

# Vėlavimų įtaka išvadsoms apie kointegruotumą mažose imtyse \*

Virmantas KVEDARAS (VU)

el. paštas: vkved@takas.lt

## 1. Įvadas

Dauguma makroekonominių rodiklių yra nestacionarūs, todėl jų ryšiams tirti taikytina kointegruotumo analizė. Pastarojoje įprasta nagrinėti vienalaikius ryšius tarp kintamųjų. Tačiau ekonominiuose procesuose informacijos sklaida yra ribota, sprendimai dažniausiai priimami ir įgyvendinami su tam tikru vėlavimu, todėl dažnai stebimas vėluojantis vieno kintamųjų poveikis kitiems. Nors vėluojantys ryšiai asimptotinėms išvadsoms apie kointegruotumą įtakos neturi, tačiau Lietuvos ir kitų pereinamosios ekonomikos šalių praktikoje makroekonominių rodiklių skaičius yra nedidelis. Pavyzdžiui, dėl praėjusiame dešimtmetyje vykusių struktūrinių pokyčių ir pasikeitusios nacionalinių sąskaitų sistemos apie daugelį makroekonominių rodiklių Lietuvoje šiuo metu turima tik apie trisdešimt palyginamų ir palyginti patikimų stebėjimų. Todėl aktualu įvertinti, kokios įtakos išvadsoms apie kointegruotumą turi vėluojantys ryšiai mažose imtyse.

Šiame darbe taikant Monte Carlo (MC) modeliavimą nagrinėjama, kokią įtaką mažose imtyse kointegruotumo testų galiai daro vėluojančio poveikio ignoravimas. Tyrime apsiribojama dažniausiai kintamųjų kointegruotumo tikrinimui atskirose lygtyse taikomo požiūrio – Engle–Granger (EG) procedūros – nagrinėjimu. Analizuojami keli populiarius vienietinės šaknies bei stacionarumo testai (DF, ADF, KPSS) ir nagrinėjami atvejai, kai kointegravimo parametrai yra žinomi, bei kai jie yra įvertinami.

Atsižvelgiant į tai, ar kointegravimo parametrai yra žinomi, ar įvertinami, pasiūlomi procedūrų pakeitimai, leidžiantys gauti korektiškesnes išvadas mažose imtyse.

## 2. Kintamųjų kointegruotumas, Engle–Granger procedūra ir vėluojantis poveikis

Jeigu vienoda eile integruotų atsitiktinių kintamųjų tiesinė kombinacija yra integruota žemesne eile nei patys kintamieji, tai jie vadinami kointegruotais, o atitinkamas tiesinę kombinaciją nusakantis parametru vektorius vadinamas kointegravimo parametru vektoriumi. Engle ir Granger [1] pabrėžė, jog ekonominiuose procesuose aktualiausias yra

---

\*Darbą remia Lietuvos VMSF programa „Lietuvos ekonomikos matematiniai modeliai makroekonominiams procesams prognozuoti“ (registracijos Nr. C-03004).

kintamųjų, integruotų pirma eile, atvejis ir išskyrė du empirinio kointegravimo tikrinimo tipus: a) kai kointegravimo parametrų vektorius žinomas; b) kai jis nežinomas. Kai kointegravimo parametrai žinomi, tai pirma eile integruotų kintamųjų kointegravimas tikrinamas tiesiogiai nagrinėjant jų tiesinės kombinacijos stacionarumą. Kai kointegravimo parametrai nežinomi, Engle ir Granger pasiūlė taikyti dviejų žingsnių procedūrą, kurios pirmame žingsnyje mažiausių kvadratų metodu (MKM) yra įvertinama regresija tarp integruotų kintamųjų lygių, o antrame žingsnyje tikrinama, ar šios regresijos liekana turi vienetinę šaknį; jei taip, – kintamieji nėra laikomi kointegruotais.

