

EGYVÁLTOZÓS IDŐSORMODELLEKEN ALAPULÓ INFLÁCIÓS ELŐREJELZÉSEK

LIELI RÓBERT

Tanulmányomban a magyarországi inflációs folyamat egyváltozós idősormodellekkel történő leírására és előrejelzésére vállalkozom. Az inflációs rátát a fogyasztói árindex (Consumer Price Index – CPI) havi növekedési ütemével mérem. A modellek az 1994 januárjától 1999 áprilisáig terjedő időszakra vonatkoznak, az előrejelzések 1999 végéig készültek.

A fogyasztói árindexet technikailag kétféle módon lehet előrejelezni. Az első lehetőség az, hogy közvetlenül a fogyasztói árindex idősorára illesztünk modellt. A másik megoldás, hogy külön-külön modellezzük a fogyasztói árindex fő összetevőit, és az egyes részürejelzésekből súlyozott átlagot képzünk azokkal a súlyokkal, melyekkel az egyes összetevők a fogyasztói árindexben szerepelnek. A Központi Statisztikai Hivatal (KSH) által havi rendszerességgel közzétett részárindexek¹ a következő termékcsoportokat foglalják magukban (a zárójelben a továbbiakban használandó megnevezés és a fogyasztói árindexben betöltött százalékos súly szerepel):

1. élelmiszerek (ÉLELM – 27,2),
2. szeszes italok, dohányárak (SZESZ – 8,9),
3. ruházati cikkek (RUHA – 6,1),
4. tartós fogyasztási cikkek (TARTÓS – 5,5),
5. háztartási energia (ENERG – 8,9),
6. egyéb cikkek (EGYÉB – 17,0),
7. szolgáltatások (SZOLG – 26,4),
8. a teljes fogyasztói árindex (CPI – 100).

Az alkalmazott statisztikai modellek az ún. SARIMA modellcsalád tagjai, melyek tulajdonképpen az egyszerű autoregresszív mozgóátlagolású (ARMA) modellek kiterjesztései szezonaritást és nem stacionárius viselkedést mutató idősorokra. A felhasznált idősorok tehát szezonálisan nem kiigazítottak: a szezonaritást a modelleken belül kezeltem.² A SARIMA-modellek nem adnak közgazdasági magyarázatot az inflációs folyamatra:

¹ Az árindexeket a KSH változó bázissal, „előző hó=100,0” formában közli. Számításaim során tizedes törtté alakítottam ezeket az értékeket, de a végeredményeket én is hasonló formában adom meg.

² A tradicionális szemléletű idősorelemzés a szezonaritást olyan zavaró tényezőnek tekinti, melytől a modellezés előtt a vizsgált idősorot meg kell tisztítani (különösen, ha a hosszú távú tendenciák vizsgálata a cél). A szezonális kiigazítás általánosan alkalmazott módszerei azonban meglehetősen ad hoc természetűek, és újabban éles kritika célpontjaivá váltak. (Lásd [2], [3].)

egyszerűen a múltbeli értékekből extrapolálják a jövőt, azt feltételezve, hogy a bizonyos múltbeli korrelációk a jövőben is jelen lesznek.

A modellek felállításának első lépéseként találni kell egy olyan transzformációt, mely a vizsgált idősort stacionáriussá teszi. A *G. E. P. Box* és *G. M. Jenkins* nevével fémjelzett hagyományos metodológia [1] szezonalitást mutató idősorok esetében általában szezonális differencia³ képzését ajánlja e cél érdekében. Ez a látszólag egyszerű transzformáció valójában rendkívül összetett, használata olyan implicit feltevések elfogadásával jár, melyek a priori alapokon csak ritkán védhetők. Havi idősorok esetében a szezonális differenciálás ugyanis összesen 12 (egy nem szezonális és 11 szezonális) egységgyököt feltételez az eredeti idősorban, és ha ténylegesen ennél kevesebb egységgyök van jelen, akkor a szezonális differenciálásnál egyszerűbb, abba „beágyazott” transzformáció is elég a stacionaritás eléréséhez. A szezonális differencia képzése ebben az esetben rosszul modellezhető, „túldifferenciált” idősorhoz vezet.

Az *S. Hylleberg* és szerzőtársai [7] által kidolgozott módszertan explicit szezonális egységgyöktesztet javasol annak megállapítására, hogy a vizsgált idősor pontosan milyen egységgyököket tartalmaz. A próba eredményeinek ismeretében „testre szabott” stacionárius transzformáció található az idősorhoz, melynek alkalmazásával a túldifferenciálás veszélye elkerülhető. (Az ez utóbbi módszertannal felállított modellekre a „HEGY” betűszóval fogok hivatkozni.)

A modellezés során gyakorta felmerülő probléma, hogy sok ténylegesen megfigyelt idősorban – az általam tanulmányozott inflációs idősorokban is – időnként élesen kiugró, az idősor addigi és azt követő alakulásába nem illő megfigyelések jelennek meg. Ezek kezelése külön módszertant igényel, hiszen feltehető, hogy az idősormodell szempontjából külső, vissza nem térő eseményekről van szó, melyek modellezése éppen ezért nem mindig célszerű.

A különböző elméleti megfontolások alapján végül öt inflációs előrejelzést készítettem. Az első két előrejelzés közvetlenül a fogyasztói árindexre illesztett Box–Jenkins-, illetve HEGY-modellből származik; három pedig a hét CPI-komponens különböző előrejelzéseinek súlyozott átlaga. Az első súlyozott átlag a részidősorok Box–Jenkins-előrejelzéseit veszi alapul, míg a második a HEGY-modellekét. A harmadik („vegyes”) súlyozott átlag képzésénél a részaggregátumokat legjobban leíró modelleket vettem figyelembe, függetlenül attól, hogy melyik modell típus alkalmazásával készültek.

Az előrejelző modellek jóságának megállapításához számos vizsgálat nyújthat támpontot. Elsőként dinamikus, illetve statikus mintán kívüli előrejelzéseket készítettem az 1998. január és december közötti időszakra, és összevettem egymással a különböző előrejelzések hibastatisztikáit. Ezáltal képet lehet alkotni az egyes módszerek relatív előrejelzési teljesítményéről. Az előrejelzési pontosság abszolút mércéjeként két naiv előrejelző módszer hibastatisztikáit is bevontam az összehasonlításba. A felállított idősormodellekkel szemben jogos elvárás, hogy a naiv módszereknél jobb előrejelzést nyújtsanak.

Az idősormodellek előrejelzései abból a szempontból is értékelhetők, hogy az általuk jósolt értékek miként viszonyulnak a piaci szereplők előrejelzéseéhez. A különböző pro-

³ Egy havi rendszerességgel megfigyelt y_t idősor szezonális differenciáján az $y_t - y_{t-12}$ különbséget értjük. Az inflációs idősorok esetében ez a mennyiség a havi infláció változását adja meg az előző év azonos hónapjához képest.

fitorientált szervezeteknek (bankok, brókerházak, kutatócégek stb.) elvileg – a piac logikája szerint – optimális inflációs előrejelzésekkel kell rendelkezniük. Az optimális kifejezés úgy értendő, hogy az előrejelzések további pontosításához szükséges ráfordítások már meghaladnák a nagyobb pontosságból eredő többlethasznot. Éppen ezért különösen érdekes kérdés, hogy a viszonylag kis ráfordítással felállítható SARIMA-modellek előrejelzései hogyan viszonyulnak a piaci szereplőkéhez. Az összehasonlítást különböző időtávú előrejelzésekre vonatkozóan végeztem el.

A SARIMA-MODELLEK

Az egyváltozós lineáris SARIMA-modelleket tagadhatatlan egyszerűségük ellenére (vagy éppen emiatt) meglehetősen gyakran alkalmazzák különböző gazdasági idősorok leírására és előrejelzésére. A tapasztalat azt mutatja, hogy ezek az egyszerű sztochasztikus modellek is képesek megbízható előrejelzéseket adni legalábbis, ha a vizsgált időtartományban nincsenek jelentős strukturális változások, és a linearitási feltétel nem túl korlátozó leírása az adatszolgáltató folyamatnak.⁴

Az y_t sztochasztikus folyamat a definíció szerint SARIMA $(P,D,Q)^s(p,d,q)$ folyamatot követ, ha

$$a_s(L^s)a(L)(1-L^s)^D(1-L)^d y_t = c + b_s(L)b(L)\varepsilon_t, \quad /1/$$

ahol ε_t konstans varianciájú fehér zaj, s a szezonális frekvenciája (például $s=12$ havi adatok esetén), L a késleltetési operátor,⁵ és

$$a_s(L^s) = 1 - a_{1s}L^s - a_{2s}L^{2s} - \dots - a_{Ps}L^{Ps}, \quad /2/$$

$$a(L) = 1 - a_1L - a_2L^2 - \dots - a_pL^p, \quad /3/$$

$$b_s(L^s) = 1 + b_{1s}L^s + b_{2s}L^{2s} + \dots + b_{Qs}L^{Qs}, \quad /4/$$

$$b(L) = 1 + b_1L + b_2L^2 + \dots + b_qL^q \quad /5/$$

a késleltetési operátor polinomjai, és az $a_s(L^s)a(L)=0$ egyenlet gyökei az egységkörön kívül esnek. Ez a feltétel azt jelenti, hogy az $(1-L)^d(1-L^s)^D$ szűrő már minden egységgyektől megtisztította a folyamatot, vagyis a megfelelően differenciált y idősor már stacionárius. E szűrő bármely olyan transzformációval helyettesíthető, mely stacionárius idősort eredményez. A modell eme általános formája alkalmas olyan idősorok leírására, melyek szezonalitást és nem stacionárius viselkedést mutatnak: két tulajdonság, mely sajátja az inflációs idősoroknak.

A SARIMA-modellek /1/-es specifikációjában a modell autoregresszív része két polinom – egy szezonális és egy nem szezonális – szorzataként áll elő. Az autoregresszív részt azonban szabadon, egyetlen $Ps+p$ rendű polinom formájában is fel lehetne írni; a

⁴ Az ökonometriai szakirodalom régóta hangoztatja, hogy – legalábbis rövid távon – az ARMA-modellek gyakorta jobb előrejelzéseket képesek adni, mint az összetett strukturális makromodellek. (Az előrejelzésről szóló klasszikus irodalom jó áttekintését adja [8] 17. fejezete.)

⁵ A késleltetési operátor definíciója: $L^k y_t = y_{t-k}$.

szorzat alakban való kifejezés tulajdonképpen korlátozásokat jelent a szabadon felírt autoregresszív polinom együtthatóira nézve. Természetesen elképzelhető, hogy az adatok nem támasztják alá ennek a korlátozásnak a jogosságát: a modellek felállítása során erre a lehetőségre is tekintettel kell lenni. (Ez a megfontolás nyilván érvényes a mozgóátlagolású polinomra is.)

Az eddigi általános és meglehetősen formális leírás után érdemes konkrét példán megvizsgálni a SARIMA-modellek struktúráját. A továbbiakban részletezendő identifikációs vizsgálatok szerint az élelmiszerek havi inflációs rátája jól közelíthető a következő SARIMA (1,1,0)^s (0,0,1) modellel:

$$\Delta_{12}y_t = -0,0025 - 0,50\Delta_{12}y_{t-1} + 0,44\varepsilon_{t-1} + \varepsilon_t, \quad /6/$$

ahol y_t jelöli az élelmiszerek havi inflációs rátáját, Δ_{12} a szezonális differencia képzését⁶ és ε_t az idősort érő véletlen sokkokat. Az egyenlet tehát úgy írja fel a havi infláció változásának jelen időpontbeli értékét, mint az előző havi érték és sokk, valamint egy hibatag – a jelenbeli sokk – lineáris függvényét. (A múltbeli értékeket autoregresszív, a múltbeli sokkokat pedig mozgóátlagolású tagnak hívják.)

A modell alapján az élelmiszerek árufőcsoport inflációs dinamikáját három lényeges tényező vezérli.

– A negatív konstans azt jelenti, hogy az infláció hosszú távon csökkenő tendenciát mutat: jelenleg létezik az inflációnak egy átlagos (várható) csökkenési üteme, mely azt mutatja meg, hogy az infláció adott hónapban átlagosan (várhatóan) mennyivel lesz kisebb, mint az előző év ugyanazon hónapjában. (A várható csökkenési ütem pontos nagyságát az autoregresszív tag együtthatója és a konstans tag együtt határozza meg.) Az inflációnak tehát van egy tehetetlenségi komponense, mely jelenleg a csökkenés irányába hat.

