Egzamin z ekonometrii 13.06.2006
Informatyka i Ekonometria
II semestr

Pytania teoretyczne

1. Jakie trzy testy stosujemy w kontekście estymacji MNW? Porównaj zalety i wady tych testów.

   Stosujemy trzy test: LR, LM i W. Test LR jest łatwy do przeprowadzenia ale wymaga wyestymowania modelu z ograniczeniami i bez ograniczeń. Test LM wymaga wyestymowania jedynie modelu z ograniczeniami a test W jedynie modelu bez ograniczeń. Test W w przypadku testowania hipotez nieliniowych może dać różne wyniki w małych próbach w zależności od sposobu zdefiniowania hipotezy.

2. Kiedy do modelowania dyskretnej zmiennej zależnej powinno się użyć modelu wielomianowego logita, a kiedy uporządkowanego logita lub probita?

   Model wielomianowy logitowy stosujemy w przypadku, gdy modelujemy zmienną dyskretną o z góry znanej liczbie poziomów, których jednak nie da się logicznie uporządkować. Modele uporządkowane stosujemy w przypadku, gdy zmienna ma znaną liczbę poziomów i poziomy te można w jakiś sposób uporządkować.

3. Pokazać, że estymator MNK można wyprowadzić jako estymator GMM z warunku, że $\mathrm{E}\left(\varepsilon_i | \boldsymbol{x}_i\right) = 0$

   Warunek narzucony na moment warunkowy: $\mathrm{E}\left(\varepsilon_i | \boldsymbol{x}_i\right) = 0$. Z warunku tego wynika następujące ograniczenie narzucone na moment bezwarunkowy: $\mathrm{E}\left(\boldsymbol{x}_i' \varepsilon_i\right) = 0$. Stosując ogólne zasady liczenia estymatorów GMM zamieniamy moment bezwarunkowy na średnią z próby:

   $$\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i' \varepsilon_i = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i' \left(y_i - \boldsymbol{x}_i \boldsymbol{\beta}\right) = \frac{1}{n}\left(\boldsymbol{X}'\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}\right) = \boldsymbol{0}$$

   Estymatorem GMM jest $\widehat{\boldsymbol{\beta}}$ dla którego równość ta jest spełniona a więc estymator MNK postaci: $\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}$

4. Czym różni się forma strukturalna modelu wielorównaniowego od jego formy zredukowanej?

   Forma strukturalna ma postać układu równań:

   $$\boldsymbol{A}\boldsymbol{Y}_t = \boldsymbol{B}\boldsymbol{X}_t + \boldsymbol{u}_t,$$

   Poszczególne równania formy strukturalnej mają określoną interpretację ekonomiczną. W formie strukturalnej zmienna która jest zmienną objaśnianą w jednym równaniu, może być zmienną objaśniającą w innym równaniu. Forma zredukowana powstaje z formy strukturalnej po przemnożeniu obu jej stron przez $\boldsymbol{A}^{-1}$ i ma ona postać układu równań:

   $$\boldsymbol{Y}_t = \boldsymbol{\Pi}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t$$

   W formie zredukowanej równania nie mają interpretacji ekonomicznej a parametry interpretujemy jako mnożniki. W tej formie modelu zmienne objaśniane (endogeniczne) znajdują się wyłącznie po lewej stronie układu równań.

