Egzamin poprawkowy z Ekonometrii 31.08.2010

Pytania teoretyczne

1. Wyjaśnić, jak należy rozumieć przyczynowość w sensie Grangera i jak jest ona testowana.
2. Opisać procedurę testowania stacjonarności za pomocą rozszerzonego testu Dickey-Fullera (ADF).
3. Podać założenia, które powinny spełniać instrumenty używane w trakcie estymacji MZI. Za pomocą jakich testów weryfikuje się prawidłowość tych założeń?
4. Czym różni się forma strukturalna od formy zredukowanej? Na czym polega różnica w interpretacji parametrów i równań formy strukturalnej i zredukowanej?

ZADANIE 1 Mamy następujący liniowy model efektów nieobserwowalnych

$$y_{it} = \alpha + \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i + \varepsilon_{it}$$

Załóżmy, że

$$\begin{aligned} \mathrm{E}\left(\boldsymbol{\varepsilon}|\,\boldsymbol{X}\right) &= \mathbf{0} \\ \mathrm{E}\left(\boldsymbol{u}\right) &= \mathbf{0} \end{aligned}$$

oraz, że wśród zmiennych objaśniających nie ma zmiennych stałych w czasie. W analizowanym modelu może wystąpić korelacja między efektami stałymi a zmiennymi objaśniającymi $\mathrm{E}\left(\boldsymbol{u}|\,\boldsymbol{X}\right) \neq \mathbf{0}$. Estymator wektora stałej uzyskano w sposób następujący:

1. policzono estymator efektów stałych wektora parametrów $\boldsymbol{\beta}$
2. estymator α uzyskano jako $\widetilde{\alpha} = \overline{y}_{..} - \overline{\boldsymbol{x}}_{..}\widetilde{\boldsymbol{\beta}}_{FE}$, gdzie $\overline{y}_{..} = \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$, $\overline{\boldsymbol{x}}_{..} = \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T} \boldsymbol{x}_{it}$.

Pokaż, że estymator $\widetilde{\alpha}$:

1. jest estymatorem warunkowo obciążonym $\mathrm{E}\left(\widetilde{\alpha} - \alpha|\,\boldsymbol{X}\right) \neq 0$
2. jest estymatorem bezwarunkowo nieobciążonym $\mathrm{E}\left(\widetilde{\alpha} - \alpha\right) = 0$

Podpowiedź: w zapisie macierzowym estymator efektów stałych można zapisać jako $\widetilde{\boldsymbol{\beta}}_{FE} = \left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{y}}$ przy czym elementami $\ddot{\boldsymbol{X}}$ są $\ddot{\boldsymbol{x}}_{it} = \boldsymbol{x}_{it} - \overline{\boldsymbol{x}}_{i\cdot}$ a $\ddot{y}_{it} = y_{it} - \overline{y}_{i\cdot}$, gdzie $\overline{y}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\overline{\boldsymbol{x}}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_{it}$. Ponadto z założenia, że $\mathrm{E}\left(\boldsymbol{\varepsilon}|\,\boldsymbol{X}\right) = \mathbf{0}$ wynika, że $\mathrm{E}\left(\ddot{\boldsymbol{\varepsilon}}|\,\boldsymbol{X}\right) = \mathbf{0}$, gdzie $\ddot{\varepsilon}_{it} = \varepsilon_{it} - \overline{\varepsilon}_{i\cdot}$ i $\overline{\varepsilon}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}$.