– Az a tény, hogy az autoregresszív tag együtthatója abszolút értékben egynél kisebb, biztosítja, hogy a havi infláció 12 hónappal ezelőtti értékéhez viszonyított változása valóban a várható érték körül ingadozzon, és hosszú távon konvergáljon hozzá. A negatív előjel figyelembevételével ez azt jelenti, hogy ha az infláció tényleges csökkenése az adott hónapban meghaladja a csökkenés átlagos mértékét, akkor – amennyiben nem éri az inflációt előre nem látható sokk – a 12 hónap múlva mérhető havi infláció valamivel nagyobb lesz, mint az az érték, amit egyedül az átlagos csökkenési ütem figyelembevételével kapnánk.

– Az inflációnak van egy „meglepetés” komponense, mely nem jelezhető az infláció múltbeli alakulásából. A mozgóátlagolású tag azt írja le, hogy ezek a sokkok (vagy ún. innovációk) miként épülnek be az inflációs folyamatba. Tegyük fel, hogy adott hónapban a havi inflációs ráta egy százalékponttal haladja meg az előrejelzett értékét, például az infláció csökkenése egy százalékponttal kisebb a vártnál (pozitív innováció). Az elsőrendű mozgóátlagolású tag jelenléte és a 0 és 1 közé eső együttható azt jelenti, hogy a következő hónapban az infláció értéke nulla egész valahány százalékponttal nagyobb lesz annál, mint amit kizárólag az előző két dinamikus tényező alapján feltételezni lehetne. Vagyis az infláció tehetetlenségének van egy másik dimenziója is: az ebben a hónapban jelentkező véletlen sokkok a következő hónap inflációját is közvetlenül befolyásolják, időre van szükség a teljes alkalmazkodáshoz. Ha nincs másodrendű mozgóátlagolású tag, és nincs újabb meglepetés (az innovációk értéke nulla), akkor adott sokk a felmerülése utáni második hónapban már nem gyakorol közvetlen hatást az inflációs rátára: a dinamikát az első két tényező határozza meg, a sokk hatása elenyészik, és az infláció csökkenése ismét visszatér az átlagos értékéhez.

A SARIMA-modellekből származó inflációs előrejelzések tulajdonképpen a jelzett dinamikus tényezők alkalmazásával készülnek, azon feltételezés mellett, hogy az innovációk értéke az előrejelzési periódusban nulla.

⁶ Vagyis $\Delta_{12}y_t = (1 - L^{12})y_t = y_t - y_{t-12}$.

IDENTIFIKÁCIÓS VIZSGÁLATOK

A SARIMA-modellek identifikálásakor⁷ sok tekintetben a hagyományos Box–Jenkins-metodológia előírásait követtem. Az eljárás egyik alapelve a mértékletes parametrizációra való törekvés: a $/2/-/5/$ polinomok fokszáma a gyakorlatban rendszerint igen alacsony (nulla is lehet). További fontos alapelv a felállított modell reziduumaival szemben támasztott azon követelés, hogy azok egy fehér zaj folyamat realizációinak legyenek tekinthetők.

Az identifikáció már említett alapvető kérdése, hogy a vizsgált idősor konzisztens-e a stacionaritási feltevessel, és ha nem, akkor milyen transzformáció teszi azzá. A SARIMA-modellek $/1/-$ es specifikációjában az $(1-L)^d(1-L^s)^D$ alakú transzformáció a hagyományos Box–Jenkins-módszertan ajánlását tükrözi, gyakorlati munkákban legtöbbször ilyen alakú transzformációt alkalmaznak a stacionaritás elérése érdekében.

A HEGY-módszer abból indul ki, hogy a $D \neq 0$ esetben az $(1-L)^d(1-L^s)^D$ alakú transzformáció rendkívül összetetté válik: havi adatok esetén például a folyamatban legalább 12 egységgyök jelenlétét feltételezi. A HEGY-módszer ezért előbb multiplikatív tényezőkre bontja a szűrőt, majd explicit egységgyöktesztekkel vizsgálja, hogy mely tényezők használatára van ténylegesen szükség.

Már említett probléma, hogy a vizsgált inflációs idősorokban vannak olyan kiugró megfigyelések, melyek kezelése külön módszertani megfontolásokat igényel. Ezekkel később bővebben foglalkozom. Itt elég annyit megjegyezni, hogy a SZESZ és az ENERGI változók esetében a identifikációs vizsgálatokat már a rendellenes megfigyelésektől megtisztított idősoron végeztem el.

A stacionárius transzformáció a hagyományos Box–Jenkins-metodológia tükrében

A Box–Jenkins-metodológia keretei között gondolkodva, az identifikáció első lépése a d és a D paraméterek értékének megválasztását jelenti az $(1-L)^d(1-L^{12})^D$ transzformációban. A tapasztalat azt mutatja, hogy ezek a paraméterek a 0 vagy 1 értéket veszik fel, azaz a stacionaritás általában elérhető az 1, $(1-L)$, $(1-L^{12})$ vagy $(1-L)(1-L^{12})$ transzformációk valamelyikének alkalmazásával.⁸ E lehetőségek közti választás azonban korántsem könnyű: nemegyszer egymásnak ellentmondó gondolatmeneteket és bizonyítékokat kell mérlegre tenni.

A kiterjesztett Dickey–Fuller-féle ADF-teszt viselkedése rögtön felhívta a figyelmet az identifikációval kapcsolatos problémákra. A teszt konklúziója nagyon érzékenyen függött a beiktatott késleltetett differenciák számától.⁹ Alacsony késleltetésszámoknál a próba egyértelműen elutasította az egységgyök nullhipotézisét (a konstans igen, de trendet nem tartalmazó alternatívával szemben). Nagy számú (például 11–12 vagy akár 23)

⁷ Az identifikálás ebben az összefüggésben a p , d , q , illetve P , D , Q paraméterek meghatározását jelenti. Tehát itt nem az ökonometria klasszikus identifikációs problémájáról van szó, amikor egy strukturális modell paramétereit kell visszanyerni a modell redukált (becsülhető) formájának megbecslése után.

⁸ Az 1 szimbólum az identikus transzformációt, vagyis azt a lehetőséget jelenti, hogy a vizsgált idősor már eleve stacionárius.

⁹ A tesztegyenleteket konstans beiktatásával, de determinisztikus trend nélkül becsültem. A kritikus értékeket például [6] 763. old.) tartalmazza. Az ADF-teszt érzékenysége a késleltetéshossz megválasztására ismert probléma. (Lásd [4].)

késleltetett differencia alkalmazása a legtöbb idősor esetében azonban pontosan ellentétes következtetéshez vezetett, amint azt az 1. tábla is tanúsítja.¹⁰

1. tábla

Az ADF t-teszt eredményei az 1992. január–1998. december mintában

Termékcsoport	A tesztgyenletben alkalmazott késleltetett differenciák száma			
	$p=3$	$p=11$	$p=12$	$p=18$
ÉLELM	-3,68***	-1,54	-1,78	-2,52
AC/SC	-8,041/-7,896	-8,103/-7,728	-8,103/-7,697	-8,065/-7,486
SZESZ	-3,76***	-1,97	-1,89	-1,21
AC/SC	-10,276/-10,132	-10,157/-9,780	-10,133/-9,278	-10,031/-9,452
RUHA	-9,59***	-2,46	-2,41	-3,51**
AC/SC	-9,060/-8,915	-10,060/-9,684	-10,040/-9,635	-10,043/-9,664
TARTÓS	-3,28**	-2,35	-2,30	-2,14
AC/SC	-10,305/-10,161	-10,305/-9,929	-10,307/-9,902	-10,267/-9,688
ENERG	-4,69***	-3,46**	-3,42**	-2,59*
AC/SC	-6,860/-6,716	-6,696/-6,320	-6,675/-6,270	-6,676/-6,098
EGYÉB	-3,09**	-0,74	-0,81	-1,10
AC/SC	-8,751/-8,606	-9,000/-8,623	-8,987/-8,582	-8,926/-8,348
SZOLG	-4,77***	-3,37**	-3,02**	-2,91**
AC/SC	-8,502/-8,357	-8,755/-8,379	-8,788/-8,383	-8,720/-8,142
CPI	-3,40**	-0,78	-1,03	-1,21
AC/SC	-8,983/-8,838	-9,311/-8,935	-9,328/-8,922	-9,221/-8,642

* 10 százalékos szignifikancia;

** 5 százalékos szignifikancia;

*** 1 százalékos szignifikancia.

AC – az Akaike-féle információs kritérium;

SC – a Schwartz-kritérium.

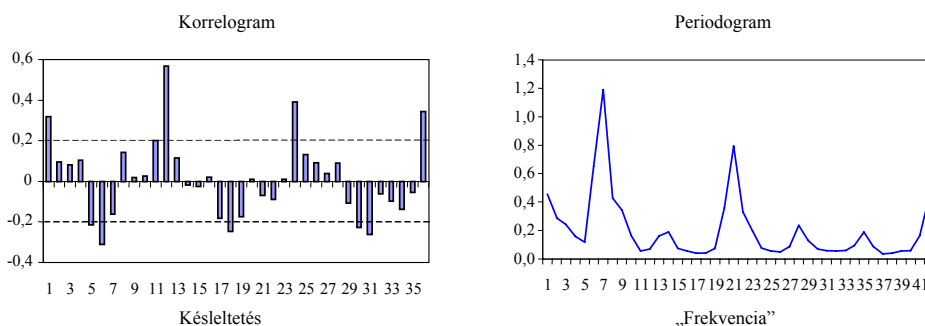
Az 1. tábla tartalmazza a becsült egyenletek Akaike-, illetve Schwartz-féle információs kritériumát (AC/SC) is, melyek segítséget nyújtanak az optimális késleltetéshossz megállapításához (ezeket minimalizáló specifikációként szokás elfogadni). A SZESZ és az ENERG változók esetét kivéve e statisztikák közül legalább az egyik 11-12 késleltetett differencia beiktatását javasolja a tesztgyenletekbe. Ezen késleltetéshosszak mellett az ÉLELM, a RUHA, a TARTÓS, az EGYÉB és a CPI idősorokban még 10 százalékon sem lehet elvetni az egységgyök nullhipotézisét. A kivételnek tetsző SZOLG változó további vizsgálata azt mutatja, hogy 1992 áprilisától (vagy később) induló mintákban az egységgyök elvetését – a magasabb késleltetésszámok mellett – már semmi sem támasztja alá.

Az ADF-teszt viselkedése (az a tény, hogy alacsony késleltetésszámoknál elveti, de 11-12 késleltetett differencia alkalmazása esetén már nem képes elvetni az egységgyök nullhipotézisét) úgy is értelmezhető, hogy az egymástól körülbelül egy évnnyire elhelyez-

¹⁰ Első pillantásra túlzásnak tűnhet 11-12, sőt 18 késleltetett differenciát alkalmazni az ADF-teszt regressziós egyenletében. Gondoljuk azonban meg, hogy havi adatok esetén például egy SARIMA (1,1,0)^s (0,0,1) folyamat is tulajdonképpen 24-ed rendű autoregresszió, hiszen $(1 - \alpha L^2)(1 - L^{12})y_t = (1 + \beta L)\varepsilon_t \Leftrightarrow y_t = (\alpha + 1)y_{t-12} - \alpha y_{t-24} + \beta \varepsilon_{t-1} + \varepsilon_t$. Mindazonáltal az ADF-teszt eredményeivel óvatosan kell bánni. A regresszióban magyarázó változóként szereplő sok késleltetett differencia például szinte elkerülhetlenül multikollinearitást indukál, rontva a becslés minőségét. (A legtöbb becsült koefficiens valószínűleg inszignifikáns lesz.)

kedő megfigyelések között különösen erős (szezónális) kapcsolat van, ami felveti az $(1-L^{12})$ szezónális szűrő alkalmazásának lehetőségét. Az idősorok korrelogramjai (vagyis empirikus autokorreláció-függvényei) szintén megerősítették ezt a hipotézist: az autokorrelációk legtöbb esetben lassan halnak ki, és erőteljesen kiugró értékek figyelhetők meg 12, 24, illetve 36 késleltetésnél. Ha az autokorrelációk ilyen sokáig szignifikánsak, az mindenképp még szezónális idősorok esetében is a stacionaritási feltétel érvénytelenségét sugallja. Az idősorok periodogramjai is jól mutatják a szezónális frekvenciák dominanciáját a vizsgált inflációs idősorok többségében.¹¹ Az 1. ábra a CPI korrelogramját és periodogramját mutatja be.