ZADANIE 1 Wyestymowano model wielomianowego logita na próbie BAEL. Model ten wyjaśnia wybory dotyczące typu pracy. Analizowane kategorie: na własny rachunek, pracownik najemny, pomagający członek rodziny. Zmiennymi objaśniającymi było wiek (AGE), wiek^2 ($AGE2$) oraz płeć (1 mężczyzna, 2 kobieta). Z regresji tej uzyskano następujące wyniki:

```
Multinomial logistic regression                   Number of obs   =       9887
                                                  LR chi2(6)      =     923.09
                                                  Prob > chi2     =     0.0000
Log likelihood = -7062.4065                       Pseudo R2       =     0.0613

------------------------------------------------------------------------------
     JOBTYPE |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   na wlasny |
```

```
     _ISEX_2 |  -.4521757   .0498364    -9.07   0.000    -.5498534   -.3544981
         AGE |  -.0146795   .0130072    -1.13   0.259     -.040173    .0108141
        AGE2 |   .0007082   .0001532     4.62   0.000     .0004079    .0010085
       _cons |  -1.555082    .265707    -5.85   0.000    -2.075858   -1.034306
-------------+----------------------------------------------------------------
  pomagajacy |
     _ISEX_2 |   .4042479   .0899746     4.49   0.000     .2279009     .580595
         AGE |  -.3076462   .0165635   -18.57   0.000      -.34011   -.2751824
        AGE2 |   .0038472   .0001926    19.98   0.000     .0034698    .0042246
       _cons |   2.761419   .3234381     8.54   0.000     2.127492    3.395346
------------------------------------------------------------------------------
(Outcome JOBTYPE==      pracownik najemny       is the comparison group)
```

