



PÉCSI TUDOMÁNYEGYETEM  
♦ JUBILEUM 650 ♦  
UNIVERSITY OF PECS JUBILEE

# A modellezés sajátosságai anomáliákkal terhelt idősorok esetén

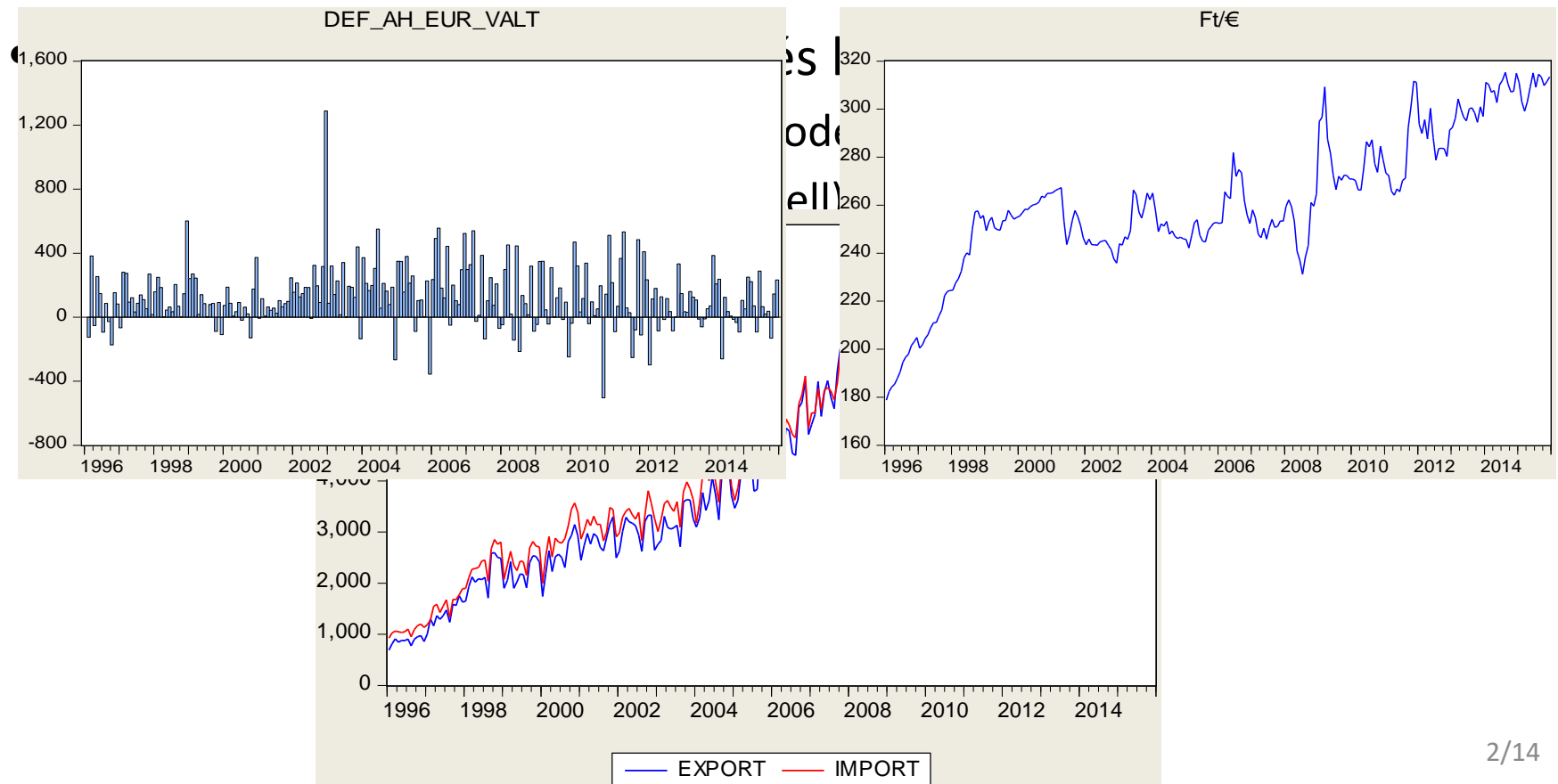
MÓDSZERTANI DILEMMÁK A STATISZTIKÁBAN – 40 ÉVE ALAKULT A  
JÖVŐKUTATÁSI BIZOTTSÁG

SJTB Tudományos ülés, 2016. november 18.



