

Munkapiaci áramlások, konzisztencia és gereblyezés*

Cseres-Gergely Zsombor,
az MTA Közgazdaságtudományi Intézet tudományos munkatársa és a Budapesti Szakpolitikai Elemző Intézet partnere
E-mail: zgergely@mtakti.hu

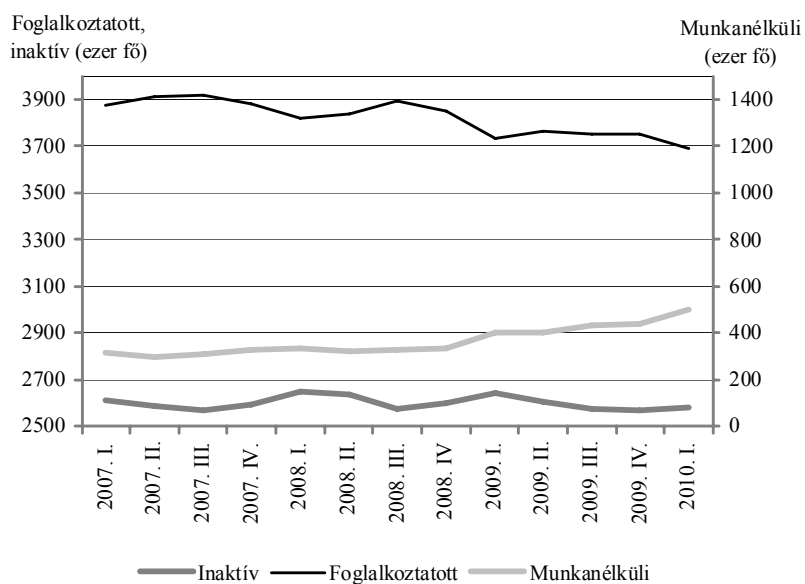
Az aktuális foglalkoztatási helyzet magyarországi elemzései szinte kizárólag a foglalkoztatottak, a munkanélküliek és az inaktívak számának változását vizsgálják, de nem foglalkoznak érdemben a közöttük megfigyelhető áramlásokkal. A szerző írásában az állományok közötti átmenetek számításának lehetséges konzisztens módszerét, valamint ennek segítségével a 2008 után kibontakozott válság foglalkoztatási hatásait ismerteti, elválasztva a támogatott és nem támogatott foglalkoztatás alakulását. A bemutatott módszer a napi elemzői, kutatói gyakorlatban alkalmazható és olyan többlet-információval szolgál, ami az állományok elemzésével nem szerezhető meg

TÁRGYSZÓ:
Munkaerőpiac.

* A tanulmány elkészítését a Foglalkoztatási és Szociális Hivatalnak az MTA Közgazdaságtudományi Intézetével kötött kutatási megállapodása segítette elő. A megalapozó munkálatokat 2007 nyarán, a Magyar Nemzeti Bankban végeztem vendégkutatóként. A felhasznált adatokat az MTA KTI Adatbankjának munkatársai tették könnyen használhatóvá, a számítások egy részét *Szabó Zsuzsa* végezte, és e téren segítséget nyújtott még *Szabó-Morvai Ágnes* is. Ezúton köszönöm e támogatásukat, valamint a Bureau of Labor Statistics (BLS – az Egyesült Államok Munkaügyi Statisztikai Hivatalának) munkatársai, *Ed Robinson* és *Greg Erkens* konzultációs készségét; továbbá *Kátay Gábor* korai kritikáját, *Molnár György* és *Nagy Gyula* segítő megjegyzéseit. Külön köszönöm a lektor, *Mihályffy László* megjegyzéseit és javaslatait. A tévedések és hibák természetesen csak engem terhelnek.

Magyarország 15–64 éves népességének foglalkoztatási rátája 1998 és 2008 között 53,6 százalékról 56,7 százalékra nőtt, elsősorban a nyugdíjkorhatár emelésének és a népesség javuló iskolázottságának köszönhetően (Kátay–Nobilis [2009]). A 2008 második felében kibontakozott világgazdasági válság hatására jelentős visszaesés volt megfigyelhető e mutatót tekintve: 2009-ben már csak 55,4 százalék volt és innen azóta sem mozdult el érdemben. Mindeközben a munkanélküliségi ráta 2008-ról 2009-re éves szinten 2,2 százalékponttal, közel 25 százalékkal nőtt, értéke 2010 első negyedévében 11,2 százalék volt (KSH stADAT, www.ksh.hu). A változás abszolút számokban mérve is jelentős: a munkanélküliek száma mintegy 330 ezerről félmillióra nőtt. (Lásd az 1. ábrát.) A regisztrált munkanélküliek aránya a 15–64 éves aktív népességben 2008 áprilisában 10,7, 2009-ban 13,7, míg 2010-ben 14,4 százalék volt ugyanebben a hónapban. A munkanélküliség jelentős növekedése ellenére kedvező jel, hogy az aktivitás nem csökkent az érintett időszakban.

