathbb{E}\eta_t = 0$
 - $\mathbb{E}\eta_t^2 = \sigma_\eta^2$
 - $\mathbb{E}\eta_t\eta_s = 0$ dla $t \neq s$

Rozważmy jakie parametry rozkładu charakteryzować będą kolejne składniki zakłócające ε_t , generowane przez model autoregresyjny. Przedstawiamy obecnie składnik zakłócający jako funkcję wszystkich przeszłych i bieżących zakłóceń η_t . Drogą kolejnych podstawień otrzymamy

$$\begin{aligned}\varepsilon_t &= \rho_1 \underbrace{(\rho_1\varepsilon_{t-2} + \eta_{t-1})}_{\varepsilon_{t-1}} + \eta_t = \rho_1^2\varepsilon_{t-2} + \rho_1\eta_{t-1} + \eta_t \\ \varepsilon_t &= \rho_1^2 \underbrace{(\rho_1\varepsilon_{t-3} + \eta_{t-2})}_{\varepsilon_{t-2}} + \rho_1\eta_{t-1} + \eta_t = \rho_1^3\varepsilon_{t-3} + \rho_1^2\eta_{t-2} + \rho_1\eta_{t-1} + \eta_t \\ &\vdots \\ \varepsilon_t &= \rho_1^{s+1}\varepsilon_{t-s-1} + \rho_1^s\eta_{t-s} + \cdots + \rho_1^2\eta_{t-2} + \rho_1\eta_{t-1} + \eta_t\end{aligned}$$

Jeśli proces generujący składniki zakłócające jest stacjonarny, tj $|\rho_1| < 1$ wtedy

$$\rho_1^{s+1}\varepsilon_{t-s-1} \rightarrow 0$$

i w postaci końcowej

$$\varepsilon_t = \rho_1^s\eta_{t-s} + \cdots + \rho_1^2\eta_{t-2} + \rho_1\eta_{t-1} + \eta_t$$

Składnik losowy ε_t jest przedstawiony jako ważona funkcja wszystkich (przyszłych i bieżącego) sferycznych składników η . Ponieważ parametry rozkładu η_t są zadane, zatem można wyznaczyć parametry rozkładu ε_t . W przypadku wariancji możemy zapisać, że

$$\begin{aligned}\mathbb{E}\varepsilon_t^2 &= \mathbb{E}\left(\rho_1^s \eta_{t-s} + \dots + \rho_1^2 \eta_{t-2} + \rho_1 \eta_{t-1} + \eta_t\right)^2 = \\ &= \rho_1^{2s} \mathbb{E}\eta_{t-s}^2 + \dots + \rho_1^4 \mathbb{E}\eta_{t-2}^2 + \rho_1^2 \mathbb{E}\eta_{t-1}^2 + \mathbb{E}\eta_t^2\end{aligned}$$

gdzie wszystkie kowariancje niejednoczesnych składników zakłócających η_t, η_s , dla $t \neq s$ przyrównano do zera. W konsekwencji można zapisać, że

$$\mathbb{E}\varepsilon_t^2 = \eta_1^{2s} \sigma_\eta^2 + \dots + \eta_1^4 \sigma_\eta^2 + \eta_1^2 \sigma_\eta^2 + \sigma_\eta^2 = \sigma_\eta^2 (\eta_1^{2s} + \dots + \eta_1^4 + \eta_1^2 + 1)$$

ponieważ suma elementów w nawiasie stanowi szereg geometryczny zbieżny o ilorazie ρ_1^2 , zatem jego suma jest równa $\frac{1}{1-\rho_1^2}$. Ostatecznie więc mamy:

$$> E\varepsilon_t^2 = \sigma_\eta^2 \cdot \frac{1}{1-\rho_1^2} = \sigma_\varepsilon^2$$

Z powyższego zapisu widać, że skorelowane w czasie składniki zakłócające mają stałe w czasie wariancje. Aby obliczyć kowariancje przeprowadzimy niestępującą procedurę kolejnych podstawień

$$\begin{aligned}\varepsilon_{t+1} &= \rho_1 \varepsilon_t + \eta_{t+1} \\ \varepsilon_{t+2} &= \rho_1 \underbrace{(\rho_1 \varepsilon_t + \eta_{t+1})}_{\varepsilon_{t+1}} + \eta_{t+2} = \rho_1^2 \varepsilon_t + \rho_1 \eta_{t+1} + \eta_{t+2} \\ \varepsilon_{t+3} &= \rho_1 \underbrace{(\rho_1^2 \varepsilon_t + \rho_1 \eta_{t+1} + \eta_{t+2})}_{\varepsilon_{t+2}} + \eta_{t+3} = \rho_1^3 \varepsilon_t + \rho_1^2 \eta_{t+1} + \rho_1 \eta_{t+2} + \eta_{t+3} \\ &\vdots \\ \varepsilon_{t+\tau} &= \rho_1^\tau \varepsilon_t + \rho_1^{\tau-1} \eta_{t+1} + \dots + \rho_1^2 \eta_{t+\tau-2} + \rho_1 \eta_{t+\tau-1}\end{aligned}$$

Mnożąc obie strony ostatniej równości przez ε_t i obliczając wartość oczekiwaną otrzymamy

$$\mathbb{E}\varepsilon_t \varepsilon_{t+\tau} = \rho_1^\tau \mathbb{E}\varepsilon_t^2 + \rho_1^{\tau-1} \mathbb{E}\varepsilon_t \eta_{t+1} + \dots + \rho_1^2 \mathbb{E}\varepsilon_t \eta_{t+\tau-2} + \rho_1 \mathbb{E}\varepsilon_t \eta_{t+\tau-1} = \rho_1^\tau \sigma_\varepsilon^2$$

Ponieważ ε_t zależy tylko od opóźnionych i bieżących składników zakłócających η_{t-i} oraz η_t , nie zależy natomiast od przyszłych składników η_{t+j} . Jeśli składniki zakłócające generowane są przez proces autoregresyjny rzędu

pierwszego, to prawdziwe są następujące własności:

$$\begin{aligned}\mathbb{E}\varepsilon_t &> 0 \\ \mathbb{E}\varepsilon_t^2 &= \sigma_\varepsilon^2 = \frac{\sigma_\eta^2}{1 - \rho_1^2} \\ \mathbb{E}\varepsilon_t\varepsilon_{t+\tau} &= \rho_1^\tau \sigma_\varepsilon^2\end{aligned}$$

Możemy pokazać, że $T \times T$ macierz wariancji-kowariancji składników zakłócających $\sigma_\varepsilon^2\Omega$ ma elementy będące funkcjami współczynnika ρ_1 . Możemy zapisać, że

$$\begin{aligned}\Omega &= \begin{bmatrix} 1 & \rho_1 & \rho_1^2 & \rho_1^3 & \dots & \rho_1^5 \\ \rho_1 & 1 & \rho_1 & \rho_1^2 & \dots & \rho_1^4 \\ \rho_1^2 & \rho_1 & 1 & \rho_1 & \dots & \rho_1^3 \\ \rho_1^3 & \rho_1^2 & \rho_1 & 1 & \dots & \rho_1^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_1^5 & \rho_1^4 & \rho_1^3 & \rho_1^2 & \dots & 1 \end{bmatrix} \\ \Omega^{-1} &= \frac{1}{1 - \rho_1^2} \begin{bmatrix} 1 & -\rho_1 & 0 & 0 & \dots & 0 & 0 \\ -\rho_1 & 1 + \rho_1^2 & -\rho_1 & 0 & \dots & 0 & 0 \\ 0 & -\rho_1 & 1 + \rho_1^2 & -\rho_1 & \dots & 0 & 0 \\ 0 & 0 & -\rho_1 & 1 + \rho_1^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 + \rho_1^2 & -\rho_1 \\ 0 & 0 & 0 & 0 & \dots & -\rho_1 & 1 \end{bmatrix}\end{aligned}$$

Macierz transformacji P może być przedstawiona w postaci

$$P = \frac{1}{\sqrt{1 - \rho_1^2}} \begin{bmatrix} \sqrt{1 - \rho_1^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho_1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho_1 & 1 \end{bmatrix}$$

Rozważmy równanie $Py = PX\beta + P\varepsilon$. Ponieważ obie strony tej równości możemy podzielić przez stałą $\frac{1}{\sqrt{1 - \rho_1^2}}$, zatem transformowanymi obserwacjami zmiennej endogenicznej, zmiennych objaśniających oraz składników zakłóca-

jących są

$$\begin{aligned}\sqrt{1-\rho_1^2}Py &= \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & 0 & \dots & 0 \\ -\rho_1 & 1 & 0 & \dots & 0 \\ 0 & -\rho_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} y_1\sqrt{1-\rho_1^2} \\ y_2 - \rho_1 y_1 \\ y_3 - \rho_1 y_2 \\ \vdots \\ y_T - \rho_1 y_{T-1} \end{bmatrix} \\ \sqrt{1-\rho_1^2}Px &= \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & 0 & \dots & 0 \\ -\rho_1 & 1 & 0 & \dots & 0 \\ 0 & -\rho_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & x_{11} & x_{12} & x_{14} & \dots & x_{1T} \\ 1 & x_{21} & x_{22} & x_{24} & \dots & x_{2T} \\ 1 & x_{31} & x_{32} & x_{34} & \dots & x_{3T} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{T1} & x_{T2} & x_{T4} & \dots & x_{TT} \end{bmatrix} = \\ &= \begin{bmatrix} \sqrt{1-\rho_1^2} & \sqrt{1-\rho_1^2}x_{11} & \sqrt{1-\rho_1^2}x_{12} & \dots & \sqrt{1-\rho_1^2}x_{1T} \\ 1-\rho_1 & x_{21}-\rho_1x_{11} & x_{22}-\rho_1x_{12} & \dots & x_{2T}-\rho_1x_{1T} \\ 1-\rho_1 & x_{31}-\rho_1x_{21} & x_{32}-\rho_1x_{22} & \dots & x_{3T}-\rho_1x_{2T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\rho_1 & x_{T1}-\rho_1x_{(T-1)1} & x_{T2}-\rho_1x_{(T-1)2} & \dots & x_T-\rho_1x_{(T-1)T} \end{bmatrix}\end{aligned}$$