Iš kointegravimo apibrėžimo nesunku matyti, kad kointegravimo savybė nepriklauso nuo laiko, t.y., nuo to, ar ryšys tarp kintamųjų yra vienalaikis, ar poveikis vėluoja. Pavyzdžiui, tegu analizuojama pirma eile integruotų atsitiktinių kintamųjų  $X_t \sim I(1)$  ir  $Y_t \sim I(1)$  stacionari tiesinė kombinacija

$$Y_t - \beta X_{t-k} = \varepsilon_t, \quad k > 0, \quad (1)$$

kur  $\varepsilon_t$  yra su  $X_t$  pokyčiais nekoreliuotas silpnas baltasis triukšmas. Jei vietoje (1) būtų nagrinėjamas vienalaikis ryšys tarp  $Y_t$  ir  $X_t$ , būtų gauta liekana

$$v_t = Y_t - \beta X_t = \varepsilon_t - \beta \sum_{i=0}^{k-1} \Delta X_{t-i}, \quad (2)$$

kur  $\Delta X_t = X_t - X_{t-1}$ . Liekana  $v_t$  taip pat yra stacionari ir nuo  $\varepsilon_t$  skiriasi tik stacionariu dydžiu  $-\beta(X_t - X_{t-k})$ . Tačiau baigtinėje realizacijoje, augant  $k$  ir jam artėjant prie realizacijos dydžio  $T$ , liekana  $v_t$  tampa vis panašesnė į generuojamą integruoto proceso, kadangi narys  $\sum_{i=0}^{k-1} \Delta X_{t-i}$  pradeda dominuoti. Todėl galima numanyti, kad vėlavimų specifikavimas baigtinėse imtyse gali turėti reikšmingos įtakos išvadoms apie kointegravimą, net kai parametrų reikšmės – nagrinėjamu atveju  $\beta$  – yra žinomos.

Kai  $\beta$  nežinomas ir jį reikia vertinti, atsiranda papildoma problema. Yra gerai žinoma [2], jog esant koreliuotoms kointegravimo regresijos liekanoms, nepaisant MKM įverčių supersuderinamumo pirmo EG procedūros žingsnio kointegravimo parametrų įverčiai mažose imtyse yra labai paslinkti, kas atitinkamai gali papildomai komplikuoti korektiškos išvados apie kintamųjų kointegravimą gavimą. Iš (2) lygties matyti, kad neteisingas vėlavimų specifikavimas dėl nario  $\sum_{i=0}^{k-1} \Delta X_{t-i}$  dažniausiai sąlygoja sekos  $v_t$  autokoreliuotumą, net jei  $\varepsilon_t$  ir  $\Delta X_t$  yra tarpusavyje nepriklausomi vienodai pasiskirstę dydžiai. Pastaruoju atveju nesunku parodyti (žr. [3]), kad pirmos eilės autokoreliacijos koeficientas priklauso nuo  $k$  ir, kai  $k \neq 0$ , turi išraišką  $\rho_v^k(1) = \frac{k-1}{k + \frac{\sigma_\varepsilon^2}{\sigma_{\Delta X}^2}}$ ,

kur  $\sigma_\varepsilon^2$  ir  $\sigma_{\Delta X}^2$  atitinkamai žymi homoskedastiškų procesų  $\varepsilon_t$  ir  $\Delta X_t$  dispersijas. Pavyzdžiui, kai  $\beta = \sigma_\varepsilon^2 = \sigma_{\Delta X}^2 = 1$ , tai  $\rho_v^1(1) = 0$ ,  $\rho_v^2(1) = 1/3$ ,  $\rho_v^3(1) = 1/2$ , o  $\rho_v^4(1) = 3/5$ . Matyti, jog augant  $k$ ,  $v_t$  pirmos eilės autokoreliacijos koeficientas artėja prie vienetinės šaknies atvejo.

Visa tai rodo galimą EG procedūra pagrįstų kointegravimo testų problemiškumą baigtinėse imtyse, kai ignoruojamas vėluojantis kintamųjų poveikis. Tą patvirtina 5 skyriuje aptariamai MC modeliavimo rezultatai, rodantys, jog mažose imtyse

vėlavimų ignoravimas gali žymiai sumažinti minėtų testų galią. Todėl tolimesniuose skyriuose siūlomi procedūrų pakeitimai, kurių savybės 5 skyriuje lyginamos su tradicinėmis.

