

<h2>Panel modellek alapjai</h2>	<p><b>Olvasási idő:</b> 30 perc</p> 	<p><b>Készítette:</b> Szládek Dániel</p> 
---------------------------------	---	--

### Definíció

- A panel adatbázisok keresztmetszeti ( $N$ ) és idősoros ( $T$ ) dimenzióval is rendelkeznek, ezáltal többlet információt nyújtanak a pusztán keresztmetszeti vagy csak idősoros adatbázisokhoz képest. A panel adatok elemzésekor a függő változó ( $y$ ) értékét olyan magyarázó változókkal ( $x$ ) becsüljük meg, amelyekről tehát keresztmetszeti és idősoros adatok is elérhetők. Ennél a modelltípusnál is léteznek nem megfigyelt változók, amelyeket a hibatag ( $u_i$ ) foglal magába.
- A keresztmetszeti dimenziót *csoportoknak* nevezzük, például országok, vállalatok, de akár egyének is jelenthetik az egyes csoportokat. Az *idősoros* dimenzió lehet év, negyedév, hónap, de elképzelhető bármilyen gyakoriságú időegység, akár napi szintű adatok is.
- Mivel három ismérvet (csoportok, idősor, változók értékei) tartalmaz a panel adatbázis, ezért ezek felépítése a következő: a változók (függő és magyarázó változók is) értékeit oszlopokban jelenítjük meg, úgy, hogy csoportonként időrendbe rendezve következnek egymás után az adatok. Az adatbázis tartalmaz csoportazonosítót, valamint idősoros azonosítót is<sup>34</sup>.

### Gravitációs modellek

A gravitációs modelleket elsősorban a mag- és periféria térségek közötti kereskedelmi kapcsolatok log-lineáris modellekkel mérésére hozták létre (Greene 2003):

$$\ln y = \text{const.} + \alpha_1 \ln X_1 + \dots + \alpha_n \ln X_n + \varepsilon$$

ahol minden  $k$  változóra ( $1 \leq k \leq n$ )  $X_k > 0$  és  $i$  országra ( $1 \leq i \leq m$ )  $X_k = x_{k,\text{mag}} - x_{k,i}$ .

A gazdasági aszimmetriák érzékeltetésére az első változó jellemzően a GDP, egy főre eső GDP vagy egyéb output vagy potenciális gazdasági erő kifejezésére szolgáló változó szokott lenni. A hagyományos kereskedelmi kapcsolatok vizsgálata (Brakman – Bergeijk 2010) mellett egyéb összeállításokban is lehetséges a használata, például a devizaárfolyam volatilitásával kiegészítve (Simáková 2016).

#### *a, Bemeneti követelmény: stacionaritás*

- Stacionárius bemenet (input) (Pesaran és Shin (2003) teszt,  $p < 0,05$  -> elvethető a nullhipotézis -> stacionárius bemenetek)
  - *mean of individual ADF statistics: P-value of the  $W_{\text{bar}}$  statistic = 0.0000*
  - *standardized IPS statistic based on the moments of the DF distribution: P-value of the  $Z_{\text{bar}}$  statistic = 0.0000*

<sup>34</sup> Hasznos információk a panel modellekről: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-panel-data/>

- *mean of individual DF statistics: P-value of the  $Z_{bar\_DF}$  statistic = 0.0000*
- *Cross-unit 1: Lag = 0 Adj. sample size = 531 ADF statistic = -22.9833 ADF p-value = 0.0100*
- *Cross-unit 2: Lag = 0 Adj. sample size = 531 ADF statistic = -10.0041 ADF p-value = 0.0100*

#### *b, Keresztmetszeti függőség*

- Események, például recessziók, globális sokkok, vagy a pénzügyi és gazdasági világválság a teljes mintánkra hatással lehetnek, emiatt keresztmetszeti függőség alakulhat ki a keresztmetszeti csoportok, a magyarázó változók és a hibatagok között. Ebben az esetben az eredmények értelmezése hibás következtetésekhez vezethet.
- A keresztmetszeti függőségét a Pesaran CD teszttel vizsgáljuk, amellyel megbizonyosodhatunk arról, hogy nincs keresztmetszeti függőség (korreláció) a reziduumok között. Ha  $p < 0,05$  -> nincs keresztmetszeti függőség (Pesaran 2007).
- Modell szelekció:
  - (1) **Chow Test:** olyan teszt, amely azt hivatott eldönteni, hogy Pooled OLS vagy fix hatás (fixed effect, FE) modell a megfelelő panelbecsléskor.
    - Pooled OLS esetén azt feltételezzük, hogy a meredekség ugyanaz minden csoportra minden időszakban, a Chow teszt kiterjesztése azt a nullhipotézist vizsgálja, hogy a magyarázó változók meredeksége (együtthatója) megegyezik minden egyénre vonatkozóan az összes  $k$  magyarázó változó esetében:  $H_0: \beta_{i,k} = \beta_k$ .
    - $H_0$ : Pooled OLS-t válasszuk ( $p > 0.05$ )
    - $H_1$ : FE modellt válasszuk ( $p < 0.05$ )
  - (2) **Hausman Test:** olyan teszt, amely segítségével arról döntünk, hogy a fix hatás (FE) vagy a véletlen hatás (random effect, RE) modell a megfelelő panelbecsléskor.
    - $H_0$ : RE modellt válasszuk ( $p > 0.05$ )
    - $H_1$ : FE modellt válasszuk ( $p < 0.05$ )
  - (3) **Test Lagrange Multiplier (LM):** olyan teszt, amellyel arról döntünk, hogy a RE vagy a Pooled OLS modell a megfelelő panelbecsléskor.
    - $H_0$ : Pooled OLS-t válasszuk ( $p > 0.05$ )
    - $H_1$ : RE modellt válasszuk ( $p < 0.05$ )