Następnie oszacowano efekty krańcowe dla alternatywy "pomagający członek rodziny" i uzyskano następujące oszacowania tych efektów:

```
Marginal effects after mlogit
      y  = Pr(JOBTYPE==30) (predict, outcome(30))
         =   .0501925
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 _ISEX_2*|    .025047      .00439    5.71   0.000    .01645  .033644   .461717
     AGE |  -.0144944      .00077  -18.73   0.000  -.016011 -.012978   39.9254
    AGE2 |   .0001751      .00001   19.34   0.000   .000157  .000193   1735.44
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

1. Wypisać założenia modelu wielomianowego logitowego.
2. Sprawdzić, czy zmienne w modelu są łącznie istotne. Zinterpretować Pseudo R2.
3. Podać, które zmienne w modelu są istotne. Zinterpretować znaki parametrów przy zmiennej SEX.
4. Wyjaśnić dlaczego w tym przypadku efekty cząstkowe liczymy osobno dla poszczególnych alternatyw. Zinterpretować policzony efekt cząstkowy dla zmiennej SEX dla alternatywy "pomagający członek rodziny".
5. Policzyć cząstkowy wieku dla alternatywy "pomagający członek rodziny" i zinterpretować obliczoną wartość.
6. Postanowiono sprawdzić hipotezę, że miejsce stan cywilny ma wpływ na typ podejmowanej pracy. Dodana do modelu zmienna MARSTAT ma 4 poziomy. W modelu uwzględniającym tą dodatkową zmienną uzyskano wielkość funkcji wiarygodności na poziomie -7042.06. Zweryfikuj hipotezę o tym, że stan cywilny nie wpływa na rodzaj podejmowanej pracy.

   Podpowiedź: $\chi^2_{0.95}(1) = 3.84$, $\chi^2_{0.95}(2) = 5.99$, $\chi^2_{0.95}(3) = 7.81$, $\chi^2_{0.95}(4) = 9.49$

Rozwiązanie:

1. W modelu wielomianowego logita prawdopodobieństwo wyboru J-tej alternatywy jest dane wzorem

$$p_j(\boldsymbol{x}) = \frac{\exp(\boldsymbol{x}\boldsymbol{\beta}_j)}{1 + \sum_{h=1}^{J} \exp(\boldsymbol{x}\boldsymbol{\beta}_h)}$$

2. Wielkość statystyki LR, wskazuje na to, że hipotezę zerową o nieistotności wszystkich zmiennych należy silnie odrzucić $[923.09, 0.000 < 0.05]$. Wielkość statystyki Pseudo R^2 wskazuje na to, że model wyjaśnił około 6.13 procent zróżnicowania zmiennej zależnej.
3. Zmiennymi istotnymi w równaniu dla alternatywy związanej z samozatrudnieniem jest płeć $[-9.07, 0.000 < 0.05]$, $AGE2$ $[4.62, 0.000 < 0.05]$ i stała $[-5.85, 0.000 < 0.05]$. Dla drugiej alternatywy istotnymi parametrami są płeć $[4.49, 0.000 < 0.05]$, AGE $[-18.57, 0.000 < 0.05]$, $AGE2$ $[19.98, 0.000 < 0.05]$ i stała $[8.54, 0.000 < 0.05]$.

Znaki przy parametrach mają związek z prawdopodobieństwami w stosunku do alternatywy bazowej. Alternatywą bazową jest bycie pracownikiem najemnym. Ujemny znak przy zmiennej SEX_2 w przypadku alternatywy "na własny rachunek" oznacza, w stosunku do prawdopodobieństwa alternatywy bazowej kobiety mają niższe prawdopodobieństwo pracy na własny rachunek niż mężczyźni. Podobnie dodatni współczynnik przy zmiennej SEX_2 w przypadku alternatywy "pomagający członek rodziny" oznacza, w stosunku do prawdopodobieństwa alternatywy bazowej kobiety mają wyższe prawdopodobieństwo bycia pomagającym członkiem rodziny niż mężczyźni.