Zatem typowe przekształcenia autoregresyjne mają postać

$$\begin{aligned}\text{dlat} = 1 & \quad y_1^* = y_1\sqrt{1-\rho_1^2} \quad x_{ti}^* = x_{ti}\sqrt{1-\rho_1^2} \quad i = 1, \dots, k \\ \text{dlat} = 2, \dots, T & \quad y_t^* = y_t - \rho_1 y_{t-1} \quad x_{ti}^* = x_{ti} - \rho_1 x_{t-1,i} \quad i = 1, \dots, k\end{aligned}$$

Jeśli nie przeprowadzone jest autoregresyjne przekształcenie pierwszych obserwacji zmiennych, wtedy zamiast macierzy P występuje macierz Q o wymiarach $)T + 1 = < timesT$

$$Q = \begin{bmatrix} -\rho_1 & 1 & 0 & \dots & 0 \\ 0 & -\rho_1 & 1 & \dots & 0 \\ 0 & 0 & -\rho_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\rho_1 \end{bmatrix}$$

Macierz ta nie spełnia warunków takich jak P . W ogólnym przypadku zachodzi bowiem, że $Q'Q \neq P'P = \Omega^{-1}$. Różnica występuje w górnym lewym elemencie obu macierzy.

Nieznany współczynnik ρ_1 estymujemy następująco

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T \varepsilon_t \varepsilon_{t-1}}{\sum_{t=1}^T \hat{\varepsilon}_t^2}$$

Utwórzmy macierz $\hat{\Omega}$ zgodnie z

$$\hat{\Omega} = \begin{bmatrix} 1 & \tilde{\rho}_1 & \tilde{\rho}_1^2 & \dots & \tilde{\rho}_1^{T-1} \\ \tilde{\rho}_1 & 1 & \tilde{\rho}_1 & \dots & \tilde{\rho}_1^{T-2} \\ \tilde{\rho}_1^2 & \tilde{\rho}_1 & 1 & \dots & \tilde{\rho}_1^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\rho}_1^{T-1} & \tilde{\rho}_1^{T-2} & \tilde{\rho}_1^{T-3} & \dots & 1 \end{bmatrix}$$

Wykorzystując powyższą macierz zdefiniujemy dwukrotnie estymator Aitkena w następujący sposób

$$\hat{\beta} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} y$$

Estymator ten jest przybliżeniem optymalnego estymatora Aitkena za znaną macierzą wariancji-kowariancji zakłóceń modelu. Jego macierz wariancji-kowariancji jest przybliżeniem macierzy wariancji-kowariancji estymatora Aitkena. Można tę macierz zapisać

$$\Sigma_{\hat{\beta}}^2 \approx (X' \hat{\Omega}^{-1} X)^{-1}$$

5.4 Konsekwencje występowania niesferycznych składników losowych

Istotnym pytaniem jest jakie własności posiada estymator MNK, jeśli prawdziwym modelem generującym obserwacje zmiennej endogenicznej nie jest model klasyczny, lecz uogólniony. Aby to prześledzić rozważmy raz jeszcze estymator MNK zapisany w postaci

$$\tilde{\beta} = \beta + (X' X)^{-1} X' \varepsilon$$

Niesferyczność składników losowych nie wpływa na nieobciążenie estymatora MNK, jeśli tylko zmienne objaśniające pozostają nieskorelowane ze składnikami losowymi modelu. Istotne konsekwencje związane są z szacowaniem wariancji składników losowych oraz z szacowaniem macierz wariancji-kowariancji błędów estymacji. Wariancja reszt oblicza na "klasycznie" jako suma kwadratów reszt podzielona przez liczbę stopni swobody nie jest już nieobciążonym estymatorem wariancji składników losowych. Zauważmy, że wariancja reszt KMNK jest zdefiniowana jako

$$\tilde{\sigma}_\varepsilon^2 = \frac{1}{T - K - 1} \sum_{t=1}^T \hat{\varepsilon}_t^2 = \frac{1}{T - K - 1} \hat{\varepsilon}' \hat{\varepsilon} = \frac{1}{T - K - 1} \varepsilon' M \varepsilon$$

gdzie

$$M = I_T - X(X'X)^{-1}X'$$

jest macierzą idempotentną o wymiarach $T \times T$, taką, że $M = M'$ oraz $MM = M$.

Suma kwadratów reszt jest kwadratową formą wektora składników losowych. Obliczając jej wartość oczekiwaną, przy wykorzystaniu własności śladu macierzy otrzymujemy

$$\begin{aligned} \mathbb{E}\hat{\varepsilon}'\hat{\varepsilon} &= \\ &= \mathbb{E}\varepsilon' M \varepsilon = \\ &= \mathbb{E} \text{tr}(\varepsilon' M \varepsilon) = \\ &= \mathbb{E} \text{tr}(M \varepsilon' \varepsilon) = \\ &= \text{tr} \mathbb{E}(M \varepsilon' \varepsilon) = \\ &= \text{tr} M \sigma_\varepsilon^2 \Omega = \\ &= \sigma_\varepsilon^2 \text{tr} M \Omega \end{aligned}$$

Aby wariancja reszt była nieobciążonym estymatorem wariancji składników losowych, wyrażenie powyższe powinno równać się

$$\mathbb{E}\hat{\varepsilon}'\hat{\varepsilon} = \sigma_\varepsilon^2 (T - K - 1)$$

Jest to możliwe wtedy, gdy $\Omega = I_T$. W takim przypadku, z uwagi na własności macierz M zachodzi

$$\begin{aligned} \text{tr} M &= \\ &= \text{tr}(I_T - X(X'X)^{-1}X') = \\ &= T - \text{tr} X(X'X)^{-1}X' = \\ &= T - \text{tr}(X'X)^{-1}X'X = \\ &= T - \text{tr} I_{K+1} = \\ &= T - K - 1 \end{aligned}$$

Jeśli zatem $\Omega \neq I_T$, wtedy $\mathbb{E}\hat{\varepsilon}'\hat{\varepsilon} \neq \sigma_\varepsilon^2 (T - K - 1)$ i w konsekwencji

$$\mathbb{E}\sigma_\varepsilon^2 \neq \mathbb{E} \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-K-1} = \sigma_\varepsilon^2$$

zatem wariancja reszt jest w takim przypadku obciążonym estymatorem wariancji składników losowych. Jeszcze poważniejszy problem związany jest z macierzą wariancji-kowariancji błędów estymacji. Jest ona równa macierzy

$\Sigma_{\hat{\beta}}^2 = \sigma_{\varepsilon}^2 (X'X)$ tylko wtedy, gdy $\Omega = I_T$. Jeśli powyższa równość nie zachodzi, wtedy macierz wariancji-kowariancji przyjmuje postać

$$\Sigma_{\hat{\beta}}^2 = \mathbb{E} \left(\hat{\beta} - \beta \right) \left(\hat{\beta} - \beta \right)' = \sigma_{\varepsilon}^2 (X'X)^{-1} (X'\Omega X)^{-1}$$

Wykorzystując estymator macierzy wariancji-kowariancji, prawdziwy dla sferycznych składników losowych, tj. $\hat{\Sigma}_{\hat{\beta}} = \hat{\sigma}_{\varepsilon}^2 (X'X)^{-1}$, w sytuacji, gdy składniki te nie są sferyczne popełniamy błąd zarówno z tytułu niewłaściwej postaci macierzy wariancji-kowariancji, jak również z tytułu obciążonej estymacji wariancji składników losowych. Na dodatek nie potrafimy w ogólnym przypadku określić jakiego rodzaju obciążenie towarzyszy wykorzystaniu macierzy $\hat{\Sigma}_{\hat{\beta}} = \hat{\sigma}_{\varepsilon}^2 (X'X)^{-1}$ sytuacji niesferycznych składników zakłócających. Wykorzystanie tej macierzy prowadzi do wyznaczenia niepoprawnych średnich błędów ocen parametrów strukturalnych. Sugestia jest zatem następująca: budując model ekonometryczny należy dążyć do wyeliminowania niesferycznych składników losowych (drogą respecyfikacji modelu) lub jeśli nie jest to możliwe, oszacować parametry modelu UMNK, tj. eliminować skutki występowania niesferycznych składników losowych.

6 Wybrane jednowymiarowe modele zmienności cen i stóp zwrotu

Analiza szeregów czasowych rynku finansowego o wysokiej częstotliwości obejmuje

- cechy analizy krótkookresowej (szeregi z tzw. krótką pamięcią)
- podstawowy i uogólniony model grupowania wariancji ARCH
- testowanie efektu ARCH i uogólnionego GARCH
- niestandardowe modele ARCH (tzw. in mean, z symetrią, E-GARCH)
- estymację i wykorzystanie modeli klasy GARCH

6.1 Cechy analizy krótkookresowej

Do cech procesów losowych (najczęściej procesów na rynkach finansowych) charakteryzującą się wysoką częstotliwością zalicza się:

- naprzemienne występowanie okresów o zwiększonej fluktuacji i okresów niskiej zmienności zmiennej finansowej

- skupiania wariancji w kolejnych jednostkach czasu, tj. dodatniej korelacji w dziedzinie zmienności zmiennej będącej przedmiotem zainteresowania, co przejawia się w wysokiej wariancji zmiennej powodowanej wzrostem tej wariancji w okresie poprzedzającym i analogicznie spadkiem wariancji na skutek niskiej wariancji w okresie poprzedzającym.