1. ábra. A CPI korrelogramja és periodogramja



Megjegyzés. Az autokorreláció-függvény az 1992. január–1999. április mintaperiódus alapján, 88 megfigyelés felhasználásával készült. Az autokorreláció-értékekre a $\pm 2/\sqrt{T}$ képlet segítségével számolható közelítőleg 95 százalékos konfidencia-intervallum (T a mintanagyság). Ezek a határok körülbelül $\pm 0,2$ -del egyenlők. A periodogram az 1992. január–1998. december mintaperiódus alapján, 84 megfigyelés felhasználásával készült ($T=84$). Az abszcissa-tengelyen az idősort felépítő ciklusok frekvenciája szerepel az $1/(84 \text{ hónap})$ egységben. Egyszerűbben fogalmazva: a j abszcissa-értékhez $T/j=84/j$ hónap periódusú ciklus tartozik. Az éves ciklusok így például a 7 abszcissa-értéknél jelennek meg.

Statistikai tesztek mellett az elméleti megfontolások is amellett szólnak, hogy a havi inflációs idősorokat valamilyen módon még differenciálni szükséges. Más különben ugyanis konstans várható értéket kényszerítenénk az inflációs idősorokra, ami annak a feltételezésnek felelne meg, hogy hosszú távon nem lehet defláció [5] (110. old.) rámutat arra, hogy ha egy egyváltozós idősormodellben a konstans várható érték feltételezése hamis, akkor a jóslott értékek megbízhatatlanok lesznek, és sokszor jobb előrejelzéseket lehet kapni, ha a kétséges esetekben a differenciálás mellett dönt a kutató. Végezetül: a differenciálás sokszor eredményez olyan autokorreláció-struktúrát, mely jóval könnyebben értelmezhető, mint az eredeti idősoré.

A Box–Jenkins-metodológia keretei között végrehajtott identifikációs vizsgálat első lépésének eredményeit a 2. tábla foglalja össze. A döntések egy része egymásnak ellentmondó bizonyítékok mérlegelése alapján született. Lehet, hogy mások eltérő módon súlyoznák a különböző vizsgálatok eredményeit, és ezért más következtetésre jutnának.

¹¹ Szezónális frekvenciákon a 2, 2.4, 3, 4, 6 és 12 hónapos periodicitású ciklusok értendők. Ezek ugyanis azok a ciklusok, melyeket a szezónális differenciálás közömbösít. (Lásd [6] 170–172. old.)

2. tábla

<i>Box-Jenkins-féle identifikációs vizsgálat</i>			
Termékcsoport	$(1-L^{12})$	$(1-L)$	Alkalmazott szűrő
ÉLELM	++	-	$(1-L^{12})$
SZESZ	+ (-)	-	$(1-L^{12})$
RUHA	++	+	$(1-L^{12})(1-L)$
TARTÓS	+	-	$(1-L^{12})$
ENERG	+ (-)	-	$(1-L^{12})$
EGYÉB	+	-	$(1-L^{12})$
SZOLG	++	-	$(1-L^{12})$
CPI	++	-	$(1-L^{12})$

Megjegyzés. A két kereszt azt jelenti, hogy a megjelölt szűrő szükségességére számottevő bizonyíték van; az egy kereszt gyengébb bizonyíték létezésére, a mínuszjel pedig bizonyíték hiányára utal.

A stacionárius transzformáció a HEGY-metodológia tükrében

Annak ellenére, hogy elég meggyőzően lehet érvelni (legalább) a szezonális differenciálás szükségessége mellett, vizsgálni kell a túldifferenciálás lehetőségét is. Az $(1-L^{12})$ szűrő alkalmazása tulajdonképpen azt feltételezi, hogy a vizsgált idősor egy nem szezonális és tizenegy szezonális egységgyökkel rendelkezik. Könnyen belátható ez a megállapítás, ha a kérdéses szűrőt elemi szűrők szorzatára bontjuk az $L^{12}-1=0$ egyenlet gyökeinek felhasználásával. Ha $\lambda_1, \lambda_2, \dots, \lambda_{12}$ jelöli ezeket a gyököket, akkor a $\prod_{j=1}^{12} (1-\lambda_j L)$ szorzat pontosan az $(1-L^{12})$ szűrőt adja vissza.

Gyök	Megfelelő elemi szűrő
$\lambda_1 = 1$	$(1-L)$
$\lambda_2 = -1$	$(1+L)$
$\lambda_{3,4} = \pm i$	$(1-iL)(1+iL) = (1+L^2)$
$\lambda_{5,6} = \frac{\sqrt{3}}{2} \pm \frac{1}{2}i$	$(1-\lambda_5 L)(1-\lambda_6 L) = (L^2 - \sqrt{3}L + 1)$
$\lambda_{7,8} = -\frac{\sqrt{3}}{2} \pm \frac{1}{2}i$	$(1-\lambda_7 L)(1-\lambda_8 L) = (L^2 + \sqrt{3}L + 1)$
$\lambda_{9,10} = \frac{1}{2} \pm \frac{\sqrt{3}}{2}i$	$(1-\lambda_9 L)(1-\lambda_{10} L) = (L^2 - L + 1)$
$\lambda_{11,12} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$	$(1-\lambda_{11} L)(1-\lambda_{12} L) = (L^2 + L + 1)$

Az $(1-L^{12})$ szűrő összetettsége miatt jogosan merül fel a kérdés, hogy vajon a vizsgált inflációs idősorok az egységgyökök mindegyikét tartalmazzák-e. Példának okáért tegyük fel, hogy az egyik idősor nem tartalmazza a $\pm i$ konjugált párt. Ebben az esetben már az $(1-L^{12})/(1+L^2)$ szűrő alkalmazásával stacionárius idősorhoz jutunk; az $(1-L^{12})$ szűrő használata túldifferenciált idősort eredményez, abban az értelemben, hogy egységgyök kerül a

folyamat $MA(\infty)$ reprezentációjába. Ez pedig bonyodalmakat okoz a modellek becslésénél és előrejelzésénél. (Egységgyökök tartalmazó MA-folyamatok ugyanis nem invertálhatók.)

A szezonális egységgyökök jelenlétének tesztelése tehát a gyakorlat számára is fontos probléma. Az [2] által bemutatott eljárás havi adatokat feltételező változata a következő segédregresszióra épül:

$$\varphi(L)y_{8,t} = \mu_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} + \pi_5 y_{4,t-1} + \pi_6 y_{4,t-2} + \pi_7 y_{5,t-1} + \pi_8 y_{5,t-2} + \pi_9 y_{6,t-1} + \pi_{10} y_{6,t-2} + \pi_{11} y_{7,t-1} + \pi_{12} y_{7,t-2} + \varepsilon_t, \quad /7/$$

ahol az $y_{1,t}, y_{2,t}, \dots, y_{8,t}$ segédidősorok a vizsgálat tárgyát képező y_t idősor meghatározott transzformáltjai,¹² és

$$\mu_t = \sum_{s=1}^{12} \delta_s D_{s,t}. \quad /8/$$

A /8/-ban $D_{s,t}$ ($s=1, \dots, 12$) olyan szezonális dummy változót jelöl, mely az 1 értéket veszi fel az év s -edik hónapjában, egyébként pedig nulla. A /7/ egyenlet OLS-sel becsülhető; a $\varphi(L)$ autoregresszív polinomot úgy kell megválasztani, hogy a reziduumok egy fehér zaj folyamat realizációinak legyenek tekinthetők.

3. tábla

A szezonális egységgyöktesztek eredményei az 1992. január–1998. szeptemberi mintában

Termék-csoport	$t(\pi_1)$	$t(\pi_2)$	$F(\pi_3, \pi_4)$	$F(\pi_5, \pi_6)$	$F(\pi_7, \pi_8)$	$F(\pi_9, \pi_{10})$	$F(\pi_{11}, \pi_{12})$
	egységgyökök						
	$1-L$	$1+L$	$1+L^2$	$1+\sqrt{3}L+L^2$	$1-\sqrt{3}L+L^2$	$1+L+L^2$	$1-L+L^2$
	szűrő alkalmazásával						
ÉLELM	-1,46	-5,01***	12,9***	15,0***	4,93*	8,78***	3,26
SZESZ	-1,44	-2,04	3,27	5,35*	3,14	8,41***	6,79**
RUHA	-1,35	-2,40*	9,79***	7,86**	9,80***	2,80	0,95
TARTÓS	-1,51	-1,33	4,24	3,50	6,93**	6,13**	8,69***
ENERG	-1,94	-2,29	4,56	7,00**	2,65	3,55	4,97*
EGYÉB	-0,77	-2,95**	5,55*	6,64**	7,64**	8,77***	5,20*
SZOLG	-1,17	-2,26	4,23	16,7***	4,82	12,4***	9,94***
CPI	-0,71	-3,92***	11,5***	39,0***	4,27	18,4***	3,90

Megjegyzés. $t(\pi_1)$ a π_1 koefficiens szignifikanciáját tesztelő t statisztika. $F(\pi_3, \pi_4)$ a $\pi_3=\pi_4=0$ nullhipotézis tesztelésére szolgáló F statisztika értéke.

* 10 százalékos szignifikancia;

** 5 százalékos szignifikancia;

*** 1 százalékos szignifikancia.

A bemutatott egységgyökök (vagy gyökpárok) és az $y_{i,t}$ segédváltozók között kölcsönösen egyértelmű hozzárendelés létesíthető. Be lehet továbbá látni, hogy a π_i koefficiens zérus értéket vesz fel abban az esetben, ha a kérdéses változóhoz rendelhető egységgyökök

¹² A pontos specifikációkat [2] (158. old.) tartalmazza. Az elméleti háttérrel pedig [7] nyújt tájékoztatást.

jelen van az idősorban. Például $\pi_1=0$ azt jelenti, hogy a folyamat nem szezonális egységgyököt tartalmaz, és differenciálásra van szükség a stacionaritás eléréséhez. Ha továbbá $\pi_3=\pi_4=0$, akkor belátható, hogy a $\pm i$ egységgyökpár is jelen van, és indokolt az $(1+L^2)$ szűrő alkalmazása. Ezen hipotézisek tesztelése a becsült koefficiensekre szokásos módon konstruált t , illetve F statisztikákkal lehetséges, de – mint az egységgyökteszteknél általában – a nullhipotézis alatt a standard aszimptotikus elmélet nem érvényes. A lényeges aszimptotikus eloszlások percentilisei Monte–Carlo-szimulációkkal származtathatók. (Lásd [7].)

E szezonális egységgyöktesztet az inflációs idősorokra elvégezve a 3. táblában látható eredmények adódnak.

Az eredmények azt sugallják, hogy az $(1-L^{12})$ szűrő általában túldifferenciálja a havi infláció idősorait, már kevésbé összetett transzformáció segítségével is el lehet érni a stacionaritást. Az egyszerű differencia képzése minden esetben szükségesnek látszik.

A következőkben azokat a szűrőket mutatom be, melyek használata mellett a szezonális egységgyöktesztet eredményeit figyelembe véve végül döntöttem. A transzformációk megválasztásakor a 3. tábla adatait bizonyos rugalmassággal vettem figyelembe. Nehéz ugyanis eldönteni, hogy egy tíz százalékon szignifikáns koefficiens elég bizonyíték-e a szóban forgó egységgyök hiányára. Ilyenkor a megfelelő elemi szűrő bevonásával és kihagyásával is transzformáltam az idősorot, majd megvizsgáltam, hogy melyikre illeszthető jobb tulajdonságokkal rendelkező ARMA-modell. Egyes kétértelmű helyzetekben azt az alternatívát választottam, mely jobban magyarázható transzformációhoz vezetett.