4. W przypadku wielomianowego logita wielkość efektów cząstkowych jest różna dla każdej alternatywy. Dlatego musimy je liczyć z osobna. Wielkość efektu cząstkowego przy zmiennej SEX_2 oznacza, że kobiety mają o 2% wyższe prawdopodobieństwo bycia pomagającym członkiem rodziny niż mężczyźni o tych samych i równych średnim w próbie charakterystykach.

5. Licząc jak w zadaniu **??** uzyskujemy:

$$-.01449441 + 2 \times 39.92 \times .0001750 = -0.000522$$

a więc dodatkowy rok życia w przypadku osób o wieku równej średniej w próbie (39.92), obniża prawdopodobieństwo bycia pomagającym członkiem rodziny o 0.05 punktu procentowego.

6. Statystyka testu LR ma postać:

$$LR = 2\left(-7042.06 + 7062.41\right) = 40.7 > \chi^2_{0.95}\left(3\right) = 5.99$$

Testowana jest zerowość 3 współczynników (jeden poziom zmiennej dyskretnej wypada). Odrzucamy H_0 o łącznej nieistotności dochodu i zmiennej zerojedynkowej zdefiniowanej dla osób żyjących w separacji.

ZADANIE 2 W prostym modelu popytu i podaży (w zapisie pominięto indeksy obserwacji):

$$\begin{aligned} Q_D &= \alpha_0 + \alpha_1 P + \varepsilon_1 \\ Q_S &= \beta_1 P + \varepsilon_2 \\ Q_D &= Q_S \end{aligned}$$

1. Sprawdzić identyfikację równań.

2. Wyprowadzić estymator Pośredniej MNK dla parametrów w równaniach, które są zidentyfikowane.

3. Wyprowadzić postać estymatora MZI dla parametrów w równaniach, które są zidentyfikowane.

4. Policzyć granicę według prawdopodobieństwa estymatora MNK parametru β_1. Czy estymator ten jest zgodny? Założyć, że $\mathrm{E}\left(\varepsilon_1\right) = \mathrm{E}\left(\varepsilon_2\right) = 0$, $\mathrm{Var}\left(\varepsilon_1\right) = \sigma_1^2$,$\mathrm{Cov}\left(\varepsilon_1, \varepsilon_2\right) = \sigma_{12}$, $\overline{P^2} = \frac{1}{n}\sum_{t=1}^{T} P_t^2 \xrightarrow{p} P^{2*}$.

Rozwiązanie:

1. Sprawdzamy identyfikację równań

zmienne egzogeniczne	1
zmienne endogeniczne	Q_D, Q_S, P

$$\begin{array}{lllc} K=1 & G_1 = 2 & K_1 = 1 & \text{niezidentyfikowane } 1 = K < G_1 + K_1 - 1 = 2 \\ & G_2 = 2 & K_2 = 0 & \text{zidentyfikowane } 1 = K \geq G_1 + K_1 - 1 = 1 \end{array}$$

Równanie popytu nie jest zidentyfikowane, równanie podaży jest zidentyfikowane

2. Budujemy formę zredukowaną. Jedyną zmienną egzogeniczną jest stała, forma zredukowana ma postać:

$$\begin{aligned} Q_D = Q_S &= \pi_0 + \epsilon_1 \\ P &= \pi_1 + \epsilon_2 \end{aligned}$$

Rozwiązując formę strukturalną dla Q_D, Q_S i P otrzymujemy

$$Q_D = Q_S = \frac{\beta_1\alpha_0}{\beta_1-\alpha_1} + \frac{\beta_1\varepsilon_1-\alpha_1\varepsilon_2}{\beta_1-\alpha_1}$$
$$P = \frac{\alpha_0}{\beta_1-\alpha_1} + \frac{\varepsilon_1-\varepsilon_2}{\beta_1-\alpha_1}$$

Zależności między parametrami formy strukturalnej i zredukowanej

$$\pi_0 = \frac{\beta_1\alpha_0}{\beta_1-\alpha_1}$$
$$\pi_1 = \frac{\alpha_0}{\beta_1-\alpha_1}$$

Wynika z tego, że estymator Pośredniej MNK parametru β_1 można policzyć jako $\widehat{\beta}_1 = \frac{\widehat{\pi}_0}{\widehat{\pi}_1}$. Ponieważ estymatorami MNK stałej w modelu tylko ze stała są średnie zmiennych zależnych więc $\widehat{\beta}_1 = \frac{\overline{Q}}{\overline{P}}$.