6.2 Podstawowy i uogólniony model ARCH

Rodzaje nieliniowych procesów stochastycznych.

W nieliniowej analizie jednowymiarowych szeregów czasowych poszukuje się funkcji f wiążącej dany proces z ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie

$$Y_t = f(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$$

gdzie ε_t jest zmienną losową o średniej zero i jednostkowej wariancji. Powyższa reprezentacja jest na tyle ogólna, że nie wiadomo jak dobierać postać funkcji f . Najczęściej przyjmuje się, że nieliniowy proces ekonomiczny ma postać

$$Y_t = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) + \varepsilon_t h(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$$

Procesy Y_t wyrażone z nieliniową funkcją g nazywamy procesami nieliniowymi w warunkowej wariancji.

Powyższa klasyfikacja ma sens, gdyż:

1. Warunkowa wartość oczekiwana Y_t może być zapisana

$$\mathbb{E}(Y_t | \Omega_{t-1}) = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$$

funkcja g opisuje zmiany wartości średniej procesu Y_t warunkowo względem informacji z przeszłości (zbiór Ω_{t-1} oznacza zbiór wszystkich informacji dostępnych do momentu $t-1$).

2. Kwadrat funkcji h przedstawia zmiany warunkowej wariancji procesu Y_t

$$\begin{aligned} D^2(Y_t | \Omega_{t-1}) &= \\ &= \mathbb{E} \left[Y_t - \mathbb{E}(Y_t | \Omega_{t-1})^2 | \Omega_{t-1} \right] = \\ &= \mathbb{E} \left[\varepsilon_t^2 h^2(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) | \Omega_{t-1} \right] = \\ &= h^2(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) \mathbb{E}(\varepsilon_t^2) = \\ &= h^2(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) \end{aligned}$$

Do najbardziej znanych modeli nieliniowych w warunkowej wartości średniej należą: procesy dwuliniowe, nieliniowe procesy autoregresji i średniej ruchomej, autoregresyjne modele progowe, przełącznikowe i wygładzonej przejścia, procesy autogresyjne o losowych współczynnikach. Znanymi procesami o zmiennej wariancji warunkowej są procesy ARCH/GARCH oraz procesy zmienności stochastycznej.

Podstawowym modelem ARCH jest układ dwurównaniowy

$$y_t = x_t^T \beta + \varepsilon \quad h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

gdzie

- x_t jest wektorem zmiennych objaśniających (najczęściej opóźnionych zmiennych endogenicznych postaci modelu autoregresji AR)
- β jest wektorem parametrów strukturalnych
- ε_t jest składnikiem zakłócającym spełniającym warunek

$$\frac{\varepsilon_t}{\Omega_{t-1}} \sim \mathcal{N}(0, h)$$

W celu zapewnienia dodatniości warunkowej wariancji zakłada się ponadto, że $\alpha_0 > 0$ i $\alpha_i \leq 0$. Warto zauważyć, że równanie powyższe jest nieliniowe za względu na zmienne, ale nie ze względu na parametry; równanie to (czyli granica sumy q) wyznacza tzw. stopień modelu ARCH; mówimy wtedy o modelu ARCH(q). Model ARCH(q) opisuje poprawnie proces stacjonarny (lub inaczej ARCH(q) generuje proces stacjonarny), jeśli spełniony jest warunek

$$\sum_{i=1}^q \alpha_i < 1$$

Uogólnionym modelem ARCH, czyli modelem GARCH, jest

$$y_t = x_t^T \beta + \varepsilon_t \quad h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i h_{t-i}$$

gdzie oznaczenia zmiennych i parametrów jak w równaniach poprzednich.

W celu zapewnienia dodatniości warunkowej wariancji zakłada się ponadto, że $\alpha_0 > 0$ i $\alpha_i \leq 0$ i $\gamma_1 \leq 0$. Granice sumowania q i p wyznaczają stopień modelu GARCH; mówimy, o modelu GARCH(q, p).

Stacjonarność procesu (czyli skończoność bezwarunkowej wariancji) opisanego modelem GARCH(q,p) jest zapewniona, jeśli spełniony jest warunek

$$\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \gamma_i < 1$$

W zastosowaniach modelu rynków finansowych wygodnie jest korzystać z tzw. reprezentacji równoważnej modelu GARCH. Niech dany będzie proces v_t taki, że

$$\begin{aligned} v_t &= \varepsilon_t^2 - h_t = (\eta_t^2 - 1)h_t \\ h_t &= \varepsilon - t^2 - v_t \\ \eta_t &\sim \mathcal{N}(0, 1) \end{aligned}$$

Z formuły wyrażającej warunkową wariancję w modelu GARCH wiemy, że h_t należy zapisać jako

$$\begin{aligned} h_t &= \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i h_{t-i} \Leftrightarrow \\ &\Leftrightarrow h_t = \varepsilon_t^2 - v_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i h_{t-i} \\ \varepsilon_t^2 - v_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i (\varepsilon_{t-i}^2 - v_{t-i}) \\ \varepsilon_t^2 - v_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i \varepsilon_{t-i}^2 - \sum_{i=1}^p \gamma_i v_{t-i} \\ \varepsilon_t^2 &= \alpha_0 + \underbrace{\sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i \varepsilon_{t-i}^2}_{AR(\max(p,q))} - \underbrace{\sum_{i=1}^p \gamma_i v_{t-i}}_{MA(p)} + v_t \end{aligned}$$

6.3 Testowanie efektu ARCH/GARCH

Testowanie efektu ARCH/GARCH jest ekwiwalentne, tj. istniejące testy nie pozwalają odróżnić obu procesów. Wynik testu wskazujący na obecność omawianego efektu pozwala jedynie z określonym prawdopodobieństwem wnioskować o obecności ARCH lub GARCH, bez możliwości rozróżnienia. Wnioskowanie o rzędach p i q procesów ARCH/GARCH odbywa się na podstawie miar pojemności informacyjnej.

6.3.1 TEST ENGLE'A

Jest to test typu mnożnika Lagrange'a, tzn. do testowania hipotezy zerowej niezbędna jest znajomość jedynie modelu z restrykcjami nałożonymi poprzez testową hipotezę. Przypomnijmy, że równaniem pomocniczym wariancji warunkowej w modelu ARCH jest

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

Engle zaproponował postać modelu AR dla kwadratów reszt uzyskanych z tej relacji, zatem szacowany (KMNK lub MNK) model przyjmuje postać

$$\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \hat{\varepsilon}_{t-i}^2 + \mu_t$$

Jeśli efekt ARCH/GARCH nie występuje, czyli nie występuje heteroskedastyczność wariancji warunkowej, wówczas wszystkie parametry α_i powinny zanikać, tak więc hipotezami są

$$\begin{aligned} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0 \\ H_A : \exists \alpha_i \neq 0 \end{aligned}$$

Statystyką testową jest typowa statystyka Lagrange'a

$$LM = T \cdot R^2$$

gdzie R^2 jest współczynnikiem determinacji wyznaczonym dla modelu kwadratów reszt. Statystyka testowa LM ma rozkład graniczny χ^2 o q stopniach swobody; wnioskowanie o odrzuceniu H_0 lub braku podstaw do odrzucenia jest typowe dla testu prawostronnego.

6.3.2 TEST MCLEODA I LI

W omawianym teście wykorzystuje się statystykę Ljunga-Boxa do weryfikacji hipotezy o braku autokorelacji kwadratów reszt, zatem test przebiega następująco

1. Oszacować relację KMNK
2. Wyznaczyć kwadraty reszt $\hat{\varepsilon}_t^2$
3. Obliczyć współczynniki autokorelacji (rzędu od 1 do q) kwadratów reszt uzyskanych w punkcie poprzednim

$$\hat{\rho} = \frac{\text{cov}(\hat{\varepsilon}_t^2, \hat{\varepsilon}_{t-i}^2)}{D^2(\hat{\varepsilon}_t^2)} = \frac{\frac{1}{T-i} \sum_{t=i+1}^T \hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-i}^2}{\frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2}$$

4. Obliczyć statystykę Ljunga-Boxa

$$Q' = T(T+2) \sum_{i=1}^q \frac{\rho_i}{T-i}$$

5. Statystyka Ljunga-Boxa ma rozkład graniczny χ^2 o q stopniach swobody

6. Wobec zastosowanej statystyki testowej, zestawem hipotez jest

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_q = 0$$

$$H_A : \exists \rho_i \neq 0$$

6.4 Niestandardowe modele GARCH

6.4.1 MODEL GARCH-M

Model GARCH in MEAN (GARCH-M) przyjmuje postać

$$y_t = x_t^T \beta + \lambda h_t + \varepsilon_t \quad \frac{\varepsilon_t}{\Omega_{t-1}} \sim \mathcal{N}(0, h_t)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i h_{t-i}$$

GARCH-M stosowany jest do modelowania ryzyka (premii za ryzyko). Jeżeli ocena parametru λ jest dodatnia i parametr może zostać uznany za statystycznie istotny, wówczas wzrost wariancji warunkowej h_t (czyli miary ryzyka) powoduje wzrost premii za ryzyko w postaci oczekiwanej stopy zwrotu z papieru (y_t).