Rozwiązanie:

Zauważmy, że przekształcony za pomocą przekształcenia efektów stałych model ma postać

$$\ddot{\boldsymbol{y}} = \ddot{\boldsymbol{X}}\boldsymbol{\beta} + \ddot{\boldsymbol{\varepsilon}}$$

Wynika z tego, że

$$\widetilde{\boldsymbol{\beta}}_{FE} = \left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\left(\ddot{\boldsymbol{X}}\boldsymbol{\beta} + \ddot{\boldsymbol{\varepsilon}}\right) = \boldsymbol{\beta} + \left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{\varepsilon}}$$

Zauważmy, że

$$\overline{y}_{..\cdot} = \alpha + \overline{\boldsymbol{x}}_{..}\boldsymbol{\beta} + \overline{u}_{\cdot} + \overline{\varepsilon}_{..}$$

Estymator stałej jest więc równy

$$\begin{aligned} \widetilde{\alpha} = \overline{y}_{..} - \overline{\boldsymbol{x}}_{..}\widetilde{\boldsymbol{\beta}}_{FE} &= \alpha + \overline{\boldsymbol{x}}_{..}\boldsymbol{\beta} + \overline{u}_{\cdot} + \overline{\varepsilon}_{..} - \overline{\boldsymbol{x}}_{..}\,\boldsymbol{\beta} - \overline{\boldsymbol{x}}_{..}\left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{\varepsilon}} \\ &= \alpha + \overline{u}_{\cdot} - \overline{\boldsymbol{x}}_{..}\left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{\varepsilon}} \end{aligned}$$

Wynika z tego, że warunkowa wartość oczekiwana obciążenia jest równa:

$$\mathrm{E}\left(\widetilde{\alpha} - \alpha|\,\boldsymbol{X}\right) = \mathrm{E}\left(\overline{u}_{\cdot}|\,\boldsymbol{X}\right) - \overline{\boldsymbol{x}}_{..}\left(\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}\right)^{-1}\ddot{\boldsymbol{X}}'\underbrace{\mathrm{E}\left(\ddot{\boldsymbol{\varepsilon}}|\,\boldsymbol{X}\right)}_{\mathbf{0}} = \mathrm{E}\left(\overline{u}_{\cdot}|\,\boldsymbol{X}\right) \neq 0$$

a bezwarunkowe obciążenie jest równe zeru

$$\mathrm{E}(\widetilde{\alpha}-\alpha)=\mathrm{E}(\overline{u}.)=\frac{1}{T}\mathbf{1}'\,\mathrm{E}(\boldsymbol{u})=0$$

ZADANIE 2 Otrzymano następujące wyniki estymacji modelu $ARIMA(p,d,q)$:

$$\Delta y_t = 0.2\Delta y_{t-1} + 0.3\Delta y_{t-2} + \epsilon_t + 0.2\epsilon_{t-1}.$$

1. Podaj wielkość p, d, q
2. Policz prognozę dla y_{T+1} i y_{T+2} jeśli $y_T = 2$, $y_{T-1} = 1$, $y_{t-2} = 2$, $y_{t-3} = 1$, $e_T = -0.25$.
3. Znajdź rozwiązanie długookresowe dla y_t w tym modelu.

Rozwiązanie:

1. $p = 2$, $d = 1$, $q = 1$
2. $$\begin{aligned}\Delta\widehat{y}_{T+1} &= 0.2y_T + 0.3\Delta y_{T-1} + 0.2e_t \\ &= 0.2\times 1.0 + 0.3\times(-1) + 0.2\times(-0.25) = -.15, \\ \Delta\widehat{y}_{T+2} &= 0.2\Delta\widehat{y}_{T+1} + 0.3\Delta y_t \\ &= -0.03 + 0.3\times 1 = 0.27\end{aligned}$$

   w związku z tym

   $$\begin{aligned}\widehat{y}_{T+1} &= y_T + \Delta\widehat{y}_{T+1} = 2 - .15 = 1.85 \\ \widehat{y}_{T+2} &= \widehat{y}_{T+1} + \Delta\widehat{y}_{T+2} = 1.85 + 0.27 = 2.12\end{aligned}$$

3. Rozwiązanie długookresowe liczymy przyjmując, że $\mathrm{E}(y_t) = \mathrm{E}(y_{t-1}) = \ldots = y_t^*$ w naszym przypadku

   $$\underbrace{\mathrm{E}(y_t - y_{t-1})}_{0} = 0.2\underbrace{\mathrm{E}(y_{t-1} - y_{t-2})}_{0} + 0.4\underbrace{\mathrm{E}(y_{t-2} - y_{t-3})}_{0} + \underbrace{\mathrm{E}(\epsilon_t)}_{0} + 0.2\underbrace{\mathrm{E}(\epsilon_{t-1})}_{0}$$

   otrzymujemy więc

   $$0 = 0$$

   rozwiązania długookresowe (równowaga długookresowa) w tym modelu nie istnieje!