A HEGY-teszt által javasolt szűrő

Termékcsoport	Szűrő
ÉLELM	$(1-L)(1-L+L^2)$
SZESZ	$(1-L)(1+L)(1+L^2)(1+\sqrt{3}L+L^2)(1-\sqrt{3}L+L^2)$
RUHA	$(1-L)(1+L)(1+L+L^2)(1-L+L^2)$
TARTÓS	$(1-L)(1+L)(1+L^2)(1+\sqrt{3}L+L^2)$
ENERG	$1-L^{12}$
EGYÉB	$1-L$
SZOLG	$(1-L)(1+L)(1+L^2)(1-\sqrt{3}L+L^2)$
CPI	$(1-L)(1-\sqrt{3}L+L^2)(1-L+L^2)$

A transzformált idősorok modellezése

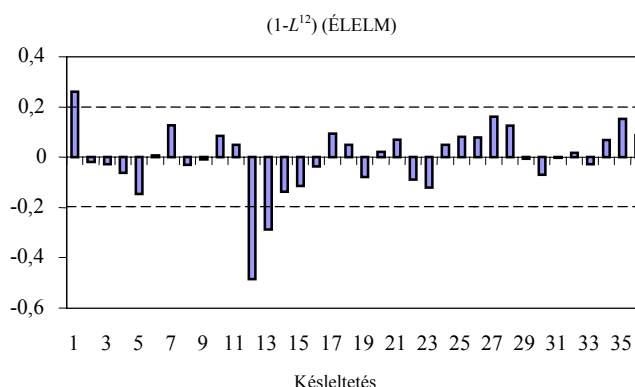
Az ARMA-tagok identifikációját a Box-Jenkins transzformáció után tekintem át. A hagyományos Box-Jenkins transzformáció alkalmazása a legtöbb esetben könnyebben értelmezhető korrelogramokhoz vezetett, mint a HEGY-metodológia. A transzformált idősorok korrelogramjai minden esetben indokolták egy szezonális (12-ed rendű) autoregresszív tag beiktatását. Általában számottevő bizonyíték volt egy elsőrendű mozgóátlagolású vagy autoregresszív tag jelenlétére is. Néhány esetben ötöd-, hatod- vagy hetedrendű autoregresszív tag beiktatására is sor került. A modellek végső kiválasztása-

nál nagy súllyal latba eső kritérium volt a becsült koefficiensek szignifikanciája,¹³ illetve a reziduumok autokorreláció-mentessége. (A pontos specifikációkat és a becsült koefficienseket a 6. tábla tartalmazza.)

Csak a CPI esetében bizonyult célszerűnek az a korlátozás, hogy a modell autoregresszív polinomja két polinom – egy szezonális és egy nem szezonális – szorzataként legyen felírva. A RUHA, a TARTÓS, az ENERG és az SZOLG változók esetében a szorzat alakú autoregresszív rész fenntartása autokorrelált reziduumokat eredményezett, vagy lényegesen csökkentette a mintán kívüli előrejelzés pontosságát. Ezért ezen időszakok modellezésekor az autoregresszív részt egyetlen (12-ed rendű) polinom formájában írtam fel, melyben azonban több tag együtthatója nullára van korlátozva. Az ÉLELM és a SZESZ változó modelljében elég volt egyetlen (12-ed rendű) autoregresszív tag használatát, így ezeknél ez a probléma természetesen nem merült fel.

Az identifikációnak ezt a lépését a szezonálisan differenciált ÉLELM idősor korrelogramjának (lásd a 2. ábrát) példáján szemléltetem.

2. ábra. A szezonálisan differenciált ÉLELM idősor korrelogramja



Szignifikáns autokorreláció-értékek figyelhetők meg az 1., a 12. és a 13. késleltetésnél. A korrelogram inszignifikáns értékeket mutat a 2.-tól a 11. késleltetésig, és az autokorrelációk újra „kihálnak” a 13. késleltetés után. Ez a minta egy MA(1) és egy AR(12) tag jelenlétét sugallja. Egy elméleti SARMA (1,0)^s (0,1) korrelogram¹⁴ csak annyiban különbözik az előbbi mintától, hogy a 11. késleltetésnél nem zérus, és az autokorrelációk visszatérnek a 23-25. késleltetéseknél. A modell együtthatói azonban vehetnek fel olyan értékeket, melyek hamar nullához közelítik az autokorrelációkat, és így a szignifikáns értékek hiánya a 24. késleltetés körül nem igazán mond ellent az elméleti mintának.

A HEGY-transzformáció után is vizsgáltam az ARMA-tagok identifikációját. A HEGY-metodológia a transzformált idősorok modellezéséhez szezonális dummy változókat is használ.

¹³ A rendellenes megfigyelések kezelése után a modellek reziduumai normális eloszlásúnak tekinthetők. A t statisztikák értelmezhetőségével tehát nincsenek gondok.

¹⁴ Az elméleti SARIMA-korrelogramoknak [5] 3. fejezete jó leírását adja.

Az ARMA-tagok identifikációjához ezért az

$$y_t = \sum_{s=1}^{12} \hat{\delta}_s D_{s,t} + \hat{u}_t$$

regresszió reziduumaival (azaz az $\hat{u}_t$ idősort) kell vizsgálni, ahol y a transzformált idősor szimbóluma. A kapott reziduumok korrelogramjai azonban meglehetősen nehezen értelmezhetőnek bizonyultak. A parciális autokorreláció-függvények inkább tiszta autoregressziókat sugalmaztak: változásukat nem a simaság, hanem a hirtelen ugrások jellemezték. A korrelogramok változása a legtöbb esetben sokkal folyamatosabb volt.

A felállított modellek tehát tiszta autoregressziók, melyek rendje 2 és 12 között változik. (A magasabb rendű modellek esetében természetesen sok tag együtthatója nullára van korlátozva.) Egy magasrendű autoregresszió egyébként akkor is jól közelítheti az igazi modellt, ha az mozgóátlagolású tagokat is tartalmaz. Az invertálható MA-polinommal rendelkező ARMA-modell ugyanis $AR(\infty)$ alakba írható fel, ahol a késleltetésszám növekedésével az együtthatók nullához tartanak, azaz a gyakorlatban elhanyagolhatóvá válnak.

A korrelogramokkal kapcsolatos bizonytalanságok miatt nagyon sok változatot kellett megvizsgálni, melyek közül az együtthatók szignifikanciája¹⁵ és a reziduumok fehér zajhoz közelsége alapján választottam ki a végső modellt. (A pontos specifikációk és a becsült együtthatók a 6. táblában láthatók.)

ENDOGENITÁS, EXOGENITÁS ÉS A RENDELLENES MEGFIGYELÉSEK

Az idősorelemzés egyik alapvető feltevése, hogy van egy ún. adatgeneráló folyamat, melyből az idősor megfigyelt értékei származnak. A felállított idősormodell tulajdonképpen az adatgeneráló folyamat reprezentációjának tekinthető. Bármely modell akkor jó, ha képes számot adni az idősor megfigyelt értékeiről anélkül, hogy a reziduumok extrém értékeket vennének fel, vagy bármilyen szabályszerűséget mutatnának.

A modellezés során gyakorta felmerül, hogy sok ténylegesen megfigyelt idősorban – a tanulmányozott inflációs idősorokban is – időnként élesen kiugró, az idősor addigi és azt követő alakulásába nem illő, rendellenes megfigyelések¹⁶ jelennek meg, melyek valószínűleg nem az idősor normális adatgeneráló folyamatából származnak. Ha ezeket a rendellenes megfigyeléseket potenciálisan vissza nem térő, az adatgeneráló folyamaton kívüli események okozták, akkor az előrejelző modell illesztésekor mindenképpen célszerű eltávolítani őket, éppen azért, mert egyszerűek és extrémek, és mert eltorzíthatják a rendes – és a jövőben feltehetőleg zavartalanul működő – adatgeneráló folyamat becslését.

Az outlierok kiszűréséhez minimálisan szükséges, hogy a fogyasztói árak alakulására ható eseményeket két nagy csoportba soroljuk:¹⁷

1. piaci események: olyan változások a keresletre vagy a kínálatra ható tényezőkben, melyek nem közvetlenül adminisztratív hatósági intézkedések következményei (ezeket az eseményeket sztochasztikusnak és endogénnek tekintem),

¹⁵ Itt is érvényes a 13. lábjegyzetben szereplő megjegyzés.

¹⁶ Az irodalomban szokásos még az „aberráns megfigyelés”, illetve az „outlier” kifejezés használata.

¹⁷ Hasonló elkülönítés más elemzésekben is megjelenik. (Lásd [10].)

2. *adminisztratív események*: a kereslet és a kínálat valamely tényezőjében, illetve közvetlenül az árban tapasztalható olyan változások, melyeknek oka valamilyen hatósági intézkedés: közvetlen ármeghatározás, egy-szeri nagyarányú forintleértékelés, az ÁFA módosítása stb. (ezeket az eseményeket egyszerűnek és exogénnek tekintem, és modellezésüktől eltekintek).¹⁸

A 90-es években a magyar inflációs folyamatot nagymértékben befolyásolták az adminisztratív intézkedések, azaz egyszeri, exogén sokkok. Feltehető azonban, hogy amint a piacgazdaság működése egyre zökkenőmentesebbé vált, annál ritkábban volt szükség hatósági beavatkozásokra (de az is lehet, hogy minél ritkábbak voltak a beavatkozások, annál tökéletesebben működött a piacgazdaság). Bármilyen legyen is a helyzet, elég nagy biztonsággal feltételezhető, hogy az utolsó nagy sokk, a Bokros-csomag óta a legtöbb áru és szolgáltatás árváltozását piaci (endogén) tényezők alakítják, és hasonló – hatósági intézkedések okozta – sokkok a jövőben már kisebb valószínűséggel (vagy egyáltalán nem) fordulnak elő. A múltbeli exogén sokkok tehát nem sokat árulnak el az infláció jövőbeli viselkedéséről, és inkább megnehezítik, mint segítik az előrejelzésre alkalmas modell felállítását.

Nyilvánvaló tehát, hogy kérdéses a statisztikai modellezés értelme azon árucsoportok esetén, melyek fogyasztói árának alakulását adminisztratív események irányítják (mint a szeszes italok és dohányárak vagy a háztartási energia esetében). E két árucsoport inflációs idősora tulajdonképpen exogén sokkok sorozata, így különösen élesen vetődik fel a kérdés, hogy az idősor múltbeli értékei mennyire informatívak a jövőbeli értékekre nézve.

Az additív outlier modell

A rendellenes megfigyelések kezelésére több elméleti lehetőség is kínálkozik. A hatósági idősorok (ENERG és SZESZ) esetén az additív outlier modell tűnt a legalkalmasabbnak.¹⁹ E megközelítés szerint a ténylegesen megfigyelt idősort (y_t) a következő módon lehet összeállítani:

$$y_t = x_t + [w_1 D_{1t}(t = s_1) + w_2 D_{2t}(t = s_2) + \dots + w_k D_{kt}(t = s_k)] \quad /9/$$

Itt x_t azt az idősort jelenti, melyet akkor figyelhetnénk meg, ha nem kerülne sor külső beavatkozásra (vagy nem lenne mérési hiba stb.). $D_{it}(t = s_i)$ olyan indikátor dummy-t jelöl ($i=1, \dots, k$), mely az egy értéket veszi fel az s_i időpontban, de máskülönben nulla. A w_i koefficiens ennek megfelelően a sokk nagyságát jelenti.

A SZESZ illetve a ENERG változókra megkíséreltem ezzel az elméleti megközelítéssel összhangban levő modellt illeszteni. A feltűnően kiugró megfigyeléseket eltávolítottam az idősorból,²⁰ és helyükre a két szomszédos megfigyelés átlagát tettem. Az így ka-

¹⁸ A hatósági intézkedések teljes exogenitása vitatható. Lehet ugyanis, hogy a hatóság passzívan reagál a gazdaságban bekövetkezett sztochasztikus eseményekre. Feltételezem azonban, hogy a hatóság nem előre rögzített szabályok alapján cselekszik, hanem minden esetben mérlegel. A hatósági intézkedések így módon elszakadnak a gazdasági folyamatok alakulásától.

¹⁹ A rendellenes megfigyelések különböző modellezési lehetőségeiről lásd [2] 6. fejezetét.

²⁰ Az ENERG változó esetén a következő időpontokhoz tartozó értékeket tekintettem outliernek: 1992. augusztus, 1992. október, 1993. január, 1995. január–március, 1996. május, 1997. február–március. A SZESZ változó esetében ugyanezek az időpontok: 1992. január–február, 1993. január, 1993. szeptember, 1994. január, 1994. augusztus, 1994. november, 1995. január, 1996. január, 1997. január, 1998. január. Az outlierok kiválasztásakor az időszak árpolitikai intézkedéseiből és az idősorok ábrájából indultam ki.

pott idősor (az elméleti modellben ez x_t -nek felel meg) remélhetőleg az eredeti változót ért sztochasztikus eseményeket reprezentálja. A SZESZ és a ENERG változó előrejelzett értékei egyedül az x_t komponens előrejelzésén alapulnak: az esetleges külső beavatkozások előrejelzésére nem tettem kísérletet.