6.4.2 MODEL GARCH Z ASYMETRIĄ REAKCJI

Asymetria reakcji na "pakietowe" zmiany zmienności zmiennej będącej przedmiotem zainteresowania (np. stopa zwrotu t_t) może być przybliżona prostym modelem GARCH-M

$$r_t = a + \lambda h_t + \varepsilon_t \quad \frac{\varepsilon_t}{\Omega_{t-1}} \sim \mathcal{N}(0, h_t)$$

$$\varepsilon_t = v_t \sqrt{h_t} \quad v_t \sim \mathcal{N}(0, 1)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i h_{t-i}$$

Wówczas możliwe jest wyznaczenie v_t jako

$$v_t = \frac{r_t - (a + \lambda h_t)}{\sqrt{h_t}}$$

Warto zauważyć, że proces opisany przez to równanie charakteryzuje się rozkładem normalnym standardowym.

6.4.3 MODEL E-GARCH

W modelu E-GARCH czyni się typowe założenia odnoszące się do równania opisującego zmienną będącą przedmiotem zainteresowania, czyli

$$\begin{aligned} y_t &= x_t^T \beta + \varepsilon & \frac{\varepsilon_t}{\Omega_{t-1}} &\sim \mathcal{N}(0, h - t) \\ \varepsilon_t &= v_t \sqrt{h_t} & v_t &\sim \mathcal{N}(0, 1) \end{aligned}$$

Funkcją warunkową wariancji jest

$$\ln h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \left(\delta_1 v_{t-1} + \delta_2 \left(v_{t-1} - \sqrt{\frac{2}{\pi}} \right) \right) + \sum_{i=1}^p \gamma_i \ln h_{t-1}$$

Powyższy model jest modelem typu wykładniczego. Z definicji funkcji wykładniczej wynika, że zapewniona jest nieujemność wariancji warunkowej.

7 Ekonomiczne modele panelowe

Definicja danych panelowych

dane przekrojowe	y_i	$i = 1, \dots, N$
szeregi czasowe	y_t	$t = 1, \dots, T$
dane przekrojowo-czasowe	y_{it}	$i = 1, \dots, N, \quad t = 1, \dots, T$

7.1 Dane panelowe

7.1.1 PRZYKŁADY DANYCH PANELOWYCH.

- Dane miesięczne o wydatkach dla 1000 gospodarstw domowych obserwowanych przez 5 lat - analiza preferencji konsumentów, badania skuteczności kompanii reklamowej.
- Makroekonomiczne wskaźniki 25 krajów UE w ostatnich 10 lat publikowane przez Eurostat - badanie skuteczności europejskich funduszy strukturalnych.

- Penn World Tables - skonstruowana przez Summersa i Hestona baza porównywalnych danych makroekonomicznych zawierająca informacje dla ponad 200 krajów dla 1960-2010 - analiza czynników wzrostu gospodarczego, weryfikacja hipotezy konwergencji gospodarczej.
- Wskaźniki zatrudnienia i wynagrodzeń w Polsce w podziale na województwa za ostatnie 15 lat - badanie skuteczności polityki makroekonomicznej, szukanie czynników determinujących zmiany na rynku pracy.
- Liczba inwestorów zagranicznych inwestujących w poszczególnych województwach w latach 1993-2015 - modelowanie częstotliwości występowania określonego zdarzenia - badanie atrakcyjności inwestycyjnej poszczególnych województw.
- Dane firmy ubezpieczeniowej o ubezpieczeniach komunikacyjnych dotyczące np. szkodowości klientów w różnych latach - np. możliwość wykorzystania wyników do budowy systemu Bonus-Malus.
- Informacje banków o kredytach konsumpcyjnych, hipotecznych i ich spłacalności na przestrzeni kilkunastu miesięcy lub lat - budowa modeli scoringowych - modeli wstępnej oceny ryzyka kredytowego.
- Informacje firmy odzieżowej o przychodach i kosztach poszczególnych sklepów w sieci w różnych miastach lub centrach dla szeregu miesięcy - ocena skuteczności poszczególnych kampanii reklamowych, szczegółowa analiza indywidualnej specyfiki poszczególnych sklepów lub miast.

7.1.2 DLACZEGO WARTO STOSOWAĆ DANE PANELOWE

- Dane panelowe pozwalają na analizę zjawiska równocześnie w czasie jak i wymiarze przekrojowym lub przestrzennym.
- Dane panelowe pozwalają na wyodrębnienie indywidualnej specyfiki poszczególnych obiektów.
- Zastosowanie paneli danych pozwala na większą heterogeniczność (większe zróżnicowanie) jednostek badania.
- Zapewnia większą liczbę stopni swobody oraz zwiększa efektywność oszacowania. Wyodrębnienie efektów wpływu nieobserwowalnych zmiennych lub efektów.

7.1.3 SPECYFIKA DANYCH PANELOWYCH

- Obserwacje dotyczące tej samej jednostki mogą być ze sobą skorelowane.
- Obserwacje dotyczące tego samego okresu mogą być ze sobą skorelowane
- Czasem mamy do czynienia z przypadkami, gdzie nie występuje ani jeden, ani drugi rodzaj korelacji (przypadek akademicki)

7.2 Ogólny zapis modelu panelowego

Model statyczny

$$Y_{it} = \beta_0 + \beta' x_{it} + \alpha_{it} + v_t + u_{it}$$

gdzie

- β - wektor parametrów
- x_{it} macierz obserwacji na zmiennych objaśniających
- α_i - efekty indywidualne, część zmienności y charakterystyczna dla efektu i -tej jednostki (N efektów)
- v_t efekty okresowe, część zmienności zmiennej y charakterystyczna dla okresu t ; (T efektów)
- u_{it} czysto losowy składnik zakłócający

Rodzaje modeli panelowych

- 1a) Jeżeli efekty indywidualne α_i i efekty v_t są nieistotne, oznacza to, że mamy do czynienia z tak zwanym homogenicznym panelem - analizowana przez nas relacja jest taka sama w każdym okresie i dla każdej badanej jednostki - Model Łącznej Regresji.
- a2) Jeżeli zarówno efekty indywidualne jak i efekty okresowe są istotne, mamy do czynienia z heterogenicznym panelem - Dwukierunkowy Model Panelowy.
- a3) Jeżeli tylko jedna grupa efektów jest istotna - Jednokierunkowy Model Panelowy

- b1) Efekty indywidualne i (lub) okresowe mogą być efektami ustalonymi, czyli stałymi w danym czasie lub dla danej jednostki, nie zależą od czynników losowych - Model z Efektami Ustalonymi (ang. *Fixed Effects Model*) - FE
- b2) Efekty indywidualne i (lub) okresowe mogą być efektami losowymi, czyli zależą od czynników losowych i mogą się zmieniać, są zmiennymi losowymi o określonych rozkładach - Model z Efektami Losowymi (ang. *Random Effects Model*) - RE.

Pozornie Niezależne Regresje

Czasami dysponując danymi panelowymi nie jesteśmy zainteresowani budową jednego modelu ekonometrycznego dla wszystkich posiadanych informacji. Szacowane są wówczas oddzielne regresje dla każdej jednostki badania, ale dla podniesienia dokładności (efektywności) oszacowania, wykorzystana jest dodatkowa informacja wynikająca z faktu, że dla wszystkich jednostek badania można określić pewną strukturę stochastyczną; stosujemy Model Pozornie Niezależnych Regresji (ang. *Seemingly Unrelated Regression Model*) - SUR

7.3 Model regresji Łącznej

Założenia

1. Mamy do czynienia z homogeniczną próbą - wszystkie jednostki badania są do siebie podobne. Oznacza to, że parametry szacowanego modelu są jednakowe dla wszystkich jednostek oraz we wszystkich okresach badania.
2. Różnice między empirycznymi wartościami zmiennej endogenicznej, a wartościami teoretycznymi są jedynie efektem działania zakłóceń losowych o tym samym rozkładzie dla każdej i -tej jednostki i dla każdego okresu t .

Wybór takiego modelu może być poprzedzony odpowiednim testowaniem. Model możemy zapisać jako

$$y_{it} = \beta_0 + \sum_{s=1}^k \beta_{sit} x_{sit} + u_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

gdzie:

- N - liczba jednostek w próbie

- T - liczba okresów
- k - liczba zmiennych objaśniających w modelu

Metody estymacji: MNK lub KMNK w przypadku niesferycznego składnika losowego.

7.4 Model z Efektami Ustalonymi (FE)

Jednokierunkowy (*one-way*)

Model możemy zapisać

$$y_{it} = \alpha_i + \beta' x_{it} + u_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

β - wektor parametrów estymacji

Zapis macierzowy

$$y_i = \alpha_i e + X_i \beta + u_i \quad i = 1, \dots, N$$

Założenia do estymacji

- u_i oraz X_i są niezależne dla $i = 1, \dots, N$ - zmienne tworzące macierz X są ściśle egzogeniczne, czyli

$$\forall_{i=1, \dots, N} \mathbb{E} \{x_{it}, u_{rs}\} = 0$$

- $\mathbb{E}(u_i) = 0$
- $\mathbb{E}(u_i, u'_j) = \sigma^2 I$, $i = 1, \dots, N$, $j = 1, \dots, N$ gdzie I jest macierzą jednostkową
- $r(X) = N + K \leq NT$

Wprowadzamy N zmiennych zerojedynkowych - każda zmienna przyjmuje wartość 1 dla danej jednostki i 0 dla pozostałych jednostek.