1. ábra. A fő munkapiaci állományok alakulása 2007 és 2010 első negyedéve között (15–64 éves népesség)



Forrás: KSH munkaerő-felmérések.

Az előbbiekhöz hasonló elemzést több szakmai műhely ad közre, és ezekből egyértelmű, hogy a munkapiacra rendkívüli és válságos állapot alakult ki. De világosak-

e a részletek is? Ebben az írásban a munkapiaci állományok közötti áramlások segítségével kísérünk meg választ adni a kérdésre.

Az áramlások vizsgálatában kiemelt szerepet tölt be egy egyszerű, de Magyarországon e célra korábban nem használt eszköz, az ún. „gereblyezés”. Az egyéni paneladatokból számított átmenetbecslések a használt adatbázisok mintájának nemvéletlen kopása miatt torzítottak, azokat korrigálni kell. Ezt *Abowd–Zellner* [1985] módszertanilag igen kifinomult módon oldotta meg, de az Egyesült Államok Munkaügyi Statisztikai Hivatala (Bureau of Labor Statistics – BLS) csak tíz év múltán egy jóval egyszerűbb módszerrel, a gereblyézéssel kezdte meg a probléma kezelését *Frazis és szerzőtársai* [2005] alapján. Hazánkban az áramlások becslésére strukturális modellezés helyett imputálást használó kísérleti javaslatot *Fraller és szerzőtársai* [2007] tettek, bármilyen módszer rendszerszerű alkalmazására pedig még várni kell. A BLS eljárása azon emberek viselkedését vizsgálja, akik időben valóban követhetők, így merőben eltér *Abowd–Zellner* [1985] megközelítésétől, amely a tárgyidőszakban már nem megfigyelhető emberekre vonatkozó adatokat is használja, lehetséges státusukat megbecsüli. A jelen írás célja ebben a szellemben olyan eszköz bevezetése, amely kellően egyszerű a napi rutinban való használathoz, mégis alkalmas az áramlások konzisztens becslésére. A számításokat a KSH munkaerő-felmérésének (MEF) adatai segítségével végeztük, amelyek időben egymás után következő értékei (anonim kóddal) egyéni szinten összekapcsolhatók.

A továbbiakban bemutatjuk az állományok közötti átmenetek számításának problémáját, majd ismertetjük a gereblyezés által kínált megoldást, az alkalmazott módszerrel kapott eredményt és a számított áramlásokat.

1. Állományok és áramlások

A bevezetőben láthattuk, hogy az alapvető munkapiaci állományok vizsgálata a foglalkoztatás csökkenését és a munkanélküliség növekedését mutatja a 2007 utáni időszakban, ebből azonban még nem világos, hogy mindez miként történt. 2008 negyedik és 2009 első negyedéve között a foglalkoztatottak állománya 114 ezer fővel csökkent, a munkanélkülieké és az inaktívoké pedig 66, illetve 42 ezer fővel nőtt (a különbséget a 15–64 évesek számának enyhe mérséklődése felelős. (Lásd az 1. táblázatot.)

Az eddigiekben a megszokott elnevezést használva *állománynak* neveztük azon emberek számát, akik adott időpontban egy meghatározott állapotban találhatók. Ezzel szemben *áramlásnak* nevezzük azon emberek számát, akik két egymást követő (bázis- és tárgy-) időpontban két meghatározott állapotban találhatók. Más szóval az áramlás a két időszak közötti *állapotváltozások* száma. Az angol nyelvű szakiroda-

lom az áramlás két formáját különbözteti meg, a *bruttó* és a *nettó* áramlást. A bruttó áramlás a két állapot közötti áramlás, ha az áramlás irányát számon tartjuk, a nettó áramlás pedig a két állapot közötti bruttó áramlások egyenlege. Mivel a mintabeli összegek hányadosa a népességi arányok konzisztens becslését adja, a bruttó áramlások és a bázisidőszaki állományok aránya a kiválasztott állapotok közötti *átmenet esélyének* becslése. Bár azok más időszakok között is meghatározhatók, ebben az írásban csak két, egymást követő negyedév közötti áramlásokkal foglalkozunk.