#### **Pooled, FE, RE**

- Hibatag: kapcsolódhat csoporthoz (időtől független -  $u$ ), de vonatkozhat a teljes regresszióra is (időtől függő -  $\varepsilon$ )
  - Autokorrelálatlan  $\varepsilon$

#### *a) Pooled OLS*

- A Pooled OLS a panel regressziók legegyszerűbb formája, hiszen ekkor az egyéni keresztmetszeti vagy időtől függő hatások hiányoznak ( $u_i = 0$ ):
  - $y_{i,t} = \omega + \beta X'_{i,t} + \varepsilon_{i,t}$ .
- Pooled OLS esetén azt feltételezzük, hogy a meredekség ugyanaz minden csoportra minden időszakban, a Chow teszt kiterjesztése azt a nullhipotézist vizsgálja, hogy a magyarázó változók meredeksége (együtthatója) megegyezik minden egyénre vonatkozóan az összes  $k$  magyarázó változó esetében:  $H_0: \beta_{i,k} = \beta_k$ .

*b) FE*

- Amennyiben egy magyarázó változó ( $x_i$ ) korrelál a nem megfigyelhető hibataggal ( $u_i$ ), akkor a fix hatás modell (FE) alkalmazására lesz szükségünk. Ha a Chow teszt nullhipotézisét elvetjük és időtől független fix hatásokat (pl.: egyén neve; vállalatok esetében az iparág, amelyben versenyeznek) szeretnénk figyelembe venni, akkor a modell a következő:

- $y_{i,t} = \omega + u_i + \beta X'_{i,t} + \varepsilon_{i,t}$ .

- A fix hatás model dummy változók segítségével is felírható:

- $y_{i,t} = u_i + \beta X'_{i,t} + \mu_1 D1_i + \mu_2 D2_i + \dots + \mu_N DN_i + \varepsilon_{i,t}$ .

- Létezik olyan változat is, amelyben az egyéntől független, de időspecifikus fix hatásokat vizsgálhatunk, pl.: szabályozási környezet vagy adórendszer változása. A kétfajta FE modell kombinálható is, azaz egyszerre figyelembe vehetjük az egyén és időspecifikus fix hatásokat is.
- A Hausman teszt nullhipotézise alapján nincs kapcsolat  $u_i$  és  $x_i$  magyarázó változó között. Ha ezt a hipotézist el tudjuk vetni ( $p < 0,05$ ), akkor a FE modellt érdemes alkalmazni a RE modellel szemben.

*c) RE*

- Ha nem vehető el a Hausman teszt nullhipotézise ( $p > 0,05$ ), azaz nincs kapcsolat  $u_i$  és  $x_i$  magyarázó változó között, akkor a véletlen hatás (random effect, RE) modellt kell alkalmazni. Az egyénekenként eltérő, de időben állandó véletlenszerű hatásokat becslő RE modell a következő formában írható fel:

- $y_{i,t} = \omega + \beta X'_{i,t} + (u_i + \varepsilon_{i,t})$ .

- A FE modellhez hasonlóan a RE modell is alkalmas arra, hogy az egyén és az idő szerepét felcseréljük. A RE eljárásban az egyénekenként eltérő, de időben állandó véletlenszerű változó helyett figyelembe vehetünk időszakonként eltérő és egyénekenként állandó véletlenszerű tényezőt is. Továbbá, a RE modellnél is figyelembe vehető mindkét hatás, azaz egy egyénekenként és egy időszakonként eltérő véletlenszerű változó beiktatásával alkalmazható egy ilyen technika.

## **Források**

- Álvarez, Inmacualda C., Javier Barbero, and José L. Zofio (2015) "A Panel Data Toolbox for MATLAB." Economic Analysis Working Paper Series Universidad Autonoma Madrid Working Paper 05/2013, ISSN: 1885–6888.
- Park, Hun Myoung (2011) "Practical Guides To Panel Data Modeling: A Step-by-step Analysis Using Stata." Tutorial Working Paper. Graduate School of International Relations, International University of Japan.
- Wooldridge, Jeffrey M. (2010) "Econometric Analysis of Cross Section and Panel Data." The MIT Press Cambridge, Massachusetts.
- Pesaran, M. H. (2007) "A simple panel unit root test in the presence of cross-section dependence." Journal of Applied Econometrics, 22, 2, 265-312.

## **Önellenőrző kérdések**

1. Hogyan kell felépíteni egy panel-adatokat tartalmazó adattáblát?
2. Hogyan döntünk egy FE és egy RE modell alkalmazása között?
3. Mit jelent a keresztmetszeti függőség?

Jelen tananyag a Szegedi Tudományegyetemen készült az Európai Unió támogatásával. Projekt azonosító: EFOP-3.4.3-16-2016-00014