ZADANIE 3 Na podstawie danych z badania dochodów i wydatków konsumpcyjnych starano się wyjaśnić wielkość miesięcznych wydatków poniesionych w restauracjach i kawiarniach przez osoby samotne za pomocą wysokość dochodu, wieku i płci. Poniżej znajdują się oszacowania parametrów i efektów cząstkowych dla modelu tobitowego oszacowanego dla tego problemu.

```
Tobit regression                                  Number of obs   =       5089
                                                  LR chi2(3)      =     800.00
                                                  Prob > chi2     =     0.0000
Log likelihood = -4375.8865                       Pseudo R2       =     0.0838

------------------------------------------------------------------------------
         wyd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dochg |   .0555346   .0034797    15.96   0.000      .048713    .0623562
        wiek |  -3.371001   .2278059   -14.80   0.000    -3.817598   -2.924403
   1.kobieta |  -54.79897   7.322538    -7.48   0.000    -69.15429   -40.44364
       _cons |    1.77951   13.57973     0.13   0.896    -24.84262    28.40164
```

```
-------------+----------------------------------------------------------------
      /sigma |   136.1669   4.530283                      127.2856    145.0482
------------------------------------------------------------------------------
  Obs. summary:       4530  left-censored observations at wyd<=0
                       559     uncensored observations
                         0 right-censored observations


Efekt czastkowy dla Pr(wyd>0)
------------------------------------------------------------------------------
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dochg |   .0000597   3.78e-06    15.81   0.000     .0000523    .0000671
        wiek |  -.0036249   .0002232   -16.24   0.000    -.0040623   -.0031875
     kobieta |  -.0661504   .0097708    -6.77   0.000    -.0853007   -.0470001
------------------------------------------------------------------------------


Efekt czastkowy dla E(wyd*)
------------------------------------------------------------------------------
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dochg |    .005689   .0003851    14.77   0.000     .0049341    .0064438
        wiek |  -.3453266   .0261011   -13.23   0.000    -.3964837   -.2941695
     kobieta |  -6.279588   .9481177    -6.62   0.000    -8.137864   -4.421311
------------------------------------------------------------------------------
```

Testy przeprowadzamy na poziomie istotności $\alpha = 0.05$. Wyniki testów należy uzasadnić liczbami z wydruku.

1. Wypisz założenia modelu tobitowego.
2. Jaką nietypową cechą będzie się najprawdopodobniej charakteryzować rozkład zmiennej zależnej?
3. Dlaczego zarówno policzenie regresji liniowej dla tej całej obserwowanej próby da najprawdopodobniej wartości dopasowane, które dla części obserwacji będą nieinterpretowalne?
4. Zbadaj istotność poszczególnych zmiennych w modelu oraz ich łączną istotność.
5. Zintepretuj efekty cząstkowe dla dochodu i wieku policzone dla zmiennej obserwowalnej.
6. Zintepretuj efekty cząstkowe policzone dla $\mathrm{E}\left(y|\, y>0\right)$ oraz efekty cząstkowe policzone dla prawdopodobieństwa niezerowych wydatków.
7. Poniżej znajduje się oszacowanie elastyczności dochodowej wydatków w restauracjach i kawiarniach policzone dla średniej wartości dochodu w próbie (a więc $\frac{\partial\,\mathrm{E}(y|x)}{\partial x}\Big/\frac{y}{x}$ policzone dla $x=\overline{x}$). Zintepretuj tę wartość i oceń, czy jest ona zgodna z intuicją ekonomiczną.