1. táblázat

Munkapiaci állományok 2009 negyedik és 2010 első negyedévében (15–64 éves népesség)
(fő)

| Munkapiaci státus | 2009. IV. negyedév | 2010. I. negyedév | Változás |
|---------------------|--------------------|-------------------|---------------|
| Foglalkoztatott | 3 850 688 | 3 736 019 | -114 669 |
| Munkanélküli | 336 633 | 402800 | 66 167 |
| Inaktív | 2 602 477 | 2 644 160 | 41 683 |
| <i>15–64 évesek</i> | <i>6 789 798</i> | <i>6 782 979</i> | <i>-6 819</i> |

Forrás: Itt és a további táblázatoknál is saját számítás a KSH MEF mikroadataiból, egészre kerekített keresztmetszeti súlyokkal.

Az áramlások pontos ismeretének jelentősége nagy. Például, ha a munkanélküli-ráta 12 százalék, igen sokat számít, hogy milyen mértékű a munkanélküli és a foglalkoztatott állapot közötti kapcsolat, mennyi az esélye annak, hogy egy dolgozó munkanélkülivé válik és mennyi annak, hogy éppen ellenkezőleg, egy munkanélküli állást kap. A munkanélküliség magas szintje mindenképpen aggasztó, de egészen más szakpolitikai beavatkozást igényel egy nagy fluktuációval jellemezhető, foglalkoztatásba jó eséllyel átlépő munkanélküli-állomány, mint az olyan, ahol kicsi a cserélődés. Nem véletlen, hogy az áramlások vizsgálata kiemelt helyen szerepel *Elsby–Hobijn–Sahin* [2010] tanulmányában is, amely a válságnak az Egyesült Államok munkapiacára gyakorolt hatását elemzi.

Az állományok nagysága kiszámolható az egy időszakra vonatkozó reprezentatív adatfelvételekből, az áramlások azonban nem. Ehhez nem pusztán két keresztmetszeti mintára van szükségünk, de a két időszakra vonatkozó mintának azonos személyeket is kell tartalmaznia. Szerencsére a MEF erre is lehetőséget biztosít a mintegy 70 ezer megfigyelést tartalmazó keresztmetszeti minta közel 5/6-át kitevő részmintáján, amelynek résztvevői és a rájuk vonatkozó adatok két negyedév között összekapcsolhatók.

Ésszerűnek tűnik az áramlásokat úgy számszerűsíteni, hogy a felvétel negyedévek között összekapcsolható almintáját használva megszámláljuk azon embereket, akik a

kiválasztott állapotpárokból voltak találhatóak a kiválasztott időpontokban. A 2. táblázatban a MEF paneladatainak a felhasználásával készített számítást látunk. A táblázat világosszürke blokkja azt mutatja meg, hogy az egyes állapotok időbeni kombinációja, tehát a köztük való átmenet hány ember esetében volt megfigyelhető. A korrekciót nem alkalmazó számítások csak erre az információra támaszkodnak. A konkrét számokat, az áramlásoknak a teljes népességre vetített becslését a mikroadatokból az adott cellában található emberek egyszerű megszámlálásával kapjuk meg, ezeket azután először a tárgyidőszakhoz tartozó keresztmetszeti súlyokkal szorzom fel – ezek a súlyok eredetileg egyrészt a teljeskörűsítést, másrészt a keresztmetszeti demográfiai reprezentativitást hivatottak biztosítani a teljes mintára. Mivel azonban itt a teljes mintánál kevesebb emberrel dolgozunk, minden cella értékét fel kell még szorozni a keresztmetszeti és a panelmintákhoz tartozó súlyok összegének hányadosával is annak érdekében, hogy csak a panel tagjait használva a teljes népességre vonatkozó összeseneket kapjunk.¹ A számítás elvégzésének tehát közvetlen akadálya nincs, azonban az mégsem pontos; akkor lenne csak az, ha a kapott áramlások megfelelnek egy igen egyszerű konzisztenciafeltételnek: az állományok változását rendre megkaphatnánk a nettó áramlások összegeként. A valóságban azonban nem ez a helyzet, mint azt a 2. táblázat „Változás becslés átmenetekből” feliratú oszlopa mutatja (ellenőrzés céljából mellette szerepel a keresztmetszeti adatokból készített, az 1. táblázatban már ismertetett becslés ugyanerre a változásra „Változás becslése keresztmetszetből” fejróval, valamint a két becslés különbsége „Eltérés” fejróval). Látható, hogy a direkt és a nettó változásokból számított indirekt becslések jelentős eltérést mutatnak. Az inaktivitás esetében a különbség elhanyagolható, a munkanélküliség növekedését és a foglalkoztatottság csökkenését azonban a naiv becslés alapján mintegy harmadával alulbecsülnénk. Ez nemcsak arányaiban nagy szám, hanem az ilyen jelenségek elemzését felhasználó szakpolitika szempontjából is, hiszen megegyezik például a 2009-ben a „Munkahelyek megőrzéséért” elnevezésű program által érintett személyek számával (mintegy 30 ezer fő), és két fontos aktív munkaerő-piaci program, a START Plusz és a START Extra résztvevőinek együttes számával 2009-ben (mintegy 28 ezer fő) (FSZH [2010]).