3. Jedynym zidentyfikowanym równaniem jest równanie podaży $\boldsymbol{Y} = [Q_{S1}, \ldots, Q_{sT}]'$ a jedyną zmienną instrumentalną stała $\boldsymbol{Z} = [1, \ldots, 1]'$. Mamy dokładnie tyle zmiennych objaśniających ($\boldsymbol{X} = [P_1, \ldots, P_T]'$) Postać estymatora MZI w tym przypadku będzie następująca:

$$\boldsymbol{b}_{MZI} = \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{Z}'\,\boldsymbol{y} = \frac{\overline{Q}}{\overline{P}}$$

a więc dokładnie ten sam wzór co w przypadku zastosowania Pośredniej MNK. Sytuacja taka zachodzi zawsze, gdy równanie jest dokładnie zidentyfikowane.

4. Wzór na estymator MNK w równaniu podaży jest następujący

$$b = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\,\boldsymbol{y} = \frac{\sum_{t=1}^{T} P_t Q_{St}}{\sum_{t=1}^{T} P_t^2} = \frac{\sum_{t=1}^{T} P_t\left(\beta_1 P_t + \varepsilon_2\right)}{\sum_{t=1}^{T} P_t^2}$$
$$= \beta_1 + \frac{\frac{1}{n}\sum_{t=1}^{T} P_t\varepsilon_2}{\frac{1}{n}\sum_{t=1}^{T} P_t^2}$$

Korzystając z wyprowadzonego wzoru na P_t w formie zredukowanej, otrzymujemy

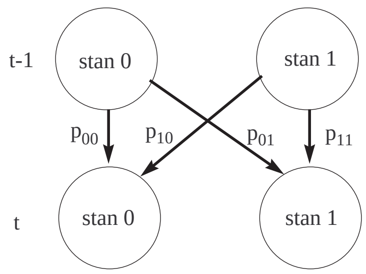
$$\frac{1}{n}\sum_{t=1}^{T} P_t\varepsilon_2 = \frac{\alpha_0}{\beta_1-\alpha_1}\frac{1}{n}\sum_{t=1}^{T}\varepsilon_2 + \frac{1}{\beta_1-\alpha_1}\frac{1}{n}\sum_{t=1}^{T}\varepsilon_1\varepsilon_2 - \frac{1}{\beta_1-\alpha_1}\frac{1}{n}\sum_{t=1}^{T}\varepsilon_1^2$$

Ponieważ jednak $\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{T}\varepsilon_2\right) = \operatorname{E}\left(\varepsilon_2\right) = 0$, $\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{T}\varepsilon_1\varepsilon_2\right) = \operatorname{Cov}\left(\varepsilon_1, \varepsilon_2\right) = \sigma_{12}$, $\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{T}\varepsilon_1^2\right) = \operatorname{E}\left(\varepsilon_1^2\right) = \sigma_1^2$ więc

$$\operatorname{plim}(b) = \beta_1 + \operatorname{plim}\left(\frac{\frac{1}{n}\sum_{t=1}^{T} P_t\varepsilon_2}{\frac{1}{n}\sum_{t=1}^{T} P_t^2}\right)$$
$$= \beta_1 + \frac{\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{T} P_t\varepsilon_2\right)}{\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{T} P_t^2\right)}$$
$$= \beta_1 + \frac{1}{\beta_1-\alpha_1}\frac{\sigma_{12}-\sigma_1^2}{P^{2*}}$$

Estymator MNK nie jest zgodny.

ZADANIE 3 Osoby w wieku produkcyjnym mogą znaleźć się w jednym z dwóch stanów 1 - praca, 0 - brak pracy. Zakładamy, że posiadamy informację z T okresów na temat stanów w jakich znajdowały się osoby w N elementowej próbie losowej osób w wieku produkcyjnym. Zakładamy dodatkowo, że prawdopodobieństwo znalezienia się w określonym stanie zależy tylko od tego w jakim stanie jednostka znajdowała się w poprzednim okresie. Jako p_{ij} oznaczono prawdopodobieństwo, że osoba znajdująca się w stanie i w okresie $t-1$ znalazła się w stanie j w



Rysunek 1: Przejścia między stanami

okresie t. Ponieważ osoba znajdująca się w stanie i w okresie $t-1$ musi znajdować się w jednym ze stanów w okresie t, więc $p_{00}+p_{01}=1$, $p_{11}+p_{10}=1$.