### 3. Kointegratumo tikrinimo modifikacija, kai kointegravimo parametrai žinomi

Šiame skyriuje analizė remiasi prielaida, kad galioja (1) ryšys, t.y., parinkus tinkamą kointegruotų kintamųjų vėlavimų struktūrą, jų tiesinė kombinacija yra baltasis triukšmas. Prieš pateikiant siūlomas modifikacijas trumpai aptariami standartiniai Dickey ir Fuller (DF), išplėstas Dickey–Fuller (ADF) [4] bei Kwiatkowski ir kt. (KPSS) [5] testai, nurodant atitinkamas vėlavimų sąlygojamas problemas.

Tuo atveju, kai parametrai žinomi, Engle ir Granger siūlo kointegratumą tikrinti iš karto nagrinėjant liekanos  $v_t$  savybes. DF ir ADF testams konstruoti, stebėjimai aprašomi tokiomis lygtimis

$$v_t = \rho v_{t-1} + w_t, \quad (3)$$

$$v_t = \rho v_{t-1} + \sum_{i=1}^m \gamma_i \Delta v_{t-i} + \eta_t, \quad (4)$$

kur paklaidos  $w_t$  bei  $\eta_t$  yra laikomos silpnu baltuoju triukšmu, ir tikrinama nulinė hipotezė apie  $v_t$  integruotumą:  $H_0: \rho = 1$ . Išplėsta ADF lygtis (4) taikoma tada, kai  $w_t$  yra autokoreliuotas. Šią autokoreliaciją siekiama pašalinti išplečiant (3) lygtį pokyčių nariais  $\Delta v_{t-i}$ ,  $i = 1, \dots, m$ , kur  $m$  parenkamas vadovaujantis standartiniais informaciniais kriterijais, pavyzdžiui, AIC.

Kitaip nei DF ar ADF testai, KPSS testas, tikrina nulinę hipotezę, kad liekanos yra stacionarios. KPSS testui konstruoti, stebėjimai yra aprašomi (5) lygčių sistema

$$\begin{aligned} v_t &= \zeta_t + w_t \\ \zeta_t &= \zeta_{t-1} + \eta_t, \end{aligned} \quad (5)$$

ir tikrinama  $H_0: \sigma_\eta = 0$ .

Nagrinėjant (3) matyti, kad kointegratumo išvadai svarbiausias veiksnys yra  $\rho$  dydis. Kuo kointegruotų kintamųjų atveju  $\rho$  yra arčiau vieneto, tuo daugiau stebėjimų reikės tikslesniam parametro įverčiui ir atitinkamai korektiškai išvadai gauti. Kadangi, kaip buvo aptarta 2 skyriuje, vėlavimų ignoravimas artina  $\rho$  prie vieneto reikšmės, tai testo galia turėtų mažėti augant  $k$  ir gerėti, didėjant  $T$ . Asimptotiškai išvada apie kointegratumą būtų gauta korektiška, kadangi, prie  $X_t$  ir  $Y_t$  kointegratumo, MKM  $\hat{\rho}$  yra suderintas. Panašias išvadas galima daryti ir nagrinėjant KPSS testo konstrukciją, – ji atitinka (2) lygtyje pateiktą išdėstymą:  $\zeta_t$  (5) lygtyje atitinka (2) lygties sumą  $\beta \sum_{i=0}^{k-1} \Delta X_{t-i}$ , o  $w_t$  atitinka  $\varepsilon_t$ . Dėl to augant santykiui  $\frac{k}{T}$ , išvadų korektiškumas turėtų mažėti.