Az innovációs outlier modell

A piacinak tekinthető idősorok esetében célszerű másképp felfogni a rendellenes megfigyelések keletkezését. Ebben az esetben az alapvető feltételezés, hogy a kiugró érték az idősort generáló folyamat hibatagjában jelentkezett. Egy ARMA (p, q) folyamat esetén például

$$a(L)y_t = b(L)[\varepsilon_t + D_t(t = s)] \quad /10/$$

alakban írható fel ez a szituáció.

A RUHA, a TARTÓS, az ÉLELM, az EGYÉB és a SZOLG változó esetén ezzel a megközelítéssel összhangban kíséreltem meg azonosítani a rendellenes megfigyeléseket.²¹ Először egy megfelelő SARIMA-modellt illesztettem a szóban forgó idősorra, és kiválasztottam azokat a megfigyeléseket, melyekre a reziduumok abszolút értéke meghaladta a regresszió standard hibájának kétszeresét. Ezeket a megfigyeléseket akkor tekintettem rendellenesnek, ha a Jarque–Bera-statisztika alapján a reziduumok hisztogramja durván megsértette a normális eloszlást, vagy ha a szóban forgó megfigyelés dátumából egyértelműen arra lehetett következtetni, hogy valamilyen központi intézkedés áll a kiugró érték hátterében. Az ily módon azonosított outlierrekhez indikátor dummy változókat konstruáltam, és ezek beiktatásával újrabecsültem a modellt. Ezt a módszert kiterjesztettem a SZESZ és a ENERG változók megtisztításával kapott idősorok modellezésére is.

A BECSÜLT MODELLEK

A becsült modelleket a 4–7. táblák mutatják be. A koefficiensek alatt zárójelben szereplő számok t statisztikák. A becslés során felhasznált minta 1992 januárjától 1999 áprilisáig terjed.

Mivel a Box–Jenkins-metodológia követésével felállított modellek tulajdonképpen egy 24 vagy 25-öd rendű autoregresszióval egyenértékűek, az első ténylegesen modellezett megfigyelés legalább két évvel későbbi a minta kezdeti időpontjánál. (Az első 24-25 megfigyelés kezdeti feltételként szolgál.) A HEGY-módszer általában ennél nagyobb tényleges minta használatát is lehetővé tenné, de az összehasonlíthatóság kedvéért a teljes minta kezdeti időpontját úgy választottam, hogy az első modellezett megfigyelés 1994 januárja legyen. A modellezett mintaperiódus ily módon általában 63 megfigyelést tartalmaz. (Eggyel kevesebbet a CPI-változó esetében, amikor az első modellezhető megfigyelés 1994. februári.)

²¹ Az ismertetendő módszer, melyet a rendellenes megfigyelések azonosítására és kezelésére ténylegesen használtam, nem teljesen konzisztens a /10/-es formulációval, amely forma a gyakorlatban igen nehezen végrehajtható paraméterkorlátozásokat igényelt volna.

4. tábla

Az inflációs idősorok Box–Jenkins-modelljei

Termék- csoport	SARIMA	Konstans	AR-tagok							MA-tag
			a szezonális	a nem szezonális						
			polinomban szereplő késleltetések							
			L^{12}	L	L^5	L^6	L^7	L^{12}	L	
			A késleltetések együtthatói							
			a_{1s}	a_1	a_5	a_6	a_7	a_{12}	b_1	
			Az együtthatók becslései							
ÉLELM	(1,1,0) ^s (0,0,1)	-0,0025 (-1,731)	-0,4974 (-6,050)						0,4448 (3,912)	
SZESZ	(1,1,0) ^s (0,0,1)	0,0003 (0,334)	-0,4179 (-3,4743)						0,5804 (5,394)	
RUHA	(0,1,0) ^s (12,1,1)	-0,0001 (-0,614)				0,4612 (4,020)		-0,5712 (-5,095)	-0,7884 (-11,1)	
TARTÓS	(0,1,0) ^s (12,0,0)	-0,0004 (-0,195)		0,6905 (8,760)		0,1668 (2,152)		-0,2857 (-3,623)		
ENERG	(0,1,0) ^s (12,0,0)	-0,0002 (-0,161)		0,3493 (3,035)			-0,3265 (-2,724)	-0,2793 (-2,351)		
EGYÉB	(0,1,0) ^s (12,0,0)	-0,0014 (-1,334)			0,2821 (2,197)			-0,2845 (-2,658)		
SZOLG	(0,1,0) ^s (12,0,0)	-0,0011 (-1,369)		0,3628 (3,790)				-0,4507 (-5,821)		
CPI	(1,1,0) ^s (1,0,0)	-0,0015 (-1,471)	-0,2879 (-3,093)	0,4385 (4,026)						

Megjegyzés. A modellek autoregresszív része a szezonális és a nem szezonális polinomok szorzataként adódik. Az AR- és MA-tagok illesztése előtt az inflációs idősorok a 2. táblában leírtaknak megfelelően voltak transzformálva. A zárójelben feltüntetett értékek t -statisztikák.

5. tábla

A Box–Jenkins-modellekben használt outlier dummy változók és további regressziós statisztikák

Termékcsoport	Dummy változók			R^2	A regresszió standard hibája (százalékpont)
ÉLELM				0,52	1,18
SZESZ				0,25	0,67
RUHA				0,65	0,50
TARTÓS	D9503	D9603		0,72	0,64
ENERG	D9505	D9511		0,48	1,03
EGYÉB	D9503	D9603	D9705	0,47	0,80
SZOLG				0,46	0,68
CPI	D9502	D9503		0,51	0,57

Megjegyzés. D9503 egy olyan dummy változót jelöl, mely az egy értéket veszi fel 1995 márciusában, és minden más időpontban nulla. A többi változó hasonlóképp értelmezendő. A regresszió standard hibája azt mutatja meg, hogy a modell illesztett értékei átlagosan hány százalékponttal térnek el az infláció valódi értékétől a becslési perióduson belül.

6. tábla

Az inflációs idősorok HEGY-modelljei

Termékcsoport	Az AR-programban szereplő késleltetések							
	L^1	L^2	L^3	L^4	L^5	L^6	L^7	L^{12}
	A késleltetések együtthatói							
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_{12}
Az együtthatók becslései								
ÉLELM	-1,4993 (-11,05)	-1,1711 (-6,038)	-0,3612 (-2,772)					
SZESZ	0,1921 (2,064)	-1,0610 (-10,22)		-0,5872 (-4,758)				
RUHA		0,3256 (2,399)	0,2545 (1,872)					
TARTÓS	2,3163 (18,31)	-2,7244 (-9,668)	2,4073 (6,630)	-2,2052 (-6,161)	1,5718 (5,806)	-0,5336 (-4,401)		
ENERG	0,3483 (2,810)						-0,3575 (-2,785)	-0,3486 (-2,670)
EGYÉB	-0,7275 (-5,166)	-0,5505 (-3,500)	-0,3276 (-3,315)					
SZOLG	-1,1313 (-11,54)	-0,6658 (-7,006)						
CPI	-2,6507 (-25,74)	-3,7902 (-15,55)	-3,5190 (-11,48)	-2,0924 (-8,696)	-0,6703 (-6,796)			-0,0316 (-1,956)

Megjegyzés. A becslt regressziók szezonális dummy változókat is tartalmaztak. Az AR-tagok illesztése előtt az inflációs idősorok a korábban leírt HEGY-teszt szűrőnek megfelelően voltak transzformálva. A zárójelben feltüntetett értékek t -statisztikák.

7. tábla

A HEGY-modellekben használt outlier dummy változók és további regressziós statisztikák

Termékcsoport	Dummy változók		R^2	A regresszió standard hibája (százalékpont)
ÉLELM	D9401	D9506	0,90	1,07
SZESZ			0,77	0,71
RUHA			0,79	0,44
TARTÓS			0,97	0,61
ENERG	D9505	D9511	0,60	1,03
EGYÉB	D9503		0,86	0,71
SZOLG			0,95	0,59
CPI			0,99	0,54

Megjegyzés. Lásd az 5. tábla megjegyzését.

ELŐREJELZÉSI EREDMÉNYEK

A CPI havi és 12 havi inflációs rátájának 1999-re vonatkozó öt különböző előrejelzését a 8. és a 9. tábla tartalmazza. A táblákban feltüntetett hibahatárok közelítőleg 95 szá-

zalékos konfidencia-intervallumot jelölnek.²² A havi inflációhoz $\pm x$ alakban megadott értékek az előrejelzések standard hibájának kétszeresével egyenlők. Az éves rátákhoz a $(-x; +y)$ formában felírt intervallumok az ún. bootstrap-eljárás felhasználásával készültek. Ezen intervallumok előállításához a kérdéses modell reziduumaiknak becsült varianciájával megegyező varianciájú normális eloszlásból fehér zajt generáltam, és ezt az idősort hibatagként felhasználva szimuláltam az egyenletet az előrejelzés időszakában. (A Jarque–Bera-statisztika alapján mindegyik modellben normális eloszlásúnak tekinthetők a reziduumok.) Az ily módon kapott adatokból éves inflációs rátákat számoltam, és meghatároztam az előrejelzésektől való eltérésüket. Az eljárást ezerszer ismételve az előrejelzési hibák empirikus bootstrap-eloszlása és a kívánt percentilisek kielégítő pontossággal meghatározhatók.

8. tábla

A CPI előrejelzései 1999-re

Hónap	Havi inflációs ráta (előző hónap=100,0)				Implikált éves inflációs ráta (előző év azonos hónapja=100,0)			
	Box–Jenkins	eltérés	HEGY	eltérés	Box–Jenkins	eltérés	HEGY	eltérés
I (tény)	102,5		102,5		109,8		109,8	
II (tény)	101,3		101,3		109,4		109,4	
III (tény)	101,2		101,2		109,3		109,3	
IV (tény)	101,2		101,2		109,4		109,4	
V	101,2	$\pm 1,2$	101,5	$\pm 1,2$	109,5	$(-1,1; +1,0)$	109,9	$(-1,5; +1,8)$
VI	100,6	$\pm 1,2$	100,8	$\pm 1,6$	109,8	$(-2,0; +1,7)$	110,4	$(-1,9; +2,1)$
VII	99,7	$\pm 1,2$	99,5	$\pm 1,8$	109,7	$(-2,7; +2,2)$	110,1	$(-5,3; +2,1)$
VIII	99,6	$\pm 1,2$	99,3	$\pm 1,8$	109,7	$(-3,4; +2,7)$	109,7	$(-6,9; +2,1)$
IX	100,7	$\pm 1,2$	100,3	$\pm 1,8$	109,7	$(-3,8; +3,1)$	109,4	$(-9,5; +1,2)$
X	100,8	$\pm 1,2$	100,6	$\pm 1,8$	109,6	$(-4,2; +3,4)$	109,0	$(-8,2; +3,8)$
XI	100,4	$\pm 1,2$	100,2	$\pm 1,8$	109,7	$(-4,9; +3,7)$	108,9	$(-9,1; +4,3)$
XII	100,4	$\pm 1,2$	100,1	$\pm 1,9$	109,8	$(-5,3; +4,2)$	108,8	$(-8,2; +7,0)$

9. tábla

A CPI súlyozott átlagként számolt előrejelzései 1999-re

Hónap	Havi inflációs ráta súlyozott átlaga (Index: előző hónap=100,0)			Implikált éves inflációs ráta súlyozott átlaga (Index: előző év azonos hónapja=100,0)		
	Box–Jenkins	HEGY	VEGYES*	Box–Jenkins	HEGY	VEGYES*
I (tény)	102,5	102,5	102,5	109,8	109,8	109,8
II (tény)	101,3	101,3	101,3	109,4	109,4	109,4
III (tény)	101,2	101,2	101,2	109,3	109,3	109,3
IV (tény)	101,2	101,2	101,2	109,4	109,4	109,4

(A tábla folytatása a következő oldalon.)

²² A modellek a havi inflációs rátákat jelzik előre közvetlenül, az éves (12 havi) inflációs ráták pedig a havi inflációs ráták szorzatai. Pontosabban, ha $i(0)$ jelöli az i havi inflációs rátát, akkor az éves (12 havi) inflációs ráta a $\prod_{k=0}^{11} [1 + i(-k)] - 1$ képlet szerint számolható, ahol $i(-k)$ a k hónappal korábbi havi inflációs ráta szimbóluma.

(Folytatás.)