# Idősor-modellezés alapkérdései

- Vizsgált jelenség stacioner vagy trendet tartalmaz?
  - Kiterjesztett Dickey-Fuller-próba



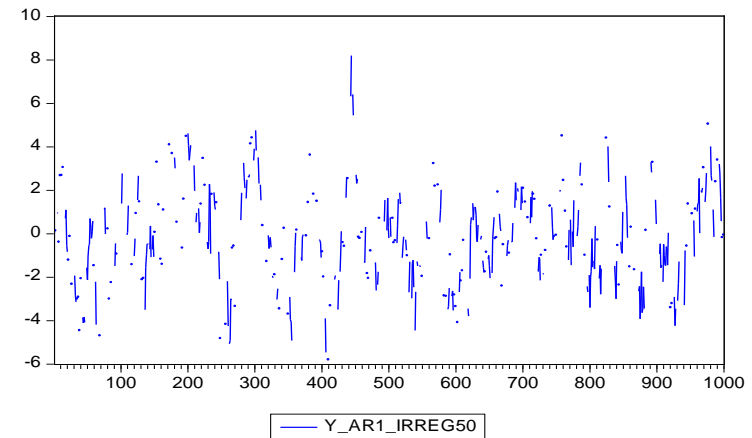
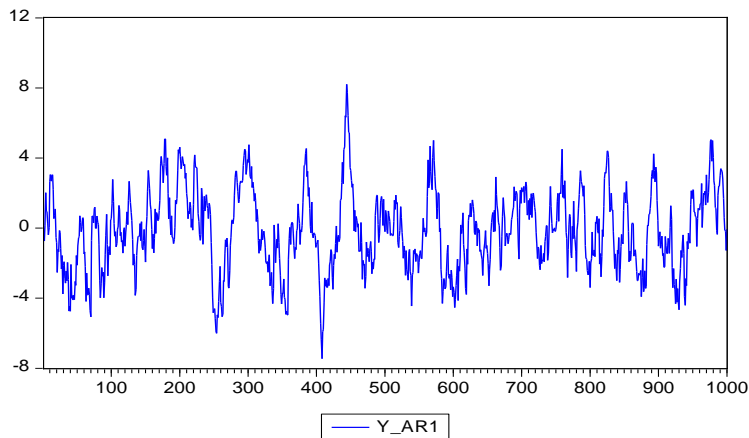


# Tankönyv és gyakorlat különbözősége

Tankönyv (elvárások)	Gyakorlat (empirikus idősorok)
Ekvidisztáns	Rendszertelen
Egy adatgeneráló folyamat	Outlierek találhatóak
Törésmentes	Strukturális törést tartalmaz
Elemi megfigyelések	Időben aggregált megfigyelések



1. Probléma és következményeinek rövid bemutatása
2. Kimutatásának, kiszűrésének, illetve kezelésének módszertana
3. Adatgeneráló folyamatok (DGP) előállítása
  - ismert tulajdonságokkal bíró (AR1, VAR, RW, EC)
  - (előbbi) ismert tulajdonságokkal bíró, de a vizsgált anomáliával terhelt
4. Monte Carlo szimulációval elemezni milyen arányban fordul elő első-, illetve másodfajú hiba, ha
  - az anomáliát figyelmen kívül hagyjuk
  - az eredeti idősoron feleslegesen végzünk anomália-kezelést