Abiem nagrinėtais atvejais (2) pateiktas išskaidymas tiesiogiai siūlo testų regresijų koregavimo būdą. KPSS testo atveju į pirmąją (5) lygčių sistemos lygtį tikslinga įtraukti narį  $\beta \sum_{i=0}^{k-1} \Delta X_{t-i}$ , kuris neleistų  $\zeta_t$  procesui priskirti į integruotą

procesą panašios dinamikos dalies, atsiradusios dėl nekorektiško vėlavimo parinkimo. DF ir ADF testų atveju pastebėtina, jog korektiškai specifikavus vėlavimų struktūrą, vienetinė šaknis būtų tikrinama procesui  $\varepsilon_t$ . Atitinkamai korektiška lygtis turėtų būti ne (3), o

$$\varepsilon_t = \rho\varepsilon_{t-1} + w_t. \quad (6)$$

Iš (2) lygties išsireiškus  $\varepsilon_t$  ir įstačius į (6) bei sutvarkius lygtį gaunama modifikuota DF regresija

$$v_t = \rho v_{t-1} - \beta \Delta_k X_t + \rho\beta \Delta_k X_{t-1} + w_t, \quad (7)$$

kur  $\Delta_k X_t = X_t - X_{t-k}$ . Jei  $k = 0$ , tai (7) redukuotųsi į DF regresiją (3). Kadangi  $k$  nėra žinomas, tai ji analogiškai kaip ir pokyčių narius ADF lygtyje (4), galima parinkti pagal AIC ar kitą informacinę kriteriją. Be to, ekonominiai samprotavimai dažnai leidžia iš anksto apriboti maksimalų galimą  $k$ . Pavyzdžiui, nagrinėjant ketvirtinius duomenis, dėl sprendimų periodo yra tikėtinas metų vėlavimas, todėl dažniausiai  $k = 4$ . Prie nulinės hipotezės, kad kintamieji yra nekointegruoti, parametro  $\rho$  įverčio  $t$  statistikos skirstinys asimptotiškai išlieka kaip ir DF, o, kaip rodo 5 skyriuje pateikti MC modelavimo rezultatai, ir mažose imtyse penkių procentų nominalaus reikšmingumo lygmens kritinės reikšmės nedaug skiriasi nuo asimptotinių.

#### 4. Kointegravimo tikrinimo korekcija, kai kointegravimo parametrai įvertinami

Šiame skyriuje, kaip ir 3, toliau laikoma, kad, parinkus tam tikrą vėlavimų struktūrą, kointegruotų kintamųjų tiesinė kombinacija yra baltasis triukšmas. Jei kointegravimo parametrai nežinomi, 3 skyriuje pateiktos korekcijos nepakanka, kadangi vertinant parametrus mažose imtyse parametrų įverčiai – čia nagrinėjamoje EG dviejų žingsnių procedūroje įprasta taikyti MKM – yra labai paslinkti, jei liekanos autokoreliuotos. Kadangi vėlavimų ignoravimas kaip tik tai sąlygoja, tikslinga modifikuoti jau pirmą EG procedūros žingsnį. Pasinaudojant (2) lygtyje pateiktu išskaidymu, vietoje vienalaikių dydžių statinės lygių regresijos tikslinga taikyti dinaminę (1) ir (2) lygtis atitinkančią regresijos funkciją

$$Y_t = \beta X_t + \sum_{i=0}^m \delta_i \Delta X_{t-i} + \varepsilon_t, \quad (8)$$

čia  $m$  žymi maksimalų galimą vėlavimą, kurį tikslinga apriboti vadovaujantis ekonomikos teorija.  $\delta_i = -\beta 1\{i < k\}$ , jei (1) lygtimi apibrėžtas kointegravimo ryšys galioja. Pastebėtina, kad tokia parametrinė korekcija turi didelį privalumą prieš daugelį kitų neparimetrinių korekcijų, kadangi, esant  $\Delta X_{t-i}$  ir  $\varepsilon_t$  baltojo triukšmo nekoreliuotiems procesams, hipotezes apie  $\delta_i$  galima tikrinti taikant standartinius  $t$  ir  $F$  testus. Kai kuri nors iš minėtų prielaidų pažeidžiama, nedidelė standartinės  $F$  testo statistikos korekcija užtikrina asimptotinę testo statistikos pasiskirstymą pagal  $F$  skirstinį (žr., pavyzdžiui, [6]).