Hónap	Havi inflációs ráta súlyozott átlaga (Index: előző hónap=100,0)			Implikált éves inflációs ráta súlyozott átlaga (Index: előző év azonos hónapja=100,0)		
	Box-Jenkins	HEGY	VEGYES*	Box-Jenkins	HEGY	VEGYES*
V	101,0 (-0,6; +0,8)	100,7 (-0,7; +0,6)	101,1 (-1,0; +0,5)	109,3 (-0,7; +0,9)	109,0 (-1,2; +0,4)	109,5 (-1,1; +0,5)
VI	100,8 (-0,7; +0,7)	100,4 (-0,7; +0,9)	101,0 (-0,7; +0,8)	109,8 (-1,2; +1,2)	109,1 (-1,3; +0,8)	110,2 (-1,4; +0,9)
VII	99,8 (-0,7; +0,7)	100,0 (-0,8; +1,0)	100,0 (-0,5; +0,8)	109,8 (-1,5; +1,6)	109,3 (-1,5; +1,5)	110,4 (-1,7; +1,3)
VIII	99,8 (-0,7; +0,7)	99,8 (-0,8; +1,0)	99,8 (-0,6; +0,8)	110,0 (-1,7; +1,9)	109,6 (-1,8; +1,7)	110,6 (-1,8; +1,7)
IX	100,9 (-0,7; +0,7)	101,0 (-0,9; +1,0)	100,9 (-0,8; +0,7)	110,3 (-2,0; +2,1)	110,1 (-2,4; +2,0)	111,0 (-2,1; +1,9)
X	100,9 (-0,8; +0,7)	101,1 (-0,9; +1,0)	101,1 (-0,7; +0,8)	110,3 (-2,2; +2,2)	110,3 (-2,7; +2,2)	111,2 (-2,5; +2,3)
XI	100,5 (-0,8; +0,7)	100,2 (-0,9; +1,1)	100,7 (-0,7; +0,9)	110,5 (-2,6; +2,6)	110,2 (-2,8; +2,7)	111,7 (-2,6; +2,6)
XII	100,4 (-0,8; +0,7)	100,5 (-0,9; +1,1)	100,6 (-0,7; +0,9)	110,6 (-2,8; +2,6)	110,5 (-3,0; +2,9)	112,0 (-2,8; +3,0)

* A „legjobb” modellek előrejelzéseiből készített súlyozott átlag. A jóság kritériuma a mintán kívüli előrejelzési teljesítmény volt (lásd a következőkben). A vegyes súlyozott átlag konkrét összetétele: az ÉLELM, a SZESZ, a TARTÓS, az ENERG és a SZOLG változó előrejelzését a Box-Jenkins-modell adja, míg a RUHA és az EGYÉB változók előrejelzése HEGER-modellből származik.

Megjegyzés. Az előrejelzések készítésekor felhasznált legfrissebb adat tehát 1999. áprilisi.

Mintán kívüli előrejelzések 1998. január–decemberre

Valamely előrejelző modell jóságát alapvetően az szabja meg, hogy az általa előrejelzett értékek mennyire térnek el a ténylegestől adott hosszabb időszak folyamán. Ezt a kritériumot számos statisztika segítségével lehet leírni, leggyakrabban vagy az előrejelzett és a valós értékek négyzetes eltéréseinek átlagából vont gyök értékét (Root Mean Squared Error – RMSE) vagy az előrejelzési hibák abszolút értékének egyszerű számtani átlagát (Mean Absolute Error – MAE) használják e célra. A másik lehetséges mutató a Theil-féle egyenlőtlenségi koefficiens (Theil inequality coefficient) és a kapcsolódó egyenlőtlenségi arányszámok (proportions of inequality), melyek elvileg nemcsak az előrejelzési hiba mértékéről, hanem jellegéről, forrásáról is tájékoztatást adnak. (A Theil-statisztika értelmezéséről [9] 340–342. old. nyújt tájékoztatást.)

Teljesebb képet kapunk az előrejelző modellek jóságáról, ha az említett illeszkedési statisztikákat nem kizárólag önmagukban vizsgáljuk, hanem bizonyos referenciapontokat jelölünk ki. Ilyen viszonyítási alapok lehetnek például a különböző naiv módszerekkel származtatható előrejelzések. Egy összetettebb előrejelző modell felállításának ugyanis akkor van igazi értelme, ha ezáltal csökken az előrejelzések átlagos hibája.²³

Viszonyítási alapként két naiv modellt vizsgáltam. Az első a kérdéses hónap inflációs rátáját az előző év azonos hónapjában megfigyelt értékkel jelzi előre. Ez az

²³ Ez az elvárás akkor valósulhat meg, ha az idősor múltbeli alakulásában vannak olyan információk, melyeket a naiv modellek számításán kívül hagynak, de melyeket a bonyolultabb modellek tudnak használni.

eljárás elméletileg is megalapozott, abban az esetben ha az infláció szezonális véletlen ingadozást követ, azaz

$$y_t = y_{t-12} + \varepsilon_t, \quad /11/$$

ahol y_t a havi inflációs rátát jelenti, és ε_t a fehér zajt. Ebben az esetben ugyanis y_t legjobb előrejelzését valóban y_{t-12} biztosítja.

A második naiv modell inkább determinisztikus ihletésű. Ebben a modellben az infláció szintje egy lineáris időtrend, 12 szezonális dummy változó és egy hibatag függvényében van felírva. Képletben:

$$y_t = \beta t + \sum_{s=1}^{12} \delta_s D_{s,t} + v_t, \quad /12/$$

ahol $D_{s,t}$ ($s=1, \dots, 12$) olyan szezonális dummy változót jelöl, mely az egy értéket vesz fel az év s -edik hónapjában, egyébként pedig nulla. Az inflációt 1992 januárjában kezdődő minták alapján becsültem, vagyis ugyanazt az információs bázist használtam fel, mint a Box–Jenkins- és a HEGY-modellek felállításánál.

A havi inflációs idősorokra – a különböző modellek felhasználásával – kétféle módon készítettem mintán kívüli előrejelzéseket. (Az előrejelzés periódusa az 1998 januárjától decemberig tartó tizenkét hónapos időszak volt.) Az ún. *dinamikus előrejelzés* 1997 decemberéig terjedő adatok felhasználásával készült, és azt mutatja meg, hogy 1998 legelején milyen infláció várható az év tizenkét hónapjában. A dinamikus előrejelzés az első perióduson túli előrejelzések készítésekor a változó előzőleg jelzett értékeit használja ott, ahol a változó késleltetett értékeire van szükség. 1998. februári előrejelzés például felhasznál(hat)ja a januárra előrejelzett értéket, a márciusi előrejelzés pedig az előző két hónap előrejelzett (és nem tényleges) értékeit stb. A *statikus előrejelzés* ezzel szemben a tényleges késleltetett értékeket használja, így statikus előrejelzés nyilván csak akkor készíthető, ha rendelkezésre állnak a szükséges megfigyelések. A statikus előrejelzés tulajdonképpen egyhavi előrejelzések sorozata, melyek az előző hónapig rendelkezésre álló adatok alapján megbecsült modellből származnak. Az 1998. áprilisi előrejelzés például 1998 márciusáig terjedő adatok segítségével készül, az 1998. májusi már felhasználja az áprilisi megfigyelést is stb. Az első naiv modell esetében a statikus és a dinamikus előrejelzés egybeesik. Nyilvánvaló továbbá, hogy a statikus előrejelzésektől alacsonyabb hibastatisztikák várhatók.

A CPI különböző előrejelzéseire vonatkozó részletes eredményeket a 10. és 11. táblák mutatják be. A 12. tábla a részaggregátumokról közöl tömör összefoglalót.

A mintán kívüli előrejelzésekből leszűrhető általános tanulságok közül az első az, hogy mind a Box–Jenkins-, mind a HEGY-modellek csak számottevő hibával képesek előrejelezni a havi inflációs idősorok alakulását, különösen hosszabb távon. Egy közepes pontossággal előrejelezhető részaggregátumot körülbelül 0,6 százalékpontos átlagos hiba jellemez a tizenkét hónapos periódus alatt, ami havi infláció esetében jelentős bizonytalanság. Az egyhavi (statikus) előrejelzések átlagos hibája általában kisebb, de a legtöbb modellnél eléri vagy meghaladja a 0,5 százalékpontot. Ráadásul több esetben (SZESZ, TARTÓS, SZOLG) az első naiv modell adta a legjobb dinamikus és statikus előrejelzéseket, megkérdőjelezve ezáltal az összetett modellek specifikációját. 1998-ban a CPI

inflációs rátája a várakozásokat jelentősen meghaladó mértékben csökkent: a 10. és a 11. tábla tanúsága szerint ez a csökkenés az infláció múltbeli értékeiből nem volt előrejelezhető.

10. tábla

A CPI havi inflációs rátája 1998-ban

Hónap, mutató	Tény	Dinamikus				Statikus			
		mintán kívüli előrejelzés (százalék)							
		Box– Jenkins	HEGY	NAIV 1	NAIV 2	Box– Jenkins	HEGY	NAIV 1	NAIV 2
I	3,0	3,9	4,1	3,7	4,1	3,9	4,1	3,7	4,1
II	1,7	2,2	1,9	2,2	2,1	1,8	0,9	2,2	2,0
III	1,3	1,9	1,7	1,9	1,8	1,6	1,6	1,9	1,7
IV	1,0	1,4	1,8	1,4	1,4	1,1	1,7	1,4	1,3
V	1,2	1,4	1,9	1,3	1,4	1,2	1,3	1,3	1,3
VI	0,3	1,5	1,2	1,7	0,8	1,3	0,6	1,7	0,7
VII	-0,2	-0,0	0,4	-0,1	0,5	-0,6	-0,3	-0,1	0,3
VIII	-0,4	0,2	-0,4	0,2	0,7	0,0	-0,8	0,2	0,6
IX	0,6	1,4	1,0	1,4	2,0	1,1	1,3	1,4	1,8
X	0,9	1,2	0,9	1,1	1,8	0,8	0,9	1,1	1,5
XI	0,3	1,1	1,4	1,2	1,2	0,9	0,6	1,2	1,0
XII	0,3	1,0	1,4	1,1	1,1	0,6	0,4	1,1	0,8
RMSE*		0,65	0,72	0,69	0,81	0,49	0,52	0,69	0,68
MAE*		0,59	0,61	0,59	0,74	0,39	0,41	0,59	0,59
THEIL		0,22	0,24	0,24	0,27	0,18	0,19	0,24	0,23
BIAS		0,82	0,71	0,73	0,83	0,39	0,13	0,73	0,76
VAR		0,01	0,04	0,00	0,00	0,11	0,29	0,00	0,01
COV		0,17	0,25	0,26	0,17	0,50	0,57	0,26	0,23

* Százalékpontban.

11. tábla

A CPI havi inflációs rátája 1998-ban

Hónap, mutató	Tény	Dinamikus			Statikus		
		súlyozott átlagként számított mintán kívüli előrejelzés (százalék)					
		Box–Jenkins	HEGY	VEGYES*	Box–Jenkins	HEGY	VEGYES*
I	3,0	3,4	3,2	3,2	3,4	3,2	3,2
II	1,7	2,0	1,8	1,9	1,8	1,6	1,7
III	1,3	1,6	1,3	1,5	1,5	1,2	1,5
IV	1,0	1,3	1,3	1,1	1,1	1,2	1,0
V	1,2	1,1	1,2	1,1	0,9	0,7	1,0
VI	0,3	1,2	1,0	1,1	1,3	1,0	1,2
VII	-0,2	0,0	0,3	-0,0	-0,5	0,1	-0,4
VIII	-0,4	0,1	-0,2	-0,0	0,1	0,0	0,0
IX	0,6	1,3	1,0	1,2	1,1	1,3	1,1
X	0,9	1,1	1,1	1,0	0,8	1,2	0,8
XI	0,3	0,9	1,0	0,9	0,8	0,5	0,8
XII	0,3	0,9	1,0	0,8	0,7	0,4	0,6

(A tábla folytatása a következő oldalon.)

(Folytatás.)

Hónap, mutató	Tény	Dinamikus			Statikus		
		súlyozott átlagként számított mintán kívüli előrejelzés (százalék)					
		Box–Jenkins	HEGY	VEGYES*	Box–Jenkins	HEGY	VEGYES*
RMSE**		0,48	0,42	0,40	0,44	0,38	0,39
MAE**		0,43	0,33	0,33	0,37	0,32	0,31
THEIL		0,18	0,16	0,15	0,17	0,15	0,15
BIAS		0,72	0,63	0,61	0,36	0,28	0,33
VAR		0,01	0,06	0,04	0,00	0,03	0,00
COV		0,28	0,31	0,35	0,64	0,69	0,67

* Lásd a 9. táblánál.