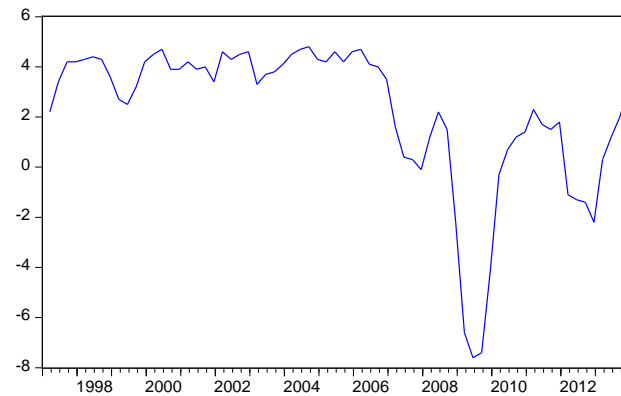
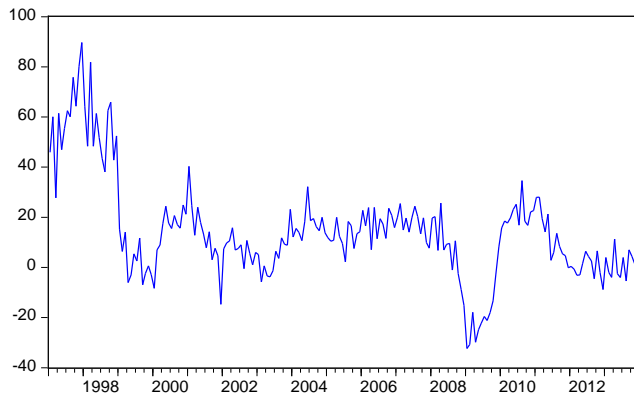
# Rendszertelen idősorok

- Nemekvidisztáns lehet, mert
  - hiányzó adat
  - valódi rendszertelenség
- Kezelése
  - (lineáris, log-lineáris, vagy spline) interpolációval kiegészítés
  - folytonos idejű autoregresszív (CAR) modellek
- Tézisek
  1. Egyváltozós esetben a **stacionaritás**, vagy **trend** meglétének **felismerését nem veszélyezteti** az interpoláció, általában már a lineáris interpoláció is kielégítő.
  2. Eredetileg Granger-okságban levő idősoroknál, ha az ok sokkal ritkábban megfigyelt, mint az okozat, akkor gyakran **„elfedődhet” az okság** (fals negatív eset).
  3. Kointegrált idősorok esetén az interpoláció **elfedheti a közös trendet** (kifejezetten veszélyes a nem azonos gyakoriságú, hosszabb szakaszon kiegészített idősorok együttmozgásának vizsgálata).



## Illusztratív példa nem azonos frekvencián megfigyelt idősorok közötti okság

- Magyarország export (havi bontás) és GDP (negyedéves bontás) változása az előző év azonos időszakának százalékában (1997-2013)



- Wald-próba (Granger-okság: export ok, interpolációval kiegészített GDP okozat)

Interpoláció módszere	F-érték	p-érték
kiegészítés	1,091	0,371
lineáris interpoláció	0,890	0,558
Catmul-Rom spline	0,865	0,584
harmadfokú, másodrendű spline	1,896	0,038