Szacowanie:

- KMNK
- Estymator Wewnątrzgrupowy

7.5 Model z Efektami Losowymi (RE)

Jednokierunkowy (*one-way*)

Ogólna postać modelu

$$y_{it} = \beta_0 + \beta' x_{it} + v_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

gdzie

- $v_{it} = u_{it} + \alpha_i$ - dwuczęściowy składnik losowy

Dodatkowe założenia

- $\mathbb{E}(\alpha_i) = 0, \quad i = 1, \dots, N$
- Efekty indywidualne α_i są niezależne od u_i oraz X_i dla $i = 1, \dots, N$
- Efekty indywidualne dla różnych jednostek są nieskorelowane, ale występuje korelacja między efektami z różnych okresów dla tej samej jednostki.

$$\mathbb{E}(\alpha_i, \alpha_j) = \begin{cases} 0 & i \neq j \\ \sigma_\alpha^2 & i = j \end{cases} \quad i, j = 1, \dots, N$$

Wariancja składnika losowego w modelu RE składa się z dwóch części

$$\sigma_v^2 = \sigma_u^2 + \sigma_\alpha^2$$

gdzie

- σ_u^2 - wariancja składnika losowego
- σ_α^2 - wariancja efektów indywidualnych

Uwaga!

Składniki losowe w modelu RE są ze sobą skorelowane, czyli KMNK przestaje być efektywna, dlatego do oszacowania tego modelu zastosować estymator Uogólnionej Metody Najmniejszych Kwadratów.

Macierz wariancji i kowariancji

$$V_{(NT \times NT)} = \begin{bmatrix} \Omega_1 & 0 & \dots & 0 \\ 0 & \Omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Omega_N \end{bmatrix} = I_N \otimes \Omega_i$$

gdzie macierz Ω_i jest w przypadku stałych efektów indywidualnych macierzą jednostkową rzędu T , a w przypadku losowym efektów indywidualnych macierzą

$$\Omega_t = \mathbb{E}(v_i v_i') = \begin{bmatrix} (\sigma_u^2 + \sigma_\alpha^2) & \sigma_\alpha^2 & \dots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & (\sigma_u^2 + \sigma_\alpha^2) & \dots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \dots & (\sigma_u^2 + \sigma_\alpha^2) \end{bmatrix} = \sigma_u^2 I_T + \sigma_\alpha^2 i i'$$

gdzie $v_i = \alpha_i + u_i$

7.6 Niespełnienie założeń

7.6.1 HETEROSKEDASTYCZNOŚĆ

- *groupwise heteroskedasticity* - macierze Ω_i są różne dla poszczególnych jednostek, co może być spowodowane zmiennością czysto losowego składnika lub (w modelu random effects) zmiennością wariancji efektów indywidualnych.
- *cross-sectional correlation* - w macierzy blokowej V poza przekątną występują macierze niezerowe; sytuację taką najczęściej opisuje się poprzez współczynnik korelacji równoczesnych realizacji "czystego" zaburzenia losowego.

7.6.2 AUTOKORELACJA

$$\mathbb{E}(u_{it}, u_{jt} | X_i, X_j) = \rho_{ij}$$

mogą wystąpić np. procesy autoregresyjne i/lub średniej ruchomej

$$\mathbb{E}(u_{it}, u_{is} | X_i) = \rho_{ts}$$

7.7 Konsekwencje niespełnienia założeń

Jeśli założenia o strukturze wariancji składnika losowego są błędne, to:

- Estymatory modelu po zostają nieobciążone, zgodne i asymptotycznie normalne, lecz są nieefektywne w porównaniu z estymatorem uwzględniającym strukturę prawdziwą.
- Dodatkowo, obciążony jest estymator macierzy wariancji kowariancji parametrów, zatem mogą okazać się błędne wyniki o istotności zmiennych wyciągane na podstawie statystyki t .

7.8 Sposoby rozwiązania problemu

- Zastosowanie UMNK z estymacją:
 - dla modelu, w którym autokorelacja rzędu pierwszego: np. Cochran-Orcutta
 - dla modelu ze zmienną wariancją tzw. Ważona Metoda Najmniejszych Kwadratów
- Estymacja macierzy wariancji i kowariancji parametrów (czyli również błędów szacunku parametrów strukturalnych) w taki sposób, że byłaby odporna (roboust) na niespełnienie założeń o sferyczności składnika losowego.

7.9 Estymatory odporne (roboust)

1. Metoda zaproponowana przez Arellana (duże N , małe T). Macierz kowariancji i wariancji może być szacowana jako

$$\Sigma_{\hat{\beta}} = NT(X'X)^{-1}\hat{\Omega}(X'X)^{-1}$$

gdzie $\hat{\Omega}$ jest zgodnym estymatorem macierzy wariancji składnika losowego; dla wszelkich postaci heteroskedastyczności i autokorelacji "czystego składnika losowego u_{it} ma postać

$$\hat{\Omega} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2 x_{it} x_{it}'$$

gdzie

- u_{it} - reszty modelu
- x_{ti} - obserwacje na zmiennych objaśniających.

Jest to tzw. korekta White'a uogólniona do danych panelowych.

2. Metoda Arellano niedoszacowuje błędów w przypadku wystąpienia znaczącej autokorelacji zakłóceń pomiędzy jednostkami w tym samym momencie czasu. W takich przypadkach stosuje się

$$\Sigma_{\hat{\beta}} = (X'X)^{-1} \left(\sum_{t=1}^N \sum_{j=1}^N \hat{\sigma}_{ij} (X_i' X_j)^{-1} \right) (X'X)^{-1}$$

gdzie $\hat{\sigma}_{ij} = \frac{\hat{u}_i' \hat{u}_j}{T}$

T Oznacza długość szeregu czasowego dla każdej jednostki.

7.9.1 HETEROSKEDASTYCZNOŚĆ

Dwa źródła

- duże zróżnicowanie jednostek w panelu
- duża zmienność zakłóceń losowych w czasie

Przy dużym N i małym T uwzględnienie zmienności wariancji w modelu RE wiąże się z estymacją zbyt dużej liczby parametrów, stąd w takim przypadku powinno się stosować model FE. Wówczas macierz wariancji i kowariancji składników losowych

$$V_{(NT \times NT)} = \Sigma_{(N \times N)} \otimes I_T$$

Model taki możemy szacować Uogólnioną Metodą Najmniejszych Kwadratów

$$\hat{\beta}_{UMNK} = (x' \hat{V}^{-1} X)^{-1}$$

A elementy macierzy wariancji i kowariancji parametrów

$$\Sigma_{\hat{\beta}} = (X' \hat{V}^{-1} X)^{-1}$$

a elementy macierzy \hat{V} obliczamy na podstawie reszt z oszacowania wstępnej regresji Klasyczną Metodą Najmniejszych Kwadratów

$$\hat{\sigma}_I^2 = \frac{\hat{u}_i' \hat{u}_i}{T} \qquad \hat{\sigma}_{ij} = \frac{\hat{u}_i' \hat{u}_j}{T}$$

7.10 Model RE a model FE

- Brak różnicy w estymacji, jeśli T dąży do nieskończoności
- Jeśli efekty indywidualne nie są skorelowane za zmiennymi objaśniającymi, to zarówno FE jak i RE są nieobciążone i zgodne, ale RE jest efektywniejszy
- Jeśli efekty indywidualne są skorelowane za zmiennymi objaśniającymi, to FE jest nieobciążony, a RE obciążony.

7.11 Model z efektami ustalonymi (FE)

Dwukierunkowy (*two-way*)

Jeżeli efekty okresowe (czasowe) mają być efektami ustalonymi to wyraz wolny dzielony jest na dwie części. Możemy zatem zapisać:

$$y_{it} = \alpha_i + \lambda_t + \beta'x_{it} + \varepsilon_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

gdzie

- α_i - ustalony efekt indywidualny stały w każdym okresie czasu, ale inny (może być inny) dla każdego obiektu w panelu
- λ_t - ustalony efekt, który ma tę samą wartość dla wszystkich jednostek w panelu w tym samym okresie, ale jest różny (może być różny) w każdym czasie.

Należy pamiętać, nie możemy wprowadzić równocześnie N efektów indywidualnych oraz T efektów okresowych, jeśli chcemy oszacować model \rightarrow dokładana współliniowość wektorów.

7.12 Testowanie modeli FE, RE oraz FE vs RE

- model FE: test Walda istotności efektów istotności efektów indywidualnych tzw. testowanie różnicy w sumie kwadratów reszt dla modelu *pooled* (hipoteza zerowa) vs FE (hipoteza alternatywna)
- model RE: test Breuscha-Pagana (wersja testu mnożników Lagrange'a) istotności wariancji przy prawdziwości hipotezy zerowej wariancja jest stała w czasie dla wszystkich badanych jednostek
- FE vs RE: test hausmana; testowanie są różnice między oszacowaniami parametrów obu modeli; duże różnice (hipoteza alternatywna) mogą być spowodowane tym, że RE jest obciążony, czyli efekty indywidualne są skorelowane ze zmiennymi objaśniającymi lub nastąpił błąd w specyfikacji modelu; małe różnice (hipoteza zerowa) wskazują, że FE i RE są zgodne...

7.13 Istotność efektów stałych w modelu FE

Pytanie: czy wprowadzanie różnych dla poszczególnych wyrazów wolnych prowadzi do uzyskania dokładniejszych oszacowań?