** Százalékpontban.

12. tábla

Az RMSE mutató
(1998. január – december)

Termékcsoport	Dinamikus				Statikus			
	mintán kívüli előrejelzés (százalékpont)							
	Box–Jenkins	HEGY	NAIV 1	NAIV 2	Box–Jenkins	HEGY	NAIV 1	NAIV 2
ÉLELM	1,13	1,44	1,56	1,38	1,00	1,15	1,56	1,34
SZESZ	0,65	0,73	0,59	0,74	0,63	0,67	0,59	0,86
RUHA	0,44	0,44	0,48	0,80	0,44	0,40	0,48	0,71
TARTÓS	0,59	0,61	0,30	0,66	0,41	0,57	0,30	0,60
ENERG	1,05	1,06	1,23	1,44	1,12	1,16	1,23	1,25
EGYÉB	0,64	0,56	0,72	0,75	0,67	0,58	0,72	0,65
SZOLG	0,65	0,65	0,54	0,72	0,59	0,75	0,54	0,68
CPI	0,65	0,72	0,69	0,81	0,49	0,52	0,69	0,68

A második fontos megállapítás, hogy a szezonális egységgyöktesztek eredményei alapján transzformált idősorokra illesztett modellek – a HEGY-modellek – előrejelzései általában nem jobbak a hagyományos Box–Jenkins-transzformációra épülő modellekénél. Az ÉLELM, a SZESZ, a TARTÓS, az ENERG, a SZOLG és a CPI változó esetében például a Box–Jenkins-metodológia a HEGY-módszernél kedvezőbb (vagy legalább olyan jó) illeszkedési statisztikákhoz vezetett mind a dinamikus, mind a statikus előrejelzéseket illetően. A maradék két változó (RUHA és EGYÉB) esetében a HEGY-modell bizonyult jobbnak (bár a RUHA változónál ez csak a statikus előrejelzésnél egyértelmű). A két modellesoport pontossága közti különbségek mindenestre korántsem jelentősek.

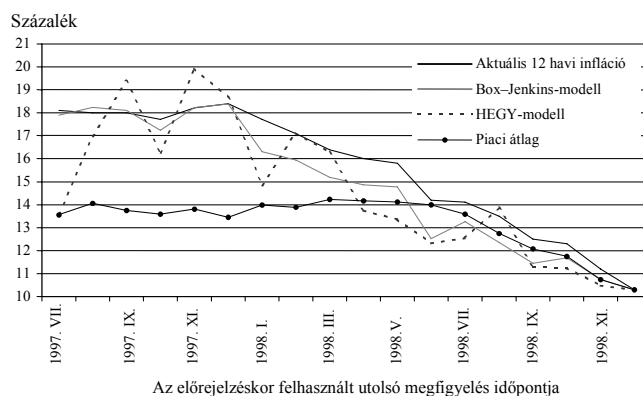
Harmadik általános tanulságként azt jegyezhetjük meg, hogy a CPI súlyozott átlagként előállított becslései megbízhatóbbnak látszanak, mint a közvetlen előrejelzések.

AZ IDŐSORMODELLEK ÉS A PIACI SZEREPLŐK ELŐREJELZÉSEI

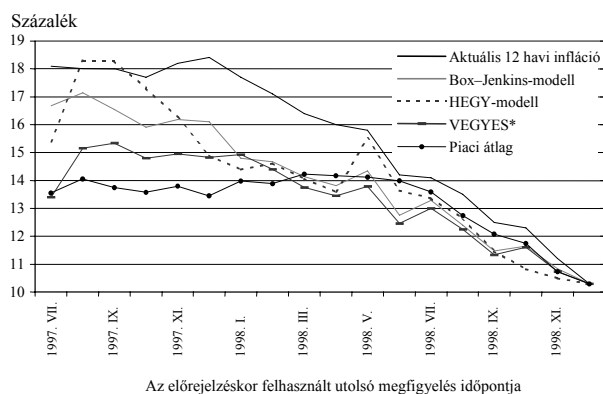
A fogyasztói árindex előrejelzésére használt öt módszer (Box–Jenkins, HEGY és a három súlyozott átlag) eredményeinek további értékelésére megvizsgáljuk, miképp válto-

zott időben a modellek által 1998 decemberére prognosztizált 12 havi infláció, amint az e dátumhoz közeledve havonta újabb és újabb adatok láttak napvilágot, és összevetjük ezeket az értékeket a piaci szereplők várakozásaival. A 3. és a 4. ábrán az 1998-as éves inflációra vonatkozó dinamikus előrejelzések sorozatai láthatók, a sorozatok megfelelő tagjainak elkészítéséhez felhasznált utolsó megfigyelés alapján.²⁴ A 3. ábra a közvetlenül a fogyasztói árindexre illesztett két modell előrejelzéseit mutatja, míg a 4. ábra a súlyozott átlagokét. A legkorábbi előrejelzés az 1997 júliusáig tartó minta alapján készült. Az ábrák tartalmazzák a 12 havi infláció előrejelzéskori aktuális értékét és a piaci szereplők várakozásait, melyet brókercegek, bankok, illetve kutatók által készített előrejelzések átlagaként határoztam meg.

3. ábra. A CPI 12 havi inflációs rátájának előrejelzése 1998 decemberére



4. ábra. A CPI 12 havi inflációs rátájának előrejelzése 1998 decemberére (súlyozott átlagok)



* Lásd a 9. táblánál.

²⁴ Adott hónap inflációs rátáját a KSH a következő hónap 11-én közli. Az 1998. márciusi előrejelzések tehát 1998. április 11 után készültek, és 1998. május 11-ig voltak érvényesek. Az egyszerűsítés érdekében azonban 1998. márciusi előrejelzésen azt az előrejelzést fogom érteni, mely az 1998. márciusi adatok alapján készült, tekintet nélkül arra, hogy ez az adat csak később kerül nyilvánosságra.

Az első szembetűnő tény, hogy az idősormodellekből származó előrejelzéseket az új információk megjelenése folyamatosan módosította. A HEGY-metodológia alkalmazásával készült előrejelzések a legérzékenyebbek az új adatokra. A 3. ábrán megfigyelhető, hogy az 1997 végéig terjedő időszakban a Box–Jenkins- és a HEGY-modellek által 1998-ra jelzett éves infláció értéke a 12 havi infláció tényleges (aktuális) értéke körül ingadozott. Ezzel szemben a piac már ekkor is jelentős csökkenésre számított, és a súlyozott átlagok előrejelzései is jóval alatta maradtak a tényleges értékeknek (kivétel: a HEGY-modell 1997. augusztus–szeptemberi súlyozott átlaga). 1998 januárjában a Box–Jenkins és a HEGY-modellek számára is láthatóvá vált az infláció bizonyos mértékű jövőbeni csökkenése; ettől az időponttól kezdve a jelzett értékek alacsonyabbak voltak a tényleges értéknél. Az infláció csökkenése azonban újra és újra meghaladta az egyes modellek által jelzett mértéket, ezért egy bizonyos időponttól kezdve valamennyi előrejelzés csökkenő tendenciát mutat. A piaci előrejelzésnek kellett a legkésőbb igazodnia: először 1998 júniusában teszik kétségesse az addigi 13,5–14 százalékos prognózist.

A 3. és a 4. ábra azt sugallja, hogy az idősormodelleken alapuló inflációs előrejelzések rövid távon – és csak rövid távon – felveszik a versenyt a piaci előrejelzésekkel. A kérdés pontos megválaszolása a különböző időhorizontokhoz rendelhető átlagos előrejelzési hibák formális meghatározását kívánja az egyes modellek esetében. Ezt a következők szerint végeztem el:

– minden egyes előrejelzési hónapban meghatároztam az 1998. decemberi 12 havi infláció jelzett és tényleges értéke közti különbség abszolút értékét, vagyis az adott hónap adataival bezárólag készített előrejelzéshez tartozó abszolút hibát;

– az egyes hónapokhoz hozzárendeltem a szóban forgó és az azt követő hónapok abszolút előrejelzési hibáinak átlagát, formálisan:

$$t \mapsto \frac{1}{\{98:12-t\}} \sum_{i=t}^{98:11} |f_i - 10,3\%|, t = 97:07, \dots, 98:11,$$

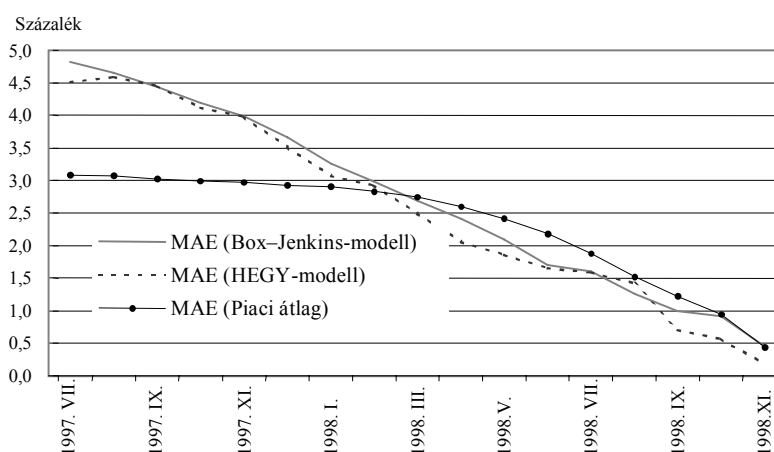
ahol f_i az 1998-as éves inflációnak az i hónap adataival bezárólag készült előrejelzése, a $\{98:12-t\}$ szimbólum a t hónap és 1998 decembere közti hónapok számát jelenti (például $\{98:12-98:11\}=1$, $\{98:12-97:11\}=13$ stb.).

Az így definiált hozzárendelést az 5. és a 6. ábra jeleníti meg. Az ábrák 1998 márciusához tartozó értéke például azt az átlagos hibát mutatja, amit a szóban forgó modell 9 előrejelzési időszak folyamán, 1998 márciusától 1998 novemberéig vét. Viszonyítási alapként mindkét ábra tartalmazza a piaci várakozások átlagos abszolút hibáját is.

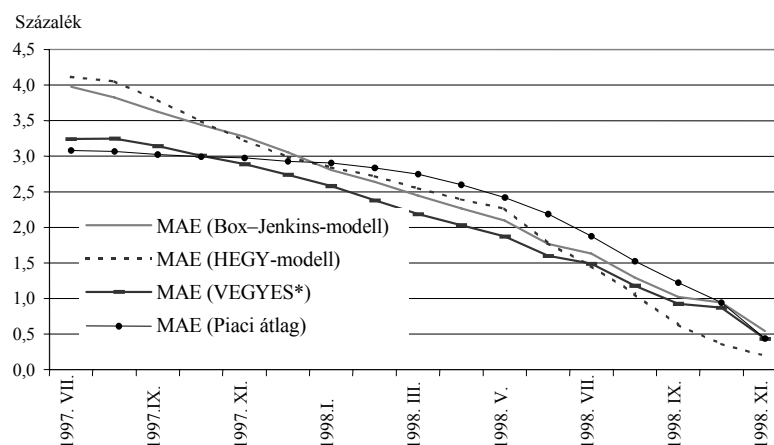
Az 5. és a 6. ábra igazolja az előbb megfogalmazott feltételezést, miszerint a vizsgált idősormodellek rövid távon viszonylag pontos, a piac előrejelzéseinél jobb előrejelzéseket adnak. Az 5. ábra jól mutatja, hogy a Box–Jenkins- és a HEGY-modellek előrejelzései 1998 márciusától kezdve – vagyis a decemberi adat megjelenése előtti utolsó 9 előrejelzési időszakban – már valamivel kisebb átlagos hibát követek el, mint a piaci előrejelzés. A súlyozott átlagok esetében ez a pont még korábbra tehető: a egyes súlyozott átlag már 12 hónapra visszatekintve is kisebb átlagos hibával jelzett előre, mint a piac. Az idősormodellek teljesítménye az utolsó 7-8 hónap tekintetében előzte meg legjobban a piaci előrejelzését, az átlagos abszolút hibák közti különbség ilyenkor a 0,5-0,7 százalékpontot is elérte. Érdekes viszont, hogy nagyon rövid távon (1-2 hónap) az

idősormodellek fölénye – a CPI HEGY-modelljét és a részaggregátumok HEGY-modellekből képzett súlyozott átlagát kivéve – elenyészik, a piaci és az idősormodellek által elkövetett előrejelzési hiba gyakorlatilag egyenlővé válik.