Oznaczono jako $n_{ij,t}$ liczbę osób, które były w okresie $t-1$ w stanie i a w okresie t w stanie j. Podobnie oznaczono przez $n_{i,t}=n_{0i,t}+n_{1i,t}$ ogólną liczbę osób, która w okresie t znajduje się w stanie i. Próba jest losowa a więc przejścia między stanami dla poszczególnych osób są niezależne.

1. Znaleźć funkcję wiarogodności dla postawionego problemu.

   Podpowiedź: prawdopodobieństwo liczby przejść ze stanu i do stanu j w okresie t jest dane rozkładem dwumianowym z liczbą prób równą $n_{i,t-1}$, liczbą sukcesów równą $n_{ij,t}$ i liczbą porażek równą $n_{i,t-1}-n_{ij,t}$.

2. Znaleźć postać estymatora MNW parametrów p_{00} i p_{11}.

3. Dysponujemy jedynie informacją na temat liczby $n_{i,t}$ jednostek znajdujących się w danym stanie. Sformułować układ równań, z którego można wyznaczyć estymator GMM parametrów p_{11} i p_{00}.

   Podpowiedź: skorzystać z postaci warunkowej wartości oczekiwanej $\mathrm{E}\left(n_{0,t}\right| n_{0,t-1}, n_{1,t-1})$.

4. Uogólnić ten estymator na przypadek, gdy prawdopodobieństwa przejść zależą od zmiennych makroekonomicznych, założyć, że $p_{11}\left(\boldsymbol{x}_t\right)=F\left(\boldsymbol{x}_t\boldsymbol{\beta}_0\right)$ i $p_{00}\left(\boldsymbol{x}_t\right)=F\left(\boldsymbol{x}_t\boldsymbol{\beta}_1\right)$, gdzie $F\left(\cdot\right)$ jest dystrybuantą pewnego rozkładu.

Rozwiązanie:

1. Prawdopodobieństwo, że $n_{11,t}$ osób znajdujących spośród $n_{1,t-1}$ znajdujących się w stanie 1 pozostanie w stanie 1 a $n_{1,t-1}-n_{11,t}$ przejdzie do stanu 0 wynosi:

   $$\Pr\left(n_{11,t}\right| n_{1,t-1}) = \binom{n_{1,t-1}}{n_{11,t}} p_{11}^{n_{11,t}}\left(1-p_{11}\right)^{n_{1,t-1}-n_{11,t}}$$

   Podobnie prawdopodobieństwo $n_{00,t}$ osób znajdujących spośród $n_{0,t-1}$ znajdujących się w stanie 0 pozostanie w stanie 0 a $n_{0,t-1}-n_{00,t}$ przejdzie do stanu 1 wynosi:

   $$\Pr\left(n_{00,t}\right| n_{0,t-1}) = \binom{n_{0,t-1}}{n_{00,t}} p_{00}^{n_{00,t}}\left(1-p_{00}\right)^{n_{0,t-1}-n_{00,t}}$$

   Zdarzenia te są niezależne więc funkcja prawdopodobieństwa dla obserwacji t ma postać:

   $$L_t\left(p_{11},p_{00}\right) = \binom{n_{1,t-1}}{n_{11,t}}\binom{n_{0,t-1}}{n_{00,t}} p_{11}^{n_{11,t}}\left(1-p_{11}\right)^{n_{1,t-1}-n_{11,t}} p_{00}^{n_{00,t}}\left(1-p_{00}\right)^{n_{0,t-1}-n_{00,t}}$$

   Oznaczmy jako $c_t=\binom{n_{1,t-1}}{n_{11,t}}\binom{n_{0,t-1}}{n_{00,t}}$.Z racji na niezależność obserwacji funkcja wiarygodności ma postać

   $$L\left(p_{11},p_{00}\right)=\prod_{t=1}^{T} c_t p_{11}^{n_{11,t}}\left(1-p_{11}\right)^{n_{1,t-1}-n_{11,t}} p_{00}^{n_{00,t}}\left(1-p_{00}\right)^{n_{0,t-1}-n_{00,t}}$$

Logarytm funkcji wiarygodności wynosi:

$$\begin{aligned}\ell\left(p_{11}, p_{00}\right) = &\sum_{t=1}^{T} \ln c_t + \ln p_{11} \sum_{t=1}^{T} n_{11,t} + \ln\left(1 - p_{11}\right) \sum_{t=1}^{T} \left(n_{1,t-1} - n_{11,t}\right) \\ &+ \ln p_{00} \sum_{t=1}^{T} n_{00,t} + \ln\left(1 - p_{00}\right) \sum_{t=1}^{T} \left(n_{0,t-1} - n_{00,t}\right)\end{aligned}$$

2. Rozwiązując warunki pierwszego rzędu otrzymujemy estymator MNW

