Egzamin z ekonometrii dla IiE14.09.2007
II semestr

ZADANIE 1 Analizowany jest model liniowy z heteroskedastycznością:

$$y_i = \beta x_i + \varepsilon_i$$

$$\varepsilon_i | \, x_i \sim N\left(0, \exp\left(\alpha x_i\right)\right)$$

oraz wiadomo, że ε_i i ε_j są niezależne dla $i \neq j$.

1. Pokazać, że znalezienie estymatora MNW wymaga w tym przypadku rozwiązania nieliniowego układu równań.
2. Znaleźć estymator UMM dla parametrów tego modelu

Podpowiedź: wykorzystać to, że z model można zapisać jako

$$y_i = \beta x_i + \exp\left(\alpha x_i\right) u_i$$

gdzie $u_i | \, x_i \sim N(0,1)$ a $\varepsilon_i = \exp\left(\alpha x_i\right) u_i$. W rezultacie $\mathrm{E}\left(\ln \varepsilon_i^2 \middle| \, x_i\right) = 2ax_i + \mathrm{E}\left(\ln u_i^2 \middle| \, x_i\right)$ a dla $u_i \sim N(0,1)$, $\mathrm{E}\left(\ln u_i^2\right) = \gamma \approx -1.27$.

Rozwiązanie:

1. Funkcja gęstości rozkładu normalnego, dla każdej z obserwacji ma postać

   $$f\left(y_i | \, x_i\right) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\left(y_i - x_i\beta\right)^2}{2\sigma_i^2}\right)$$

   Logarytm funkcji wiarygodności (pomijamy elementy stałe) będzie więc miał postać:

   $$\begin{aligned}\ell\left(\alpha, \beta, \sigma_\alpha\right) &= -\sum_{i=1}^n \frac{1}{2}\ln\left(\sigma_i^2\right) - \sum_{i=1}^n \frac{\left(y - x_i\beta\right)^2}{2\sigma_i^2} \\ &= -\sum_{i=1}^n \frac{1}{2}\ln\left[e^{\alpha x_i}\right] - \frac{1}{2}\frac{\sum_{i=1}^n \left(y - x_i\beta\right)^2}{e^{\alpha x_i}}\end{aligned}$$

   Pierwsze pochodne funkcji wiarygodności mają postać:

   $$\begin{aligned}\frac{\partial \ell\left(\alpha, \beta\right)}{\partial \alpha} &= -\frac{1}{2}\frac{x_i e^{\alpha x_i}}{e^{\alpha x_i}} + \frac{1}{2}\sum_{i=1}^n \frac{e^{\alpha x_i} x_i \left(y_i - x_i\beta\right)^2}{\left(e^{\alpha x_i}\right)^2} \\ \frac{\partial \ell\left(\alpha, \beta\right)}{\partial \beta} &= \sum_{i=1}^n \frac{x_i\left(y_i - x_i\beta\right)}{e^{\alpha x_i}}\end{aligned}$$

   Znalezienie estymatorów MNW wymaga więc znalezienia rozwiązania nieliniowego układu równań:

   $$\sum_{i=1}^n x_i = \sum_{i=1}^n \frac{x_i\left(y_i - x_i\widehat{\beta}\right)^2}{e^{\widehat{\alpha} x_i}}$$

   $$\sum_{i=1}^n \frac{x_i\left(y_i - x_i\widehat{\beta}\right)}{e^{\widehat{\alpha} x_i}} = 0$$

2. Estymator UMM możemy wyprowadzić zauważając, że $\mathrm{E}\left(\varepsilon_i\right) = 0$ i $\mathrm{E}\left(\ln \varepsilon_i^2 - ax_i - \gamma\right) = 0$ gdzie $\varepsilon_i = y_i - x_i\beta$. Zastępując momenty teoretyczne empirycznymi uzyskujemy:

   $$\overline{y} - \overline{x}\widehat{\beta} = 0$$

   $$\frac{1}{n}\sum_{i=1}^n \ln\left(y_i - x_i\widehat{\beta}\right)^2 = \alpha\frac{1}{n}\sum_{i=1}^n x_i + \gamma$$

Wzór na $\widehat{\beta} = \frac{\overline{y}}{\overline{x}}$ uzyskujemy z pierwszego równania. Oznaczmy

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\widehat{\beta}\right)^2 = \overline{e^2}$$
$$\frac{1}{n}\sum_{i=1}^{n}\ln\left(y_i - x_i\widehat{\beta}\right)^2 = \overline{\ln e_i^2}$$

Mając estymator $\widehat{\beta}$ możemy policzyć $\overline{\ln\left(e^2\right)}$ i uzyskać $\widehat{\alpha}$ z drugiego równania

$$\widehat{\alpha} = \frac{\overline{\ln\left(e^2\right)} - \gamma}{2\overline{x}}$$

ZADANIE 2 Dany jest następujący model popytu i podaży pracy

$$l_{1t} = \alpha_0 + \alpha_1 w_t + \alpha_2 h_t + \varepsilon_{1t}$$
$$l_{2t} = \beta_1 w_t + \beta_2 k_t + \varepsilon_{2t}$$
$$l_{1t} = l_{2t}$$

gdzie $l_t = l_{1t} = l_{2t}$ oznacza logarytm zatrudnienia, w_t to logarytm płacy realnej, h_t logarytm liczby osób w wieku produkcyjnym, a k_t logarytm wielkości kapitału w gospodarce. Zmiennymi endogenicznymi są l_{1t}, l_{2t} i w_t. Zmienne h_t i k_t są traktowane jako egzogeniczne.