5. ábra. Az 1998-as év/év infláció előrejelzéseinek átlagos hibája különböző hosszúságú előrejelzési időszakokban



6. ábra. Az 1998-as év/év infláció előrejelzései átlagos hibájának súlyozott átlaga különböző hosszúságú előrejelzési időszakokban



* Lásd a 9. táblánál.

A havi statikus előrejelzésekhez 1998 novemberétől intézményenkénti bontásban is rendelkezésre állnak az egy hónapra előretekintő piaci várakozások, melyeket a 13. tábla mutat. (Az előző, hosszabb távra előretekintő piaci előrejelzéseknél sajnos csak az előre-

jelzések értékei állnak rendelkezésre, a készítői nem.) Ezen előrejelzések közlésének időpontjában a kérdéses hónapot megelőző havi infláció már ismert volt (statikus előrejelzésekről van szó). Az idősormodellek által meghatározott megfelelő előrejelzések a 14. táblában láthatók. A novemberi és a decemberi értékek konzisztensek azokkal a havi inflációra vonatkozó statikus előrejelzésekkel, melyeket a 10. és a 11. tábla tartalmaz.

13. tábla

*Piaci szereplők által előrejelzett 12 havi inflációs értékek**
(Index: előző év azonos hónapja = 100,0)

Piaci szereplő	1998. november	1998. december	1999. január	1999. február	1999. március	1999. április
AB-Moneta	112,1		109,3	109,4	109,1	109,2
Bank of America				109,3		
Budapest Bank		110,8			109,3	109,2
CA IB Securities		110,4		109,5	109,0	109,0
Citibank	112,0	110,6	109,5	109,4	109,3	109,1
Erste Bank			109,5	109,5		
Fundamenta			109,7	109,3	109,0	109,2
GKI	111,2	110,7	109,4	109,7		
IE-NYB			109,4			
ING Barings				109,2	108,8	108,9
K&H Brókerház		110,4				
MKB Securities	112,0				109,1	109,0
Nomura				109,3	109,1	
OTP Securities			109,6		109,0	109,1
Postabank	111,6		109,4	109,3	109,2	109,1
Rabobank			109,3			
Raiffeisen	112,1			109,3	108,8	108,9
S&P MMS	111,8	110,6	109,8		109,1	109,2
Takarék Bróker	112,1	110,6	109,7		109,1	109,1
Warburg Dillon		110,6		109,4	109,1	109,0
Átlag	111,86	110,61	109,52	109,38	109,09	109,08
Medián	112,0	110,6	109,5	109,4	109,1	109,1
Legmagasabb	112,1	110,8	109,8	109,7	109,3	109,2
Legalacsonyabb	111,2	110,4	109,3	109,2	108,8	108,9
Tényleges	111,2	110,3	109,8	109,4	109,3	109,4

* Az előrejelzés készítésekor az előző hónap inflációs rátája már rendelkezésre állt.

Forrás: Reuter's Monthly Survey of Forecasts for the Hungarian Economy. Nov. 1998. Apr.. 1999.

A 13. és a 14. tábla összevetéséből látható, hogy az idősormodellek és a piaci szereplők egy hónappal előre történő előrejelzései általában elég hasonlóak. 1998 végén az infláció gyors csökkenése mindegyik idősormodellel és piaci előrejelző számára meglepetés volt. Az idősor-előrejelzések mind novemberben, mind decemberben a legalacsonyabb és a legmagasabb piaci előrejelzés által kijelölt intervallumban voltak. 1998 novemberében az idősormodellek átlaga esett közelebb a tényleges inflációs értékhez, míg decemberben egy hajszálnyival a piaci várakozások bizonyultak jobbaknak.

Az 1999. januári piaci és idősor-előrejelzések már nem voltak ilyen közel egymáshoz. Az idősormodellek által jelzett legalacsonyabb érték (9,9%) is magasabb volt, mint a

piaci szereplők várakozásainak maximuma (9,8%). A két csoport előrejelzéseinek átlagához azonban ugyanakkora abszolút hiba járul, hiszen a januári infláció körülbelül ugyanolyan mértékben haladta meg a piaci várakozásokat, mint amilyen mértékben az idősormodellek felülbecsülték azt. Az 1999 februárjára vonatkozó idősor-előrejelzések átlaga szintén magasabb volt a piaci előrejelzésekenél, de a két csoport maximuma megegyezett, sőt a legkisebb előrejelzést (8,8%) a HEGY idősormodell produkálta. Végül a piaci előrejelzések átlaga bizonyult valamivel pontosabbnak; a legtöbb idősormodell ugyanis 2-3 tized százalékponttal felülbecsülte a februári inflációt. Márciusra és áprilisra szintén az idősormodellek vártak valamivel magasabb inflációs rátát, és ezzel jobbnak bizonyultak a piaci előrejelzéseknél, melyek átlagban 0,2-0,3 százalékponttal alulbecsülték a szóban forgó két hónap inflációs rátáját.

14. tábla

*Az idősormodellek által előrejelzett 12 havi inflációs értékek**
(Index: előző év azonos hónapja=100,0)

Módszer, érték	1998. november	1998. december	1999. január	1999. február	1999. március	1999. április
Box–Jenkins	112,0	110,7	110,2	109,6	109,3	109,2
HEGY	111,6	110,5	110,2	108,8	110,0	109,0
Súlyozott átlag						
Box–Jenkins	111,9	110,8	110,1	109,7	109,3	109,2
HEGY	111,5	110,5	110,0	109,5	109,1	109,0
Vegyes	111,9	100,7	109,9	109,7	109,2	109,3
Átlag	111,78	110,64	110,08	109,46	109,38	109,14
Medián	111,9	110,7	110,1	109,6	109,3	109,2
Legmagasabb	112,0	110,8	110,2	109,7	110,0	109,3
Legalacsonyabb	111,5	110,5	109,9	108,8	109,1	109,0
<i>Tényleges</i>	<i>111,2</i>	<i>110,3</i>	<i>109,8</i>	<i>109,4</i>	<i>109,3</i>	<i>109,4</i>

* Lásd a 13. táblánál.

Sokféle szempontból összevethetnénk még a piaci szereplők és az idősormodellek előrejelzéseit, de a további vizsgálatok is valószínűleg csak megerősítenék, hogy:

1. a hosszú távú (8-9 hónapnál hosszabb időszakra vonatkozó) előrejelzések tekintetében a piac egyértelműen pontosabb;
2. a rövid távú, de nem a közvetlen jövőre vonatkozó előrejelzések tekintetében (vagyis kb. 3-8 hónapos távlatban) az idősormodellek nyújtanak valamivel jobb teljesítményt;
3. a közeljövőre (1-2 hónap) vonatkozóan a piaci és az idősor-előrejelzések gyakorlatilag egyenértékűek.

KÖVETKEZTETÉSEK

A tanulmány egyik fő eredménye természetesen a tényleges inflációs előrejelzések készítése. A vizsgálatok azonban általános következtetések levonását is lehetővé teszik a SARIMA-modellek előrejelzési teljesítményével kapcsolatban.

Alapvető megfigyelés, hogy ezek a modellek számottevő hibával tudják csak leírni és előrejelezni a magyar inflációs idősorok alakulását, különösen hosszabb távon. Az előrejelzések pontatlanságát kiemeli az a tény is, hogy néhány inflációs részaggregátum eseté-

ben még a naiv extrapolációs módszerek is jobb illeszkedést jeleznek. A teljes fogyasztói árindex és a részidősorok többsége esetében azonban a SARIMA-modellek bizonyulnak pontosabb előrejelzőnek. Tehát az infláció múltbeli alakulása általában tartalmaz olyan extra információt a jövőbeli inflációra nézve, melyet a naiv modellek nem használnak ki. Ez az információ azonban legfeljebb 0,2-0,3 százalékpontos javulást eredményez az illeszkedési statisztikákban.

A piaci szereplők előrejelzéseivel való összevetésben a vizsgált idősormodellek közép- és rövid távon (1-8 hónap) teljes mértékben versenyképesnek mutatkoznak. Sőt, 3-7 hónapos előrejelzési időszakot tekintve az idősormodellek akár 0,5-0,6 százalékponttal alacsonyabb átlagos hibával is képesek az inflációs rátát előrejelezni. Hosszú távon azonban tagadhatatlan a piaci előrejelzések fölénye.

A szezonális egységgyöktesztek eredményeinek megfelelően transzformált idősorokra illesztett SARMA-modellek nem hoznak látványos és egyértelmű javulást az előrejelzési teljesítményben. A hagyományos Box–Jenkins-metodológia által alkalmazott szezonális differenciálás elméletileg nem mindig megfelelő ugyan, de ez a tény a legtöbb esetben nem befolyásolja hátrányosan a SARIMA-modellek előrejelzési pontosságát.

A fogyasztói árindex legmegbízhatóbb becsléseit az egyes részaggregátumok egyedi előrejelzéseiből képzett súlyozott átlagok adják. Ez az eredmény korántsem közismert, hiszen a CPI egyes komponenseit (például ÉLELM vagy ENERGI) igen nagy hibával lehet csak előrejelezni.

A modellek felállításakor természetesen számos módszertani probléma adódik, de ezek több-kevesebb nehézséggel áthidalhatók. Egy strukturális makromodell felállítása során felmerülő problémák sokkal bonyolultabbak lennének, és megoldásuk sokkal több időt kívánná. Az előrejelzési pontosság és a szellemi ráfordítás tekintetében a SARIMA-modellek mindenképpen esélyeseek a győzelemre. A SARIMA-modellekből származó előrejelzések összehasonlítási alapként szolgálhatnak, vagy más módszerekkel kapott értékekkel kombinálhatók. Éppen ezért ezeket a modelleket még akkor is érdemes előállítani, ha a kutató sokkal kifinomultabb előrejelző-apparátus alkalmazását fontolgatja.

IRODALOM

- [1] Box, G. E. P. – Jenkins, G. M.: Time series analysis; Forecasting and control. Holden Day. San Francisco. 1970.
- [2] Franses, P. H.: Periodicity and stochastic trends in economic time series. Oxford University Press. Oxford. 1996. 230 old.
- [3] Franses, P. H.: Time series models for business and economic forecasting. Cambridge University Press. Cambridge. 1998.
- [4] Gordon, D. V.: Optimal lag length in estimating Dickey-Fuller statistics: an empirical note. *Applied Economics Letters*. 1995. évi 2. sz. 188–190. old.
- [5] Granger, C. W. J. – Newbold, P.: Forecasting economic time series. Academic Press. New York. 1986.
- [6] Hamilton, J. D.: Time series analysis. Princeton University Press. Princeton. 1994. 799 old.
- [7] Hylleberg, S. – Engle, R. F. – Granger, C. W. J. – Yoo, B. S.: Seasonal integration and cointegration. *Journal of Econometrics*. 1990. évi 215–238. old.
- [8] Kennedy, P.: A guide to econometrics. MIT Press. 1992. 410 old.
- [9] Pindyck, R. S. – Rubinfeld, D. L.: Econometric models and economic forecasts. McGraw-Hill. New York. 1991. 596 old.
- [10] Vincze János – Zsoldos István: A fogyasztói árak struktúrája, szintje és alakulása Magyarországon 1991–1996-ban. Magyar Nemzeti Bank. Füzetek 5. Budapest. 1996.
- [11] Zsoldos István: Kimagvazott infláció. *Figyelő*. 1998. február 26. 32–35. old.

TÁRGYSZÓ: Infláció. Idősormodellek. Előrejelzés.

SUMMARY

The aim of the paper is to forecast the rate of inflation in Hungary using simple SARIMA time series models. The applied methodology accounts for nonstationary behaviour, seasonality and outlier observations: these are the features that seem to dominate the inflationary process in Hungary in the 90s. The presented forecasts are for the second half of 1999.

Although *ex post* forecasts show that there is a great deal of uncertainty associated with the predictions of the time series models, most of the SARIMA based forecasts of the CPI outperform „naive” forecasting methods. In the short run, past inflation does seem to contain information about future inflation beyond what is assumed by the naive models.

In comparison with the consensus forecast of the market participants (commercial banks, brokerage firms, etc.), these time series models prove to be perfectly competitive in the short run. In the long run, however, the performance of the market is obviously better.