$$\begin{aligned}\frac{\partial \ell\left(p_{11}, p_{00}\right)}{\partial p_{11}} &= \frac{1}{p_{11}} \sum_{t=1}^{T} n_{11,t} - \frac{1}{1 - p_{11}} \sum_{t=1}^{T} \left(n_{1,t-1} - n_{11,t}\right) = 0 \\ \frac{\partial \ell\left(p_{11}, p_{00}\right)}{\partial p_{00}} &= \frac{1}{p_{00}} \sum_{t=1}^{T} n_{00,t} - \frac{1}{1 - p_{00}} \sum_{t=1}^{T} \left(n_{0,t-1} - n_{00,t}\right) = 0\end{aligned}$$

ponieważ c_t nie zależy od p_{ij}. Rozwiązując otrzymujemy

$$\begin{aligned}\widetilde{p}_{11} &= \frac{\sum_{t=1}^{T} n_{11,t}}{\sum_{t=1}^{T} n_{1,t-1}} \\ \widetilde{p}_{00} &= \frac{\sum_{t=1}^{T} n_{11,t}}{\sum_{t=1}^{T} n_{1,t-1}}\end{aligned}$$

Estymatory pozostałych prawdopodobieństw znajdujemy jako $\widetilde{p}_{10} = 1 - \widetilde{p}_{11}$ a $\widetilde{p}_{01} = 1 - \widetilde{p}_{00}$.

3. Wartość oczekiwana

$$\begin{aligned}\mathrm{E}\left(n_{0,t} \middle| n_{0,t-1}, n_{1,t-1}\right) &= \mathrm{E}\left(n_{00,t} + n_{10,t} \middle| n_{0,t-1}, n_{1,t-1}\right) = \mathrm{E}\left(n_{00,t} \middle| n_{0,t-1}\right) + \mathrm{E}\left(n_{10,t} \middle| n_{1,t-1}\right) \\ &= p_{00} n_{0,t-1} + \left(1 - p_{11}\right) n_{1,t-1}\end{aligned}$$

przy czym skorzystaliśmy z własności wartości oczekiwanej rozkładzie dwumianowym i tego, że $p_{10} = 1 - p_{00}$. Wynika z tego, dla

$$\varepsilon_t = n_{0,t} - p_{00} n_{0,t-1} - \left(1 - p_{11}\right) n_{1,t-1}$$

Wartość oczekiwana

$$\mathrm{E}\left(\varepsilon_t \middle| n_{0,t-1}, n_{1,t-1}\right) = 0$$

Zgodnie z zasadami wyprowadzania estymatorów GMM formułujemy momenty bezwarunkowe

$$\begin{aligned}\mathrm{E}\left(n_{0,t-1} \varepsilon_t\right) &= 0 \\ \mathrm{E}\left(n_{1,t-1} \varepsilon_t\right) &= 0\end{aligned}$$

Estymator GMM znajdujemy rozwiązując dla p_{00} i p_{11} układ równań:

$$\begin{cases} \frac{1}{N} \sum_{t=1}^{T} n_{0,t-1} \left[n_{0,t} - p_{00} n_{0,t-1} - \left(1 - p_{11}\right) n_{1,t-1}\right] = 0 \\ \frac{1}{N} \sum_{t=1}^{T} n_{1,t-1} \left[n_{0,t} - p_{00} n_{0,t-1} - \left(1 - p_{11}\right) n_{1,t-1}\right] = 0 \end{cases}$$

dla p_{00} i p_{11}.

4. W takim przypadku

$$\begin{aligned}\varepsilon_t = n_{0,t} - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_0\right) n_{0,t-1} - \left[1 - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_1\right)\right] n_{1,t-1} \\ \mathrm{E}\left(\varepsilon_t \middle| n_{0,t-1}, n_{1,t-1}, \boldsymbol{x}_t\right) = 0\end{aligned}$$

Jedną z opcji byłby estymator GMM wychodzący z następujących momentów bezwarunkowych:

$$\begin{aligned}\mathrm{E}\left(\boldsymbol{x}_t n_{0,t-1} \varepsilon_t\right) &= 0 \\ \mathrm{E}\left(\boldsymbol{x}_t n_{1,t-1} \varepsilon_t\right) &= 0\end{aligned}$$

a więc będący rozwiązaniem dla $\boldsymbol{\beta}_0$ i $\boldsymbol{\beta}_1$ układu równań:

$$\begin{cases} \frac{1}{N} \sum_{t=1}^{T} \boldsymbol{x}_t n_{0,t-1} \left[n_{0,t} - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_0\right) n_{0,t-1} - \left[1 - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_1\right)\right] n_{1,t-1}\right] = 0 \\ \frac{1}{N} \sum_{t=1}^{T} \boldsymbol{x}_t n_{1,t-1} \left[n_{0,t} - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_0\right) n_{0,t-1} - \left[1 - F\left(\boldsymbol{x}_t \boldsymbol{\beta}_1\right)\right] n_{1,t-1}\right] = 0 \end{cases}$$