A bruttó áramlások számításának természetesen nem célja az állományi változások visszszámítása, hiszen amíg ezeket lehetséges direkt módon becsülni, addig az indirekt út választása teljesen értelmetlen. A változások helytelen becslése azonban rámutat arra, hogy maguk az áramlások becslései sem lehetnek pontosak – a hiba mértékét akkor számszerűsítjük majd, ha a korrigált becslés rendelkezésre áll. Mivel adott állomány nagyság mellett az átmentek száma és az átlépési esélyek között egyszerű és közvetlen kapcsolat van, az érintett emberek számának helytelen becslésén túl a hiba egyben azt is jelenti, ez utóbbiakat is rosszul becsüljük. Megjegyzendő, hogy a számszerű eltérések olyan mértékűek és jellegűek is lehetnek, hogy az áramlásokból származtatott

¹ A kétféleképpen kapott népességszám 6 782 979 és 6 782 585, a 394 fős eltérést a súlyok kerekítése okozza.

változásoknak nemcsak a nagyságát, de előjelét is helytelenül becsülnék meg. Példa erre 2008 második és harmadik negyedéve között a munkanélküliség változása, amely jobban jellemzi a nagy változások közötti „békeidőt”. Itt a hiba következtében a több mint 8 ezer fős növekedése helyett 7 ezer fős *csökkenést* becsülnék. A hiba jelenléte tehát nyilvánvaló – a következőkben megvizsgáljuk, honnan ered.

2. táblázat

*Munkapiaci állományok közötti átmenetek száma, naiv és hibás becslés esetén,
2009 negyedik és 2010 első negyedéve között
(fő)*

| Megnevezés | | Tárgydőszak | | | Változás- becslés átmenetekből | Változás becslé- se keresztmet- szetből | Eltérés |
|-------------------|----------------------------------------------------------|----------------------|-------------------|------------------|--------------------------------------|-----------------------------------------------|----------|
| | | Foglalkozta- tott | Munkanél- küli | Inaktív | | | |
| Bázis- időszak | Foglalkoztatott | 3 673 955 | 75 025 | 77 399 | -81 787 | -114 669 | 32 882 |
| | Munkanélküli | 36 044 | 280 989 | 22 830 | 41 922 | 66 167 | -24 245 |
| | Inaktív | 34 593 | 25 771 | 2 505 195 | 39 865 | 41 683 | -1 818 |
| | <i>Tárgydőszaki állo- mány az áramlá- sokból</i> | <i>3 744 592</i> | <i>381 785</i> | <i>2 605 424</i> | <i>0</i> | | <i>0</i> |
| | Tárgydőszaki állo- mány keresztmet- szetből | 3 736 019 | 402 800 | 2 644 160 | -6 819 | | |
| | Eltérés | 8 573 | -21 015 | -38 736 | | | |
| | | | | | | | |

2. Az áramlások becslésének kérdései

A különböző állapotok közötti átmenetekre irányuló vizsgálatok tartalmi célja az egyénre koncentrálna a következőképpen írható le a legegyszerűbb esetben, ha csak két állapot lehetséges. A t időpontban jelöljük I_t -vel a szereplő állapotát, aki az egymást kizáró A vagy B állapotban lehet – ennek megfelelően I_t értéke az $\{A, B\}$ halmazból adódik. A két állapotban levő szereplők számának összege megegyezik a népesség nagyságával, tehát $\#(I_t = A_t) + \#(I_t = B_t) = N$, ahol $\#()$ azt a függvényt jelöli, amely az argumentumban jelzett esemény gyakoriságát adja vissza.