1. Które z tych równań jest równaniem popytu na pracę, a które podaży pracy? Odpowiedź uzasadnić.
2. Sprawdzić identyfikację poszczególnych równań.
3. Wyjaśnić jaką postać będzie miała forma zredukowana tego modelu.
4. Jaka jest różnica między interpretacją parametru przy h_t w równaniu dla popytu na pracę w formie strukturalnej i interpretacją parametru przy h_t w równaniu dla wielkości zatrudnienia w formie zredukowanej tego modelu?
5. Wyjaśnić, dlaczego równania popytu nie można poprawnie wyestymować za pomocą MNK.
6. W jaki sposób można użyć formy zredukowanej do policzenia estymatorów Pośredniej MNK? Wyprowadzić estymator β_1 tej metody i wyjaśnić, czy uzyskane w ten sposób oszacowanie jest jednoznaczne.
7. Jeśli zastosowalibyśmy do estymacji równania popytu MZI, to jakie zmienne mogłyby być użyte jako instrumenty?

Rozwiązanie:

1. Równaniem podaży pracy jest pierwsze równanie: podaż pracy zależy od płacy i liczby osób w wieku produkcyjnym ale nie zależy od kapitału. Równaniem popytu na pracę jest równanie drugie: popyt na pracę zależy od płacy i wielkości kapitału w gospodarce. Zależność między popytem na pracę i kapitałem wynika z tego, że im mniejszy kapitał, tym mniejsza produkcyjność krańcowa i niższa płaca przy tej samej wielkości zatrudnienia.
2.

zmienne egzogeniczne	$1, p_t, q_t$
zmienne endogeniczne	l_t, w_t

$K = 3$	$G_1 = 2$	$K_1 = 2$	zidentyfikowane $3 = K \geq G_1 + K_1 - 1 = 3$
	$G_2 = 2$	$K_2 = 1$	zidentyfikowane $3 = K \geq G_1 + K_1 - 1 = 2$

3. W formie zredukowanej po lewej stronie znajdują się zmienne endogeniczne a po prawej wyłącznie zmienne egzogeniczne. W przypadku analizowanego modelu, forma zredukowana będzie więc miała postać:

$$l_t = \pi_{10} + \pi_{11}h_t + \pi_{12}k_t + \epsilon_{1t}$$
$$w_t = \pi_{20} + \pi_{21}h_t + \pi_{22}k_t + \epsilon_{2t}$$

4. Parametr α_2 w równaniu podaży pracy interpretujemy jako elastyczność podaży pracy względem zmiany liczby osób w wieku produkcyjnym. Opisuje on ile procent zmieni się podaży pracy, gdy liczba osób w wieku produkcyjnym wzrasta o 1% *a wysokość płacy pozostaje niezmieniona*. π_{11} jest mnożnikiem wielkości zatrudnienia względem liczby osób w wieku produkcyjnym. Opisuje on o ile procent wzrosłoby zatrudnienie, gdyby liczba osób w wieku produkcyjnym wzrosła o 1% *a płaca dostosowała się do nowego poziomu równowagi*.

5. Równania popytu nie da się wyestymować za pomocą MNK, ponieważ jedna ze zmiennych objaśniających jest zmienną endogeniczną (płaca) co oznacza, że jest skorelowana z błędem losowym. W tym przypadku MNK daje estymatory, które nie są zgodne.

6. Aby znaleźć postać estymatorów pośredniej MNK należy znaleźć zależności między parametrami formy strukturalnej i zredukowanej. Rozwiązując formę strukturalną dla l_t i w_t uzyskujemy

$$l_t = \underbrace{-\frac{\beta_1\alpha_0}{\alpha_1-\beta_1}}_{\pi_{20}} + \underbrace{\frac{\alpha_1}{\alpha_1-\beta_1}}_{\pi_{21}}k_t \underbrace{-\frac{\beta_1\alpha_2}{\alpha_1-\beta_1}}_{\pi_{22}}h_t + \underbrace{\frac{\alpha_1\varepsilon_{2t}-\beta_1\varepsilon_{1t}}{\alpha_1-\beta_1}}_{\epsilon_{2t}}$$
$$w_t = \underbrace{-\frac{\alpha_0}{\alpha_1-\beta_1}}_{\pi_{10}} + \underbrace{\frac{\beta_2}{\alpha_1-\beta_1}}_{\pi_{11}}k_t \underbrace{-\frac{\alpha_2}{\alpha_1-\beta_1}}_{\pi_{12}}h_t + \underbrace{\frac{\varepsilon_{2t}-\varepsilon_{1t}}{\alpha_1-\beta_1}}_{\epsilon_{1t}}$$