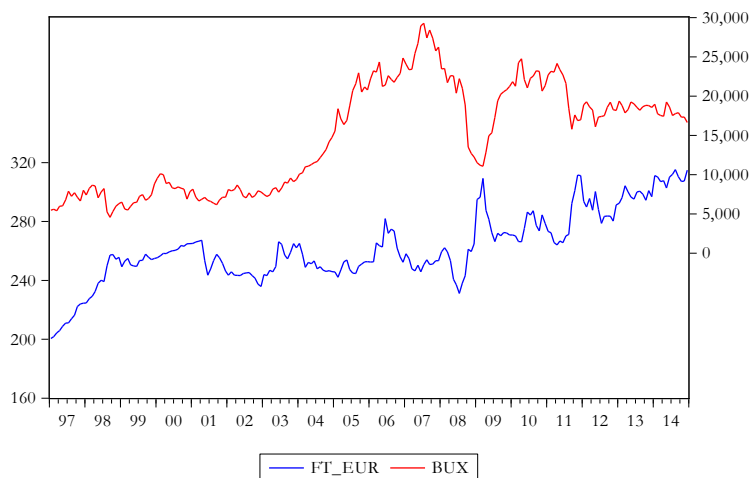
# Outlierek

- Outlier (kiugró érték, vélhetően más DGP-ből származik)
  - gross error (adatrögzítés, stb.)
  - true (várakozásoknak ellentmondó)
- Kiszűrése
  - modell független technikák (SD-, z-score, Mahalanobis-, stb.)
  - modellalapú (LR, érzékenységvizsgálat, saját)
- Kezelése
  - elhagyás (→ nemekvidisztáns idősorok)
  - helyettesítés (csonkolás)
  - becslési módszer változtatás
- Tézisek
  1. Stacionaritás helyett trendet ritkán lehet találni, de a trend gyakran elfedődik (fals negatív eset).
  2. Okság esetén (akár stacionárius változókról, akár kointegrációról van szó) az outlierek elfedhetik az ok-okozati összefüggést (fals negatív eset). Az okban megjelenő outlier veszélyesebb.
  3. Felesleges winsorizálás elfedheti a sztochasztikus trendet, vagy a közös trendet.