Pirmu EG žingsniu (8) lygtyje įvertinus paklaidą  $\varepsilon_t$ , antrame procedūros žingsnyje tada galima taikyti standartinius testus jos integruotumo savybėm tirti.

### 5. Imitacinio modeliavimo rezultatai

Siekiant įvertinti vėlavimų struktūros ignoravimo kointegravimo ryšiuose problemiškumą ir pasiūlytos korekcijos prasmingumą, atliktas 500 iteracijų Monte Carlo eksperimentas.

Naudotas duomenis generuojantis procesas:

$$Y_t = \beta X_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim n.v.p.N(0, 1); \quad \beta = 0, 1; \quad k = 1, \dots, 4; \quad (9)$$

$$\Delta X_t = \xi_t, \quad \xi_t \sim n.v.p.N(0, 1); \quad E(\xi_t \varepsilon_t) = 0. \quad (10)$$

Pagal generuotus duomenis tirti standartinės ir modifikuotos EG procedūrų rezultatai bendresniam atvejui, t.y., kai nėra žinomi parametru įverčiai (modifikacijų palyginimas su standartinių procedūrų rezultatais kai parametru įverčiai yra žinomi darbe nepateikiami dėl apimties apribojimų, tačiau jų rezultatai yra geresni nei toliau apibūdinamam atvejui, kai parametru reikšmės nežinomos; taip pat neaptariami ir KPSS testo rezultatai dėl prastesnių šio testo savybių tikrinant kointegratumą mažose imtyse). Tuo tikslu pirmiausiai įvertinta standartinė vienalaikė EG pirmo žingsnio regresija

$$Y_t = \beta X_t + v_t, \quad (11)$$

bei jos koreguota lygtis (8), kurioje  $m$  iš anksto apribotas 5, o optimalus išplėtimas parinktas pagal AIC kriterijų. Antrame žingsnyje nulinė hipotezė, ar įvertintų regresijų liekanos turi vienetinę šaknį ( $H_0: \rho = 1$ ), tikrinta taikant standartinius DF ir ADF testus, naudojant (3) ir (4) lygtis, kur ADF testo išplėtimas parinktas pagal AIC kriterijų.

1–2 lentelėse pateikti EG procedūroje pagal DF ir ADF testų  $t$  statistikos empirinius skirstinius įvertinti testų galios prieš kointegruotų kintamųjų alternatyvą rezultatai, vertinant pagal penkių procentų asimptotinių nominalųjų reikšmingumo lygmenį. Stulpeliuose DF ir ADF nurodyta standartinės EG procedūros galia. Modifikuotos procedūros rezultatai pateikti atitinkamai MDF ir MADF stulpeliuose.

Nagrinėjant modeliavimo rezultatus matyti, jog, esant nedideliam realizacijos dydžiui  $T$ , augantis vėlavimas  $k$  vis labiau jaučiasi ir, prie  $T = 30$ , standartinės EG procedūros taikant DF ir ADF testus galia nukrenta beveik per pusę. Tuo tarpu taikant modifikuotą procedūrą testo galia tiek MDF, tiek MADF atvejais beveik neįtakojama

1 lentelė. Kointegravimo testų galia 30 stebėjimų dydžio realizacijose (MC imitacijos rezultatai)

$T = 30$	Kintamųjų poveikio lagas	DF	MDF	ADF	MADF
$\beta = 1$	$k = 0$	0,998	0,997	0,969	0,97
	$k = 1$	0,999	0,998	0,957	0,952
	$k = 2$	0,943	0,995	0,908	0,946
	$k = 3$	0,723	0,994	0,689	0,951
	$k = 4$	0,584	0,989	0,539	0,927
$\beta = 0$	$k = 0$	0,046	0,061	0,07	0,085