7.13.1 TEST WALDA (F)

$$\begin{aligned} H_0 : \alpha_{it} &= \text{const}, & i &= 1, \dots, N & t &= 1, \dots, T \\ H_A : \forall_{i,j} \alpha_i &\neq \alpha_j, \text{ ale } \alpha_{it} &= \alpha_{is} &= \alpha_i & i &= 1, \dots, N & t, s &= 1, \dots, T \end{aligned}$$

Według hipotezy zerowej wyrazy wolne (wszystkich jednostek i okresów) mają tę samą własność. Według hipotezy alternatywnej wyrazy wolne są stałe w czasie, lecz mogą różnić się dla poszczególnych jednostek.

Statystyka testowa z próby

$$F = \frac{(R_1^2 - R_0^2)/(N - 1)}{(1 - R_1^2)/(NT - N - k)}$$

gdzie:

- R_0^2 - wartość współczynnika determinacji dla modelu prawidłowego według hipotezy zerowej, czyli

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T$$

- R_1^2 - prawidłowa wartość współczynnika determinacji dla modelu poprawionego według hipotezy alternatywnej, czyli

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T$$

Można również skorzystać ze statystyki testowej

$$F = \frac{(S_1^2 - S_0^2)/(N - 1)}{(1 - S_1^2)/(NT - N - k)}$$

gdzie:

- S_0^2 - suma kwadratów reszt dla modelu prawidłowego według hipotezy zerowej,
- S_1^2 - dla modelu prawidłowego według hipotezy alternatywnej.

W obu przypadkach obszar krytyczny testu jest prawostronnie określony przez wartość krytyczną

$$F_\alpha = (N - 1, NT - N - k)$$

7.14 Istotności efektów okresowych w FE

Pytanie: czy właściwy jest model jednokierunkowy (*one-way*), czy dwukierunkowy (*two-way*)?

$$\begin{aligned} H_0 : \gamma_{it} = \gamma_{is} = \gamma_i, & \quad i = 1, \dots, N & \quad t, s = 1, \dots, T \\ H_A : \forall_{i,j} \gamma_i \neq \gamma_j, & \quad i = 1, \dots, N & \quad t, s = 1, \dots, T \end{aligned}$$

gdzie:

- $\gamma_{it} = \alpha_i + \lambda_t$

Według hipotezy zerowej jest model jednokierunkowy, a według hipotezy alternatywnej model dwukierunkowy.

Statystyka z próby

$$F = \frac{(S_0^2 - S_1^2)/(T - 1)}{S_1^2/[NT - (N - 1) - (T - 1) - k]}$$

7.15 Istotność efektów w modelu RE

W modelu z efektami losowymi zakładamy, że składnik losowy zawiera w sobie zarówno efekty indywidualne jak i okresowe. Należy zbadać, czy wariancja składników losowych dla wszystkich obserwacji jest stała. Jeżeli jest stała, tzn. że brak istotnego zróżnicowania efektów i należy zastosować model regresji łącznej.

7.15.1 TEST BREUSCH-PAGANA (JAK W TEŚCIE MNOŻNIKÓW LAGRANGE'A)

$$\begin{aligned} H_0 : \sigma_\alpha^2 &= 0 \\ H_A : \sigma_\alpha^2 &\neq 0 \end{aligned}$$

Statystyka testowa z próby

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N \left(\sum_{t=1}^T e_{it} \right)^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} - 1 \right]^2$$

gdzie wartości reszt e_{it} to reszty z modelu regresji łącznej. Przy prawidłowości hipotezy zerowej powyższa statystyka ma rozkład χ^2 z jednym stopniem swobody.

7.16 Czy stosować model FE czy raczej RE?

- Jakie założenia można przyjąć co do efektów indywidualnych?
Czy można je ocenić jako nieprzypadkowe (systematycznie) związane z poszczególnymi jednostkami?
- Testowanie statystyczne - czy są spełnione przyjęte założenia?
 - Kluczowym założeniem dla modelu RE jest założenie o niezależności efektów indywidualnych i zmiennych objaśniających.
 - Jeżeli założenie to nie jest spełnione wówczas estymator UMNK staje się estymatorem obciążonym.
 - W przypadku spełnienia założenia zarówno estymator UMNK jak i estymator KMNK dla modelu FE są zgodne i nieobciążone, jednak estymator UMNK jest bardziej lub co najmniej tak samo efektywny.

7.16.1 TEST HAUSMANA

Polega na porównaniu wartości ocen parametrów uzyskanych przy pomocy obu estymatorów.

$$H_0 : \mathbb{E}(\xi_{it}|X) = 0$$

$$H_A : \mathbb{E}(\xi_{it}|X) \neq 0$$

gdzie:

- $\xi_{it} = \eta_i + u_{it}$
- H_0 - oba estymatory są zgodne i nieobciążone, ale estymator UMNK dla modelu RE jest bardziej efektywny.
- H_A - estymator UMNK dla RE jest obciążony, zatem należy stosować model FE, którego estymator jest nieobciążony

Statystyka testowa z próby

$$m_1 = (\hat{\beta}_{RE} - \hat{\beta}_{FE})^T (\text{Var}(\hat{\beta}_{FE}) - \text{Var}(\hat{\beta}_{RE}))^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE})$$

gdzie:

- $\hat{\beta}_{RE}$ - wektor ocen parametrów z modelu RE
- $\hat{\beta}_{FE}$ - wektor ocen parametrów z modelu FE

- $\text{Var}(\hat{\beta}_{RE})$ - macierz wariancji i kowariancji ocen parametrów modelu RE
- $\text{Var}(\hat{\beta}_{FE})$ - macierz wariancji i kowariancji ocen parametrów modelu FE

Jeżeli hipoteza H_0 jest prawdziwa to statystyka m_1 ma rozkład χ^2 z k stopniami swobody (prawostronny obszar krytyczny).

7.16.2 MACIERZ WARIANCJI I KOWARIANCJI PARAMETRÓW

Ogólna formuła wyznaczania macierzy wariancji i kowariancji parametrów strukturalnych

$$\text{Var}(\hat{\beta}) = \Sigma_{\hat{\beta}} = \mathbb{E}\left((\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

gdzie Ω jest macierzą wariancji i kowariancji składników losowych. Tylko wtedy, gdy $\Omega = \sigma^2 I$ redukuje się do:

$$\text{Var}(\hat{\beta}) = \Sigma_{\hat{\beta}} = \mathbb{E}\left((\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right) = \sigma_u^2 (X'X)^{-1}$$

7.16.3 WSPÓŁCZYNNIK DETERMINACJI W ANALIZIE

Dla modeli panelowych definiujemy różne współczynniki determinacji

- ogólny współczynnik determinacji

$$R^2 = r^2 \{ \hat{y}_{it}, y_{it} \}, \quad i = 1, \dots, N \quad t = 1, \dots, T$$

- wewnątrzgrupowy współczynnik determinacji

$$R^2 = r^2 \{ \hat{y}_{it} - \bar{\hat{y}}_{it}, y_{it} - \bar{y}_{it} \}, \quad i = 1, \dots, N \quad t = 1, \dots, T$$

gdzie

$$\begin{aligned} \bar{\hat{y}}_{it} &= \frac{1}{T} \sum_{t=1}^T \hat{y}_{it} \\ \bar{y}_{it} &= \frac{1}{T} \sum_{t=1}^T y_{it} \end{aligned} \quad i = 1, \dots, N$$

- międzygrupowy współczynnik determinacji

$$R_{BG}^2 = r^2 \{ \bar{\hat{y}}_{it}, \bar{y}_{it} \}$$

Wszystkie współczynniki wyznaczone są jako kwadraty współczynnika korelacji liniowej Pearsona i przyjmuje wartości z przedziału $[0, 1]$

Wartości przedstawionych współczynników mogą istotnie się różnić w zależności od postaci wybranego estymatora.

- Ogólny współczynnik determinacji przyjmuje z reguły największe wartości dla modelu regresji łącznej.
- Zastosowanie modelu FE pozwala na maksymalizację R^2 wewnątrzgrupowego.
- W regresji międzygrupowej uzyskamy najwyższe wartości współczynnika determinacji międzygrupowego.

7.16.4 SZACOWANIE MODELU Z EFEKTAMI USTALONYMI (FE)

Model FE ze zmiennymi zerojedynekowymi generalnie szacujemy KMNK (tzw. *LSDV - LS with Dummy Variables*), pojawia się jednak często problem zbyt dużej liczby zmiennych w modelu - czyli zbyt dużego wymiaru macierzy X . Rozwiązaniem (dla dużej liczby jednostek N) jest Estymator Wewnątrzgrupowy.