Célunk megbecsülni a $t_{XY} = \Pr(I_t = X | I_{t-1} = Y)$ valószínűséget, tehát azt, hogy ha a szereplő a $t-1$ -edik időpontban az $X \in \{A, B\}$ állapotban van, akkor milyen eséllyel lesz az $Y \in \{A, B\}$ állapotban t időpontban. Konkrét példával élve: mennyi az esélye annak, hogy valaki „ma munkanélküli, ha tegnap dolgozott”. Ez a valószínűség az analógia elvére hagyatkozva konzisztensen becsülhető, ha a két állapot és a két időszak között lezajlott átmenetek számát elosztjuk a kiinduló népesség számával, így
$$\Pr(I_t = B | I_{t-1} = A) = \#(I_t = B | I_{t-1} = A) / \#(I_{t-1} = A).$$

Az átmenetek becsléséhez szokatlanul jó minőségű adatbázisra van szükség: olyanra, amelyben ugyanazok a szereplők találhatók meg mindkét vizsgált időpontban, és amely nemcsak mindkét időpontban reprezentatív mintája a teljes vizsgált népességnek, de az átmenetekre nézve is az. Ma már számos olyan adatbázis létezik, amelyben az első feltétel teljesül, ilyen a MEF is. Ebben ún. rotációs panelt alakítottak ki, amelyben a minta körülbelül egyhatoda minden negyedévben cserélődik, így az egyének időben követhetők, de csak meghatározott ideig (a cél eredetileg a keresztmetszeti minták folyamatos frissítése volt). Gond sokkal inkább a második feltétellel akad, és éppen a rotáció miatt. Bár az adatfelvételnél igyekeznek ügyelni arra, hogy a kilépő és belépő emberek összetétele hasonló legyen, ez csak bizonyos korlátok között biztosítható. A panelben maradó minta ennek eredményeként nem véletlenszerűen alakul ki, ezért a keresztmetszeti minta megoszlásait nem képes visszaadni, az átmenetekre vonatkozóan pedig a 2. táblázatban közöltekhez hasonló, közvetett következtetéseket tudunk levonni. Eszerint a megvalósult becslés nem helyes, mert a $\Pr(I_t = X | I_{t-1} = Y)$ valószínűséget csak a panelbe került emberek mintájának felhasználásával tudjuk becsülni. Mivel a panelminta A és B állapotok szerinti megoszlása jelentősen eltérhet mind a t , mind pedig a $t-1$ időszak teljes megvalósult mintájának megoszlásától, ez alapján más eredményt kapunk, mint amire a keresztmetszeti minta összes résztvevőjének megfigyelése által jutnánk.

Az itt bemutatott átmenetek becslését megnehezítő jelenség általános probléma, a nem elhanyagolható nemválaszolás (nonignorable nonresponse) Rubin [1976] által tárgyalt alosete, rokon a szelekciónak a munkagazdaságtanon belül a munkakínálattal összefüggésben Heckman [1979] óta sokat vizsgált esetével. Mindezzel együtt most a szelekció strukturális modellezésével nem fogunk foglalkozni, ahogyan a mérést befolyásoló hiba hatásával sem.

Meg kell jegyeznünk, hogy a tárgyalt probléma nemcsak az egyszerű pontbecsléseket, de az összetettebb regressziós eredményeket is érinti. A többváltozós becslések korrekciójának szükségességét könnyen beláthatjuk, ha egy igen egyszerű lineáris regressziót tekintünk, amelynek függő változója az átmenet tényét jelzi. Legyen ebben a konstans mellett az egyetlen magyarázóváltozó a nemi identitást jelző indikátor (értékei: 0 = férfi, 1 = nő). Egyszerű számítások után látható, hogy egy, az átmenet valószí-

nőségére vonatkozó lineáris valószínűségi modellben a konstans a férfiak, a konstans és az indikátorváltozóhoz tartozó együttható összege pedig a nők esetében adja az átlagos átmenet-valószínűség becslését. Ezek a számok megegyeznek a közvetlen számítással kapott értékekkel, és ennek megfelelően maguk is inkonzisztensek – ez a helyzet pedig lényegileg sem további kontrollváltozók bevonásával, sem akkor nem változik, ha a lineáris valószínűségi modell helyett nemlineáris (például probit) modellt alkalmazunk. A többváltozós modellek alkalmazása és korrekciója érdekes és fontos terület, amellyel azonban összetettsége miatt itt nem foglalkozunk.