   ```
   Elastycznosc dla E(wyd*)
   ------------------------------------------------------------------------------
                |      ey/ex   Std. Err.      z    P>|z|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
          dochg |   .8098296   .0492449    16.44   0.000     .7133113    .9063479
   ------------------------------------------------------------------------------
   ```

8. Do zmiennych w modelu dodano zmienną wiek^2 i uzyskano wielkość logarytmu funkcji wiarygodności na poziomie -4363.6564. Zweryfikuj hipotezę, że ta dodatkowa zmienna jest nieistotna w modelu.

   Podpowiedź: $\chi^2_{0.95}\left(1\right)=3.84$, $\chi^2_{0.95}\left(2\right)=5.99$, $\chi^2_{0.95}\left(3\right)=7.81$.

Rozwiązanie:

1. Założenia modelu tobitowego są następujące:

$$y_i^* = \boldsymbol{x}_i\boldsymbol{\beta} + \varepsilon_i$$
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$
$$y_i = \begin{cases} y_i = y_i^* & \text{dla} \quad y_i > 0 \\ y_i = 0 & \text{dla} \quad y_i \leq 0 \end{cases}$$

   i poszczególne ε_i są niezależne.

2. Nie wszystkich stać na jedzenie poza domem, więc dla niektórych gospodarstw wydatki poniesione w restauracjach i kawiarniach wyniosą 0. Na podstawie wydruku widzimy, że na 5089 przebadanych gospodarstw, jedynie 559 poniosło w danym miesiącu wydatki tego typu.

3. W przypadku oszacowania dla tego modelu zwykłej regresji liniowej najprawdopodobniej część wartości dopasowanych będzie ujemna - co jednak jest bez sensu ponieważ wydatki mogą być wyłącznie zerowe bądź dodatnie.

4. Wszystkie zmienne w modelu są istotne (wartość p są dla wszystkich zmiennych mniejsze od 0.05), stała jest nieistotna ($0.896 > 0.05$) . Łącznie wszystkie zmienne są istotne o czym świadczy wartość p dla statystki LR ($0.000 < 0.05$).

5. Oczekiwanyrost wydatków odczytujemy z tablicy efektów cząskowych dla $y*$. Wzrost dochodu o 1000 zł. powoduje oczekiwany wzrost wydatków na restauracje i kawiarnie o 5 zł na miesiąc, wzrost wieku o 1 rok powoduje spadek tego typu wydatków o 34 gr.

6. Na podstawie wydruku efektów cząstkowych dla prawdopodobieństwa dochodzimy do wniosku, że prawdopodobieństwo poniesienia większych od zera wydatków w restauracjach i kawiarniach wzrośnie o 5.97 punkta procentowego jeśli dochód wzrośnie o 1000 zł, wrost wieku o rk powoduje spadek tego p-stwa o 0.36 punktu procentowego oraz, że kobiety mają 6.61 punktu procentowego niższe p-stwo poniesienia tego typu wydatków.

7. Oszcowana wielkość oznacza, że przy 1% wzroście dochodu oczekiwane wydatki na retauracje i kawiarnie wzrosną o 0.80%. Oszacowanie to wydaje się niskie ponieważ wydatki tego typu wydają się raczej wydatkami na dobra luksusowe (powinny rosnąc wraz z dochodem) - a więc oszacowanie elastyczności powinno wyjść większe od 1.

8. Statystyka testu LR ma postać: .

$$LR = 2\left(-4375.8865 + 4363.6564\right) = 24.46 > \chi^2_{0.95}(1) = 3.84$$

   Testujemy zerowość 1 współczynnika (a więc właściwą wartością krytyczną jest $\chi^2_{0.95}(1)$. Odrzucamy H_0 o nieistotności dochodu^2.