7.16.5 ESTYMATOR WEWNĄTRZGRUPOWY (*Within estimator*)

Zdefiniujemy następujące średnie wartości zmiennych dla każdej i -tej jednostki:

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it} \qquad \bar{x}_{ik} = \frac{1}{T} \sum_{t=1}^T x_{itk}$$

Mamy zatem szereg średnich, każdy po N obserwacji. Możemy zapisać wówczas następujący model

$$\bar{y}_i = \alpha_i + \bar{x}_i \beta + \bar{u}_i \qquad i = 1, \dots, N \quad (1)$$

Model taki nazywany jest regresją międzygrupową (*between group regression*) i jest modelem przekrojowym, gdzie wartości zmiennych są średnimi po czasie. Zwróćmy uwagę, że średnia wartość efektu indywidualnego dla i -tej jednostki jest równa α_i , ponieważ efekt ten jest stały w czasie. Jeżeli zapiszemy informacje dla jednej jednostki ze wszystkich okresów w jeden wektor możemy zapisać, że

$$y_i = \alpha_i e + X_i \beta + u_i \quad (2)$$

Jeżeli od równania (2) odejmiemy równanie (1) otrzymamy wówczas

$$(y_i - \bar{y}_i) = (\alpha_i - \alpha_i) + (x_i - \bar{x}_i)\beta + (u_i - \bar{u}_i)$$

Zwróćmy uwagę, że poprzez taką transformację z modelu "znika" efekt indywidualny. Dzieje się tak dla każdej badanej jednostki. Poprzez transformację pierwotnych danych do postaci odchyleń od średnich uzyskujemy nowe zmienne:

$$\begin{aligned}\tilde{y}_{it} &= (y_{it} - \bar{y}_i), \\ \tilde{x}_{it} &= (x_{it} - \bar{x}_i)\end{aligned}$$

Wówczas stosujemy KMNK dla przetransformowanych danych

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^N (\tilde{X}_i' \tilde{X}_i) \right)^{-1} \sum_{i=1}^N (\tilde{X}_i' \tilde{Y}_i)$$

Jest to tzw. estymator wewnątrzgrupowy (*within group*). Może być zapisany również skalarnie jako

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)'$$

Własności:

- Estymator wewnątrzgrupowy daje identyczne oceny parametrów jak MNK za zmiennymi zerojedynkowymi
- Nie jest konieczne szacowanie N dodatkowych parametrów reprezentujących efekty indywidualne, jeżeli nie są one przedmiotem zainteresowania
- Jeżeli zainteresowani jesteśmy konkretnymi wartościami efektów indywidualnych, mogą być one oszacowane następująco

$$\hat{\alpha}_i = \bar{y}_i - \bar{x}_i' \hat{\beta}_{FE}$$

- Estymator efektów indywidualnych jest zgodny tylko wówczas, gdy ilość obserwacji po czasie jest duża. $T \rightarrow \infty$.

Estymator macierzy wariancji-kowariancji definiujemy jako

$$\text{Var}(\hat{\beta}_{FE}) = \sigma_u^2 \left(\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1}$$

Nieznana wartość wariancji zakłóceń losowych σ_u^2 zastępujemy wartością oszacowaną

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2}{NT - N - k} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\alpha}_i - x_{it}; \hat{\beta}_{FE})^2}{NT - N - k}$$

gdzie:

- e_{it} - reszty oszacowanego przy pomocy estymatora wewnątrzgrupowego.

Z diagonalnych elementów macierzy $\text{Var}(\hat{\beta}_{FE})$ wyliczamy błędy szacunku estymatora FE. Estymator wariancji ocen parametrów jest zgodny, gdy $N \rightarrow \infty$ lub $T \rightarrow \infty$.

Uwaga!

Estymator wewnątrzgrupowy nie może być wykorzystany do modelu, w którym wśród zmiennych objaśniających występują czynniki, które nie zmieniają się w czasie (marka i rodzaj produktu, płeć, pochodzenie, miejsce zamieszkania, położenie geograficzne itp.) Czynniki te są współliniowe z efektami indywidualnymi (nie mogą wchodzić razem do macierzy X). W transformacji *within* są eliminowane z modelu tak, jak efekty indywidualne (problem ten został rozwiązany poprzez zastosowanie tzw. estymatorów Hausmana-Taylora).

7.16.6 ESTYMATOR WEWNĄTRZGRUPOWY (*within* DLA MODELU DWUKIERUNKOWEGO)

Rozważmy cztery różne modele

Model (a):- Panelowy model z dekompozycją wyrazu wolnego na dwie części:

$$y_{it} = \alpha_i + \lambda_t + \beta' x_{it} + u_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

Model (b):- Model uśredniony po czasie - *between regression*

$$\bar{y}_t = \alpha_i + \bar{\lambda}_t + \beta' \bar{x}_i + \bar{u}_i \quad i = 1, \dots, N$$

Model (c):- Model uśredniony po jednostkach

$$\bar{y}_t = \bar{\alpha}_i + \lambda_t + \beta' \bar{x}_t + \bar{u}_t \quad t = 1, \dots, T$$

Model (d):- Obliczamy średnie wartości dla wszystkich obserwacji (łącznie po czasie i po jednostkach)

$$\bar{y}_{it} = \bar{\alpha}_i + \bar{\lambda}_t + \beta' \bar{x} + \bar{u}_{it}$$

(Zauważmy, że model (d) jest tylko pojedynczym równaniem)
 Następnie wykonujemy następujące działania na poszczególnych równaniach stronami:

$$(a) - (b) - (c) - (d)$$

W ten sposób eliminujemy wszystkie efekty ustalone indywidualne i okresowe

$$(y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}_{it}) = (x_{it} - \bar{x}_i - \bar{x}_t + \bar{x}_{it}) + (u_{it} - \bar{u}_i - \bar{u}_t + \bar{u}_{it})$$

$$i = 1, \dots, N \quad t = 1, \dots, T$$

Jeżeli zdefiniujemy

$$\ddot{y}_{it} = y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}_{it}$$

$$\ddot{x}_{it} = x_{it} - \bar{x}_i - \bar{x}_t + \bar{x}_{it}$$

To możemy zapisać model postaci

$$\ddot{y}_{it} = \beta' \ddot{x}_{it} + \ddot{u}_{it}$$

Model ten może być oszacowany KMNK na przekształconych danych.

7.16.7 JAK ZNALEŹĆ OCENY EFEKTÓW INDYWIDUALNYCH I OKRESOWYCH?

Założmy, że mamy następujący model

$$y_{it} = \beta_0 + \alpha_i + \lambda_t + \beta' x_{it} + u_{it}$$

w którym wprowadzimy $(N-1) + (T-1)$ zmiennych sztucznych - pomijamy efekty dla pierwszego okresu λ_1 . Możemy oszacować wartość i -tego efektu indywidualnego jako

$$\hat{\alpha}_i = (\bar{y}_i - \bar{y}_1) - \hat{\beta}'(\bar{x}_i - \bar{x}_1)$$

Musimy pamiętać, że jakość estymacji zależy od liczby obserwacji dostępnych po czasie. Efekty okresowe można oszacować w analogiczny sposób.

7.16.8 BŁĘDY SZACUNKU W MODELACH DWUKIERUNKOWYCH.

Niech macierz kowariancji estymatorów parametrów

$$\text{Var}(\hat{\beta}) = \sigma_u^2 (\ddot{X}' \ddot{X})^{-1}$$

gdzie jako oszacowaną wartość wariancji zakłóceń stosujemy

$$\hat{\sigma}_n^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2}{NT - T - N - k}$$

gdzie

- \hat{u}_{it} - reszty wyliczone z estymacji modelu dwukierunkowego, $i = 1, \dots, N$, $t = 1, \dots, T$.

7.17 Testowanie heteroskedastyczności

Występowanie heteroskedastyczności możemy testować za pomocą trzech klasycznych testów ekonometrycznych, które są asymptotycznie równoważne (przy $T \rightarrow \infty$), jednak przy niewielkiej ilości okresów czasu mogą dawać różne rezultaty. We wszystkich trzech przypadkach testujemy:

$$\begin{aligned} H_0 : \sigma_i &= \sigma & \text{dla każdego } i \\ H_1 : \sigma_i &\neq \sigma_j & \text{dla } i \neq j \end{aligned}$$

i wszystkie trzy statystyki testujące mają rozkład $\chi^2(N-1)$

1. Dysponując resztami z modelu z ograniczeniami, czyli zakładającej homoskedastyczność metody KMNK, możemy znaleźć statystykę testu mnożnika Lagrange'a

$$LM = \frac{T}{2} \sum_{i=1}^N \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right)^2$$

gdzie:

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N \hat{\sigma}_i^2 = \frac{\hat{u}'\hat{u}}{NT}$$

2. Wstawiając estymatory uzyskane analogicznie na podstawie reszt z modelu bez ograniczeń (estymacja UMNK), otrzymujemy statystykę Walda

$$W = \frac{T}{2} \sum_{i=1}^N \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2} - 1 \right)^2$$

3. Estymując oba modele za pomocą metody największej wiarygodności (lub iteracyjnej UMNK) i znajdując $\hat{\sigma}^2$ z modelu z ograniczeniami oraz $\hat{\sigma}_i^2$ z modelu bez ograniczeń, możemy obliczyć statystykę testu ilorazu wiarygodności

$$LR = NT \ln \hat{\sigma}^2 - \sum_{i=1}^N T \ln \hat{\sigma}_i^2$$

Niestety, wszystkie statystyki są wrażliwe na założenie o normalności reszt. Jeżeli założenie to nie jest spełnione, możemy posłużyć się skorygowaną statystyką Walda

$$W' = \sum_{i=1}^N \frac{\left(\text{gra} \hat{\sigma}_i^2 - \hat{\sigma}^2\right)^2}{V_i}$$

gdzie

$$V_i = \frac{1}{T(T-1)} \sum_{i=1}^N \left(\hat{u}_{it}^2 - \hat{\sigma}_i^2\right)$$

7.18 Testowanie korelacji międzygrupowej

Testowanie korelacji międzygrupowej (przekrojowej, międzyjednostkowej, *cross-sectional*) - korelacja składników losowych pochodzących z różnych obiektów, ale z tego samego czasu. Najczęściej testowaniu podlega hipoteza zerowa o istnieniu *cross-sectional correlation*

$$H_0 : \Omega = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N \end{bmatrix}$$

$$H_1 : \Omega = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix}$$

1. Test mnożników Lagrange'a Breucha-Pagana będzie miał postać

$$LM = T \sum_{i=2}^N \sum_{j=1}^{i-1} r_{ij}^2$$

gdzie

- r_{ij} to współczynnik korelacji reszt między obiektami i -tym i j -tym uzyskanymi z oszacowania UMNK (ze względu na założenie stałości parametrów strukturalnych względem obiektów.)