2 lentelė. Kointegravimo testų galia 50 stebėjimų dydžio realizacijose (MC imitacijos rezultatai)

$T = 50$	Kintamųjų poveikio lagas	DF	MDF	ADF	MADF
$\beta = 1$	$k = 0$	1	1	0,988	0,989
	$k = 1$	1	1	0,985	0,99
	$k = 2$	0,993	1	0,971	0,984
	$k = 3$	0,901	1	0,87	0,985
	$k = 4$	0,805	1	0,742	0,989
$\beta = 0$	$k = 0$	0,048	0,053	0,056	0,059

ir nominalaus reikšmingumo lygmuo taip pat beveik neiškreipiamas. Antra vertus, augant stebėjimų skaičiui  $T$ , vėlavimų įtaka testo galiai gana greitai mažėja, ką galima paaiškinti MKM įverčių supersuderinamumu kointegravimo regresijoje. Bendrai matyti, kad ADF bei MADF testai nors ir nedaug, bet visą laiką turi mažesnę galią nei jų DF ir MDF analogai.

## 6. Išvados

Taikant standartinę Engle–Granger procedūrą kintamųjų kointegruotumui tikrinti, vėlavimų ignoravimas kointegravimo ryšyje gali sąlygoti žymų kointegruotumo testo galios praradimą. Ši išvada galioja tikrinimui panaudojant tiek integruotumui testuoti skirtų Dickey–Fuller, tiek stacionarumo nulinei hipotezei tikrinti skirtų KPSS testų atvejais.

Paprasta parametrinė standartinės Engle–Granger kointegruotumo testo korekcija leidžia žymiai padidinti testų galią, kuri prarandama dėl vėlavimų struktūros ignoravimo. Tokio paties principo korekcija galima tiek tuo atveju, kai kointegravimo parametrų vektorius žinomas, tiek tuo atveju, kai jis yra įvertinamas. Ši korekcija turi privalumą prieš daugelį kitų neparimetrinių korekcijų, nes ji turi standartinius baigtinių imčių skirstinius, tikrinti parametrų reikšmingumo hipotezes, prie gana dažnai praktikoje pasitaikančių sąlygų.

## Literatūra

1. R.F. Engle, C.W.J. Granger, Co-integration and error correction: representation, estimation and testing, *Econometrica*, **55**, 251–276 (1987).
2. A. Banerjee, J.J. Dolado, D.F. Hendry, G.W. Smith, Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence, *Oxford Bulletin of Economics and Statistics*, **48**, 253–278 (1986).
3. V. Kvedaras, *Integracijos į Europos Sąjungą pasekmių Lietuvos ūkiui makroekonometrinis modeliavimas*, disertacija, VU (2004).
4. D.A. Dickey, W.A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–473 (1979).
5. D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics*, **54**, 159–178 (1992).
6. J.D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton (1994).

## SUMMARY

***V. Kvedaras. Some effects of lagged relationship on the cointegration inferences in small samples***

A Monte Carlo simulation is performed in order to investigate the effects of lagged relationship on the cointegration inference in a single equation. Given a small data sample the standard application of Engle–Granger cointegration testing procedure is significantly affected by the presence of lagged relationship. For instance, in a sample size of 30 observations, the power of the two-step Engle–Granger cointegration testing procedure, using the Dickey–Fuller (DF) or Augmented DF (ADF) test statistic in the second step, drops from almost one hundred percent, when the correct lag structure of cointegration relationship is respected, to around sixty percent, when the effect of 4 lags is ignored.

A simple parametric correction is proposed allowing avoiding the negative influence. When the cointegration parameters are known, the correction is applied directly to DF and ADF regressions. Whenever the parameters are estimated in the first step of Engle–Granger procedure, the cointegration regression should be modified instead in order to avoid the autocorrelation caused bias of parameter estimates. A Monte Carlo simulation reveals that such simple correction retains the power of the cointegration testing procedure without having a negative effect on the nominal size.

*Keywords:* cointegration tests, lagged relationship, Monte Carlo simulation, Engle–Granger procedure.