3. Megoldási lehetőségek elméletben és gyakorlatban

A munkapiaci áramlások helyes becslése a naiv eljárás valamilyen korrekcióját igényli. A korrekció lehetséges formáját a következőkben *Fagan–Greenberg* [1988] megközelítését ismertetve egy általános statisztikai probléma, a „táblák additív tételének” speciális eseteként mutatjuk be. Az additív tétel igénye akkor merül fel, amikor egy $R \times C$ dimenziójú táblázat (mátrix) a_{ij} elemeit úgy kell módosítanunk, hogy azokat a sorok és az oszlopok irányában összegezve is egy külsődlegesen adott, \mathbf{r} és \mathbf{c} referencia-számsorral (vektorral) megegyező számsort kapjunk. Mindezt úgy kell megtennünk, hogy a módosítás után kapott tábla elemei valamilyen jól meghatározott szempont szerint a lehető legkevésbé térjenek el a kiindulásként használt, a nyers adatokból számított átmenetmátrixtól. Formális értelemben a feladat az, hogy egy ún. *megvalósítható A* kontingenciatáblából egy származtatott *B* táblát nyerjünk, egyidejűleg az elemeik között fennálló különbséget adott szempont és feltételek szerint minimalizáljuk. Ehhez *Fagan–Greenberg* [1988] bevezet néhány alapfogalmat.

Kontingenciatáblának nevezzük a nemnegatív valós számok azon tömbjét, amelyet az $\mathbf{A} = \{(a_{ij}), \mathbf{r}, \mathbf{c}\}$ hármas határoz meg. Ebben (a_{ij}) egy $R \times C$ elemű mátrix, $\mathbf{r} = (r_1, r_2, \dots, r_R)$ és $\mathbf{c} = (c_1, c_2, \dots, c_C)$ pedig vektorok, melyek elemeire igaz, hogy összegük megegyező, azaz

$$\sum_{i=1}^R r_i = \sum_{j=1}^C c_j .$$

A *B* tábla *additív*, ha elemeinek egy adott dimenzió nagyságának megfelelő vektor irányába vett összege az adott vektor megfelelő elemével megegyezik, azaz

$$\sum_{j=1}^C b_{ij} = r_i \quad i = 1, 2, \dots, R \quad \text{és} \quad \sum_{i=1}^R b_{ij} = c_j \quad j = 1, 2, \dots, C .$$

Az \mathbf{A} tábla *megvalósítható*, ha létezik olyan (b_{ij}) $R \times C$ mátrix amelyre igaz, hogy az \mathbf{A} -ból *levezetett* $\mathbf{B} = \{(b_{ij}), \mathbf{r}, \mathbf{c}\}$ kontingenciatábla additív, valamint $b_{ij} = 0$, akkor és csak akkor, ha $a_{ij} = 0$. A \mathbf{B} táblának tehát \mathbf{A} -val megegyező számú sora és oszlopa van, megfelelő elemei nulla értéket vesznek fel akkor, ha az \mathbf{A} elemei nulla értékűek, az \mathbf{A} -val ellentétben azonban igaznak kell lennie rá, hogy sorainak, illetve oszlopainak összege megegyezik az \mathbf{r} és \mathbf{c} vektorok megfelelő elemeivel. Az eljárással tehát olyan táblát szeretnénk kapni, amely kedvezőbb tulajdonságokkal bír, mint a kiindulásként használt, mégis megtartja annak bizonyos fontos tulajdonságait.

Adott megvalósítható $\mathbf{A} = \{(a_{ij}), \mathbf{r}, \mathbf{c}\}$ tábla esetében a következő a feladat: minimalizáljuk a

$$d(a_{ij}, b_{ij})$$

célfüggvényt a

$$\sum_{j=1}^C b_{ij} = r_i \quad i=1, 2, \dots, R \quad \text{és} \quad \sum_{i=1}^R b_{ij} = c_j \quad j=1, 2, \dots, C$$

korlátok mellett, ahol a $d(\cdot, \cdot)$ távolságfüggvény nem negatív, a változója szerint konvex, kétszer folytonosan differenciálható, és értéke csak akkor nulla, ha argumentumai egyenlők ($a_{ij} = b_{ij}$).

Fagan–Greenberg [1988] megmutatja, hogy a különféle módszereknek más és más célfüggvények feleltethetők meg, és bár eltérő eredményt hoznak, mindnek egyedi minimuma van (a szerzők a duális feladatot oldják meg, és azt találják, hogy minden célfüggvény Hesse mátrixa negatív definit). A bemutatott módszerek közül a maximum likelihood a tábla elemeinek igen szélsőséges eloszlását hozta, így annak alkalmazását a szerzők elvetették. A legkisebb négyzetekhez hasonló módszer pedig ugyan zárt alakban megoldható, de önmagában nem garantálja a tábla elemeinek nemnegativitását. Az ezeken felül megvizsgált minimum khi-négyzet és gereblyezés (*raking*, ami *ratio estimator* néven is ismert) módszere az adott feltételek mellett kiegyenlítettebb, nemnegatív (kedvezőtlen esetben hamis, az eredeti táblában nem jelentkező nulla) eredményt hoz, és valamilyen mértékig mindkettő megőrzi a tábla elemei közötti relációkat. Mindezek és egyszerűsége miatt választottam a gereblyézést, amihez a következő célfüggvény tartozik:

$$\sum_{i,j} d(b_{ij}, a_{ij}) = \sum_{i=1}^R \sum_{j=1}^C b_{ij} \ln \left(\frac{b_{ij}}{a_{ij}} \right).$$