Statystyka z próby ma rozkład $\chi^2(N(N-1)/2)$

2. Test ilorazu wiarygodności (statystyka testu)

$$LR = T \left(\sum_i \ln \hat{\sigma}_i^2 - \ln |\hat{\Sigma}| \right)$$

gdzie:

- $\hat{\sigma}_i^2$ - wyliczane są z modelu oszacowanego z uwzględnieniem zjawiska heteroskedastyczności grupowej
- $\hat{\Sigma}$ - elementami tej macierzy są oceny kowariancji wyznaczone na podstawie reszt modelu oszacowanego z uwzględnieniem zjawiska heteroskedastyczności i korelacji przekrojowej

Obie powyższe estymacje wykonane są przy pomocy metody największej wiarygodności.

Rozkład z próby ma rozkład $\chi^2(N(N-1)/2)$.

7.19 Szacowanie modelu z heteroskedastycznością

Zjawisko heteroskedastyczności grupowej jest szczególnym przypadkiem heteroskedastyczności składnika losowego. W takim przypadku do estymacji wykorzystać można Ważoną Metodę Najmniejszych Kwadratów (WMNK):

- próba dzielona jest w sposób naturalny na podpróby składające się z obserwacji w czasie dla poszczególnych jednostek
- na podstawie reszt KNK z kolejnych podprób, wyznacza się wariancje próbkowe $\hat{\sigma}_i^2$, które określają wagi przypisywane kolejnym jednostkom.

Wówczas estymator WMNK ma postać

$$\hat{\beta}_{WMNK} = (X^T V^{-1} X)^{-1} (X^T V^{-1} y) = \left[\sum_{i=1}^N \frac{1}{\sigma_i^2} X_i^T X_i \right]^{-1} \left[\sum_{i=1}^N \frac{1}{\sigma_i^2} X_i^T y_i \right]$$

7.20 Szacowanie modelu z korelacją przekrojową i heteroskedastycznością grupową

1. Metoda UMNK z estymacją (FGLS) - metoda Parka

Krok 1: Wyznacza się oceny kowariancji pomiędzy obiektami M_i -tym i j -tym na podstawie reszt uzyskanych ze zgodnego, choć obciążonego estymatora KMNK

Krok 2: Wyznacza się oceny parametrów strukturalnych według wzoru:

$$\hat{\beta}_{UMNK} = (X^T V^{-1} X)^{-1} (X^T V^{-1} y)$$

gdzie elementami V są kowariancje wyznaczone w pierwszym kroku estymacji.

2. Metoda *Panel-Corrected Standard Error* (PCSE)

- parametry strukturalne modelu szacowane są przy pomocy KMNK - oceny są zgodne i nieobciążone
- problem nieefektywności rozwiązany jest poprzez zastosowanie specyficznej formuły obliczania średnich błędów szacunku parametrów, tzw. "skorygowanych błędów standardowych dla danych panelowych"
- efektywność estymatora rośnie wraz ze wzrostem wymiaru czasowego danych względem jednostek

Macierz wariancji i kowariancji estymatorów

$$D^2(\hat{\beta})_{PCSE} = (X^T X)^{-1} (X^T \hat{V} X)^{-1} (X^T X)^{-1}$$

gdzie E jest macierzą reszt modelu oszacowanego KMNK, w której kolumna o numerze i zawiera reszty pochodzące z i -tego obiektu.

7.21 Testowanie "zwykłej" autokorelacji składnika losowego

W modelu panelowym może ponadto występować jeszcze klasyczna autokorelacja zakłóceń losowych. Możliwe warianty:

- 1a. Składniki losowe z różnych obiektów są nieskorelowane, autokorelacja występuje tylko "wewnątrz" poszczególnych obiektów
- 2a. Oprócz autokorelacji dopuszcza się też występowanie korelacji przekrojowej
- 1b. Współczynnik autokorelacji może być jednakowy dla wszystkich obiektów
- 2b. Współczynnik autokorelacji może być różnych dla różnych obiektów

1. Test seryjnej autokorelacji dowolnego rzędu Breuscha-Godfrey'a (oparty na mnożnikach Lagrange'a); stosowany w modelach regresji łącznej FE i RE.
2. Test Baltagi-Li (oparty na mnożnikach Lagrange'a) - bada składniki losowe dla poszczególnych jednostek, czy nie mają charakteru AR(1) lub MA(1), stosowany w modelach RE
3. Test Durбина-Watsona, stosowany w modelach regresji łącznej FE i RE.

7.22 Szacowanie modelu z autokorelacją

W celu usunięcia autokorelacji stosuje się zazwyczaj przekształcenie Prais-Winstera - szeregi czasowe dotyczące kolejnych obiektów transformuje się w następujący sposób:

$$Y_i^* = \begin{bmatrix} \sqrt{1-r_i^2}y_{i1} \\ y_{i2} - r_i y_{i1} \\ \vdots \\ y_{it} - r_i y_{i(T-1)} \end{bmatrix} \quad X_i^* = \begin{bmatrix} \sqrt{1-r_i^2}x_{i1} \\ x_{i2} - r_i x_{i1} \\ \vdots \\ x_{it} - r_i x_{i(T-1)} \end{bmatrix}$$

gdzie

- r_i jest oceną współczynnika autokorelacji
- x_{it} jest wektorem wartości wszystkich zmiennych objaśniających przyjmowanych dla i -tego obiektu w okresie t .

W modelu szacowanym na przetransformowanych danych nie występuje autokorelacja.

7.23 UMNK dla modelu z efektami losowymi

W zapisie macierzowym

$$Y = X\beta + v$$

Macierz wariancji i kowariancji dwuczęściowego składnika losowego można ogólnie zapisać jako

$$\mathbb{E}(vv') = \mathbb{E}\left(\begin{bmatrix} v_1v_1 & \dots & v_1v_N \\ \vdots & \ddots & \vdots \\ v_Nv_1 & \dots & v_Nv_N \end{bmatrix}\right) = \Omega$$

zauważmy, że z założeń o niezależności poszczególnych części składnika losowego wynika, że

- dla $i \neq j$

$$\begin{aligned}\mathbb{E}(v_i v_j') &= \mathbb{E}(\alpha_i + u_i)(\alpha_j + u_j)' = \\ &= \mathbb{E}(\alpha_i \alpha_j') + \mathbb{E}(\alpha_i u_j') + \mathbb{E}(\alpha_i u_i) + \mathbb{E}(u_i u_j') = 0\end{aligned}$$

- dla $i = j$

$$\begin{aligned}\mathbb{E}(v_i v_j') &= \mathbb{E}(\alpha_i + u_i)(\alpha_i + u_i)' = \\ &= \mathbb{E}(\alpha_i^2) + \mathbb{E}(\alpha_i u_i) + \mathbb{E}(u_i u_i') = \sigma_\alpha^2 + \sigma_u^2 I = \omega\end{aligned}$$

Macierz wariancji i kowariancji składników losowych w modelu RE jest macierzą blokowo-diagonalną:

$$\Omega = \begin{bmatrix} \omega & 0 & \dots & 0 \\ 0 & \omega & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega \end{bmatrix} \neq \sigma^2 I$$

W UMNK potrzebujemy odwrotności tej macierzy, a zatem i odwrotności ω . Można wykazać, że przy przyjętych założeniach:

$$\omega^{-1} = \frac{1}{\sigma_u^2} \left[I - \frac{\sigma_\alpha^2}{\sigma_u^2 + T\sigma_\alpha^2} \right]$$

Nieznane wariancje można zatąpić ich ocenami uzyskanymi z oszacowania:

- σ_u^2 - z modelu FE
- σ_α^2 - z regresji międzygrupowej

Ostateczną postać estymatora UMNK dla modelu z efektami losowymi (RE) można zapisać

$$\begin{aligned}\hat{\beta}_{RE} &= \left(\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' + \psi T \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})' \right)^{-1} \cdot \\ &\cdot \left(\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' + \psi T \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})' \right)^{-1}\end{aligned}$$

gdzie

$$\psi = \frac{\sigma_u^2}{\sigma_i^2 + T\sigma_\alpha^2}$$

Przybliżona wartość losowych efektów indywidualnych można wyznaczyć jako

$$\hat{\mu}_{RE} = \bar{Y} = \hat{\beta}_{RE}\bar{x}$$

Przypadki modeli panelowych niespełniających dotychczasowych założeń

1. Model ze zmiennymi stratami w czasie

Założmy, że część zmiennych objaśniających w modelu panelowym jest stała w czasie. Zauważmy, że zmienne stałe w czasie są współliniowe w ustalonych efektami indywidualnymi. Zatem nie możemy uzyskać ocen parametrów strukturalnych przy tych zamiennych w modelu FE.

Rozwiązania:

1. Stosujemy model RE - nie jest możliwe, jeśli zmienne objaśniające okazują się skorelowane ze składnikiem losowym - (test Hausmana).
2. Szacujemy model FE przy pomocy estymatora Hausmana-Tatylora
3. Stosujemy inny zastaw założeń dla efektów indywidualnych w modelu RE.