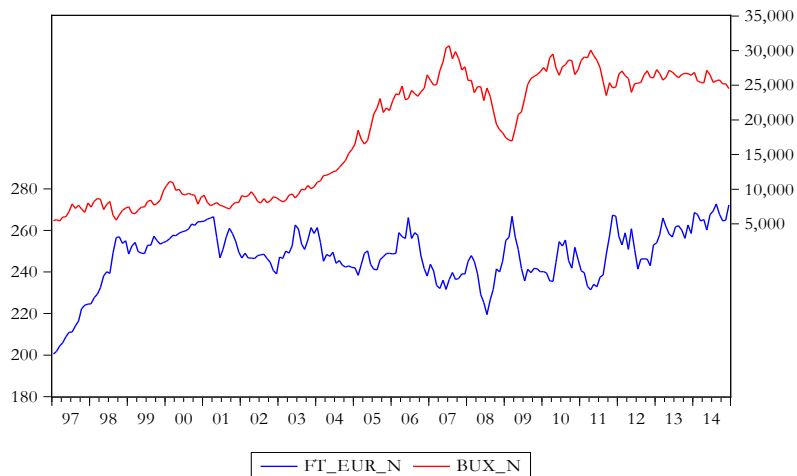


# Illusztratív példa kiugró értékeket tartalmazó tőzsdei adatokra

- BUX és €/Ft-árfolyam (1997-2014) – nem kointegráltak



- Ugyanez a winszorizált differenciák-kumulálásával – kointegráltak







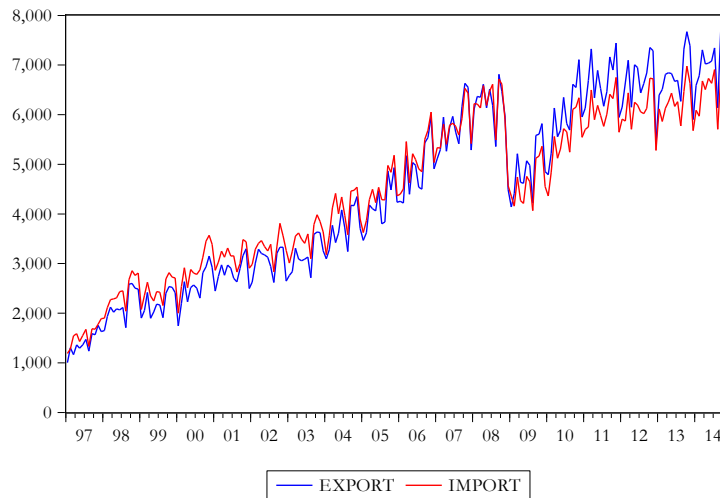
# Strukturális törés az idősorokban

- Az adatgeneráló folyamat paramétereinek megváltozása a vizsgált időhorizonton (a priori ismert, vagy ismeretlen helyen)
- Kiszűrése
  - Chow-próba (egy törés ismert hely)
  - Quandt-Andrews (egy törés ismeretlen hely)
  - Bai, Bai-Perron, Banerjee (több törés, különböző paraméterekben, ismeretlen helyen)
  - Johansen (több törés VAR-modellben)
- Kezelése: modellspecifikáció megváltoztatása (alkalmas dummy változók modellbe építése)
- Tézisek
  1. ADF-próba nagyon érzékeny a strukturális törésre, főképpen a szinteltolásra.
  2. Ha több törés van, de végül visszatérünk az eredeti DGP-re, akkor a Quandt-Andrews szinte használhatatlan.
  3. Közös trendet nagy valószínűséggel elfedi az egyik idősorban bekövetkezett törés, még akkor is, ha a másik viszonylag kis idő múlva követi.



# Illusztratív példa a magyar külkereskedelem idősorain

- Export – Import (folyó áron €-ban, 1997-2014) – nincs közös trend



- Strukturális törés feltételezésével - kointegráltak

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ D_{1t} \\ D_{2t} \end{bmatrix} + \begin{bmatrix} \gamma_{11}^x & \gamma_{12}^x & \gamma_{13}^x \\ \gamma_{21}^x & \gamma_{22}^x & \gamma_{23}^x \end{bmatrix} \begin{bmatrix} \Delta x_{t-1} \\ \Delta x_{t-2} \\ \Delta x_{t-3} \end{bmatrix} + \begin{bmatrix} \gamma_{11}^y & \gamma_{12}^y & \gamma_{13}^y \\ \gamma_{21}^y & \gamma_{22}^y & \gamma_{23}^y \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \Delta y_{t-2} \\ \Delta y_{t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$



# Időbeli aggregálás és következményei

- Típusai

- stock (szisztematikus mintavétel)
- flow (összegzés, esetleg átlagolás)

$$y_t^* = \sum_{j=0}^A w_j y_{t-j} = W(L) y_t$$

- Analitikusan belátható összefüggések

- ARMA(p,q) → ARMA(p,r)

$$r = \text{int} \left[ \frac{(p+1)(k-1) + q}{k} \right]$$

- RW ~ ARIMA(0,1,0) → ARIMA(0,1,1)
- okság módosul
- kointegráltság nem

- Tézisek

1. Trend, illetve stacionaritás felismerését nem zavarja.
2. Gyakran előfordul, hogy az aggregált idősorokban nem megfigyelhető okság, magasabb frekvencián kimutatható.
3. Ok-okozati viszonyok elfedésének valószínűsége az aggregálás rendjével párhuzamos növekszik (haviról évesre, illetve munkanapiról havira áttérésnél már zavaró mértékű).



## Devizaárfolyamok (havi záró, illetve havi átlag közötti kointegráció tesztelése)

- Euro (€) és néhány további fontos árfolyam közötti Johansen-teszt eredménye (2000-2009 havi bontás) – p-értékek

Deviza	Havi záró	Havi átlag
CHF	0,026	0,072
GBP	0,506	0,521
JPY	0,018	0,073
USD	0,014	0,081



## Tanácsok az etikus modellezéshez

- *A modell specifikációja során mindig törekedjünk a vizsgálandó jelenséget a lehető legjobban leíró, pontosan definiált idősorok használatára!*
- *Az empirikus idősor legyen a lehető leghosszabb és a lehető legnagyobb, még értelmezhető frekvenciájú, ugyanakkor igyekezzünk homogén (törésmentes) időszakot vizsgálni!*
- *Csak indokolt adat-transzformációkat használjunk, ezeket pontosan írjuk le a specifikáció során, és semmiképpen se modellezzünk olyan idősorokat, melyek a transzformációk következtében elveszítik értelmezhetőségüket!*
- *Sohase feledkezzünk meg arról, hogy a modell a valóság egyszerűsített változata, ahol az operacionalizálhatóság oltárán mindig feláldozunk valamennyit a valósághűségből! Körültekintően hívjuk fel erre a felhasználók figyelmét, jelezzük a mintából történő következtések korlátait, ismertessük a ceteris paribus elv következményeit!*



**Köszönöm a figyelmet !**