A gereblyezés keresztmetszetben egyszerűen alkalmazható eljárás, de csak a feladat megfelelő megfogalmazása után használhatjuk a munkapiaci átmenetek igazítására. *Fagan–Greenberg* [1988] rámutat arra, hogy a korrekciós eljárások bármelyikét csak akkor szabad használni, ha más út nem járható, például nem lehetséges már kiegészítő információt bevonni. A munkapiaci áramlások esetében időbeli folyamattal van dolgunk, aminek hatása két területen is jelentkezik. Egyrészt a résztvevők korának előrehaladása, másrészt az adatfelvételnek a résztvevők követése terén korlátozott képessége által. Ennek figyelembe vétele nemcsak amiatt fontos, hogy a vizsgált rendszer logikailag zárt legyen, de azért is, hogy az alkalmazott korrekció feltétele, az A tábla additivitása teljesüljön.

A vizsgált két időszak között végbemenő munkapiaci átmenetekkel párhuzamosan a személyek számát növelő és csökkentő demográfiai tényezők hatása is érvényesül, amit valamilyen formában meg kell jelenítenünk. Egyrészt az adatfelvétel első időszakban megfigyelt résztvevőinek száma *csökken* azáltal, ahogy egy részük kora előre halad és ennek következtében egy életkor szerint korlátozott csoportból (például a 15–64 évesek köréből) éppen kilép; illetve kivándorol az országból; vagy meghal. Ezzel párhuzamosan vannak olyanok, a megfigyelendők számát *növelve*, akik éppen belépnek a vizsgált korcsoportba (vagy ha annak nincs alsó határa: megszületnek), illetve bevándorolnak az országba. Emellett az adatfelvétel sajátosságai is befolyásolják, hogy mennyire képes a népességben egyébként továbbra is megtalálható egyedeket követni. Mivel a KSH által használt ELAR-mintakeret elsődleges mintavételi egysége a lakás, az adott lakásból elköltöző személyek a felvétel számára annak ellenére elvesznek, hogy a népességben megtalálhatók. Ezek a személyek a halottakhoz hasonlóan, csak éppen más okra visszavezethetően csökkentik, ezzel szemben a mintakopás ellensúlyozására bevontak növelik a vizsgált minta nagyságát (az ő hatásuk a bevándorlókéhoz hasonlatos).

A demográfiai nemstacionaritás figyelembe vételére ambíciónk és a rendelkezésre álló adatok függvényében több lehetőségünk van. Mivel paneladatokkal dolgozunk, a korcsoportokba való be- és kilépésekre valamilyen becslést minden esetben tudunk adni, ezek pedig tartalmilag is fontos részei lesznek az elemzett rendszernek. Abban az esetben, ha a további tényezők differenciálásának nincs tartalmi indoka, dönthetünk akár úgy is, hogy az egyszerűség kedvéért a korcsoportból ki- és belépők kivételével minden különbséget egy mesterséges, a korrekció céljából létrehozott állapotban jelenítünk meg. A népesség két időszakban mért nagyságának különbsége egyetlen szám, ami két kérdést vet fel. Egyrészt mivel a különbség pozitív és negatív is lehet, az áramlásokat viszont nemnegatív számok jellemzik, a különbséget az előjeltől függően pozitív ki- vagy beáramlásként könyveljük el. Másrészt valamilyen szabályt kell alkalmaznunk arra, hogy a különbség „összesenjét” milyen módon osztjuk szét a másik időszaki állapotok között.

Amennyiben külső forrásból megbízható becslésünk van, például a (vizsgált állapotok szerint differenciált) halálozási rátákra, be- és kivándorlási adatokra, és ennek szükségét érezzük,² akkor ezeket explicit módon is megjeleníthetjük. *Frazis és szerzőtársai* [2005] például külső forrásból veszik át a halálozási arányokat, de nem jelenítik meg a ki- és bevándorlási arányokat. Az új állapotok az átmenetmátrixban új sorok és oszlopok formájában jelennek meg, amelyek többsége azonban nem vesz részt az igazításban – hiszen értékük külső forrásból adott. Szerepük az, hogy a referenciasor és -oszlop értékeit növelik vagy csökkentik, ezáltal biztosítják az **A** tábla additivitását. A korrekciót mindezek után, az említett kiegészítések eredményeként kialakult táblázaton végezzük el.