Mamy więc następujące zależności między parametrami formy strukturalnej i zredukowanej

$$\begin{array}{lll} \pi_{10} = -\frac{\alpha_0}{\alpha_1-\beta_1} & \pi_{11} = \frac{\beta_2}{\alpha_1-\beta_1} & \pi_{12} = \frac{\alpha_2}{\alpha_1-\beta_1} \\ \pi_{20} = -\frac{\beta_1\alpha_0}{\alpha_1-\beta_1} & \pi_{21} = \frac{\alpha_1}{\alpha_1-\beta_1} & \pi_{22} = \frac{\beta_1\alpha_2}{\alpha_1-\beta_1} \end{array}$$

Estymatory parametrów uzyskujemy estymując formę zredukowaną, zastępując π_{ij} estymatorami π_{ij} i rozwiązując ten układ równań dla $\alpha_0, \alpha_1, \alpha_2, \beta_1$ i β_2. Z równań tych można uzyskać dwa estymatory parametru β_1 postaci: $\widehat{\beta}_1 = \frac{\widehat{\pi}_{20}}{\widehat{\pi}_{10}}$, $\widehat{\widehat{\beta}}_1 = \frac{\widehat{\pi}_{20}}{\widehat{\pi}_{10}}$. Uzyskaliśmy dwa oszacowania jednego parametru!

7. Zmiennymi instrumentalnymi mogłyby być wszystkie zmienne egzogeniczne a więc stała, h_t i k_t.

ZADANIE 3 Dany jest model logitowy, ze zmiennymi niezależnymi x_1, x_2, x_3. Hipotezy testujemy na poziomie istotności $\alpha = 0.05$.

1. Przetestować łączną nieistotność x_1 i x_2. Dla modelu z pełną liczbą zmiennych wielkość logarytmu funkcji wiarogodności wyniosła -277.5. Dla modelu bez zmiennych x_1 i x_2 wartość logarytmu funkcji wiarogodności wyniosła -281. Policzyć odpowiednią statystykę testową i podać wynik testu.

   Podpowiedź: $\chi^2_{0.05}(1) = 3.84$, $\chi^2_{0.05}(2) = 5.99$, $\chi^2_{0.05}(3) = 7.81$.

2. Policzono wielkość statystyki Walda dla hipotezy $H_0 : \beta_1^2 = \beta_2$ oraz $H_0 : \frac{\beta_1^2}{\beta_2} - 1 = 0$ i otrzymano wielkość statystyk testowych 2 i 5. Przeprowadzić testy dla tych dwóch hipotez i skomentować uzyskane wyniki.

3. Rozważmy następujący problem: chcemy przetestować hipotezę $\gamma = 1$ w modelu

$$y_i = (\beta_1 x_{1i} + \beta_2 x_{2i})^\gamma + \varepsilon_i$$

   przy czym $\varepsilon \sim N(0, \boldsymbol{I})$. Dysponujemy programem, który umożliwia jedynie szacowanie zwykłej regresji liniowej. Jakim rodzajem testu powinniśmy się posłużyć w tym przypadku?

Rozwiązanie:

1. Statystyka testowa $LR = -2(L_0 - L_1)$ i ma rozkład $\chi^2(k)$, gdzie k to liczba zmiennych o których zakładamy że są nieistotne, L_0 wartość logarytmu wiarogodności dla modelu bez ograniczeń, L_1 wartość logarytmu wiarogodności dla modelu z ograniczeniami. W tym przypadku:

$$LR = -2(-281 - (-277.5)) = 7 > \chi^2(2) = 5.99$$

   Wobec tego należy uznać, że te modele różnią się. Wobec tego zmienne są łącznie istotne.

2. W pierwszym przypadku nie podstaw do odrzucenia $H_0 : \beta_1^2 = \beta_2$ $[2 < 3.84]$, w drugim przypadku hipotezę $H_0 : \frac{\beta_1^2}{\beta_2} - 1 = 0$ odrzucamy $[5 > 3.84]$. Analizowane hipotezy są równoważne algebraicznie a więc wynik dwóch testów daje sprzeczne wyniki. W przypadku testu Walda jest to możliwe, ponieważ w małych próbach i dla hipotez nieliniowych wynik tego testu może zależeć od sposobu sformułowania H_0.

3. Statystyką mnożników Lagrange'a - umożliwia on zweryfikowanie H_0 przy znajomości oszacowań modelu z ograniczeniami. W tym przypadku model z ograniczeniami jest bardzo zwykłym modelem liniowym $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$ spełniającym założenia $KMRL$.