Noha a formális probléma célfüggvénye a kezdeti és a végeredményként adódó mátrixok hasonlóságára épül, a gyakorlatban alkalmazott algoritmus az áramlásokból származtatott és a megfigyelt állományok egyezését célozza. Az erre a célra itt alkalmazott eljárás az iteratív arányos illesztés (iterative proportional fitting – IPF) – lásd erről például *Deming–Stephan* [1940] korai írását – amelyet az input-output modellekre alkalmazva *Parikh* [1979] után RAS-ként is ismerhetünk.³ A módszer azt feltételezi, hogy az **A** megvalósítható táblát már előkészítettük, beleértve a referenciavektorok, valamint az áramlástábla kiszámítását és az előbbieken tárgyalt korrekcióját. Miután ez megvan, az iteratív arányos illesztés módszerét a következőképpen alkalmazzuk (tegyük fel, hogy a tábla sorai a bázisidőszak állapotait jelzik):

1. Sorok szerint haladva elsőként a bázisidőszak szerint igazítunk. Kiszámítjuk az első sor elemeinek összegét, majd vesszük ennek és a referenciavektor adott sorhoz tartozó elemének hányadosát. Ezzel a számmal osztjuk a sor minden elemét, amelyek összege ennek hatására megegyezik a referenciavektor megfelelő elemével. Ugyanezt meg tesszük a következő és mindegyik sorral.

2. Miután minden sorral végeztünk, oszlopok szerint kezdünk el haladni, így a tárgyidőszak sarokszámai szerint fogunk igazítani. Kiszámítjuk az első oszlop elemeinek összegét, majd vesszük ennek és a referenciavektor adott oszlophoz tartozó elemének hányadosát. Ezzel a számmal osztjuk az oszlop minden elemét, amelyek összege ennek hatására megegyezik a referenciavektor megfelelő elemével. Ugyanezt meg tesszük a következő és mindegyik oszloppal.

² Az ettől való eltérésre az lehet a motivációnk, hogy ezek a korrekciós tényezők igen nagyok lehetnek.

³ A RAS elnevezés a látszat ellenére nem rövidítés. *Molnár György*, a módszer egy hazai alkalmazója a következőképpen magyarázza kialakulását: „Az input-output táblákat tradicionálisan az **A** mátrix jelöli és a módszer alaplépése a következő transzformációt hajtja végre: $\mathbf{B} = \mathbf{rAs}$, amit gyakorta írtak $\mathbf{B} = \mathbf{RAS}$ formában, ahol **R** és **S** diagonális mátrixok. Ebből a bevett jelölésből alakult ki a név.”

3. Kerekítjük a tábla elemeit, hogy csak egész számokkal dolgozzunk. Vesszük a referenciaként használt és az átmenetekből számolt összesenek különbségének négyzetösszegét. Ha ez kisebb, mint egy korábban alkalmasan megválasztott határérték, vége az igazításnak; ha nagyobb, visszatérünk az 1. ponthoz.

Az eljárásnak megfeleltethető a korábban már bemutatott célfüggvény, amit *Fagan és Greenberg* [1988] vizsgál. Az általuk javasolt iteratív és az itt bemutatott módszer közötti azonosságot többek között *Frazis és szerzőtársai* [2005] cikke is felhasználja. Mindezek miatt a bemutatott eljárással véges számú lépésben megtalálható a probléma egyedi optimuma.

Bár a felsorolt lépések emlékeztethetnek az egyszerű, például az egyéni jellemzők szerint egyéni súlyokat igazító módszerhez – ahol a megfigyelt és a referenciaként adott számok hányadosaként egy lépésben kialakulnak a kívánt súlyok –, iteratív jellege miatt eltér azoktól. Magyarországon is a gereblyezéshez hasonló eljárást használnak a statisztikában: így készülnek többek között a MEF és a Háztartási Költségvetési Felvétel súlyai is, csak több dimenzió használatával és nem több időpontra vonatkozó referenciát használva. Ezt a módszert alkalmazzák a felvételt használó kutatók illetve tanulmányok, például *Mihályffy* [1995], valamint a módszert átvevő és módosító *Molnár* [2005]. A megoldandó feladatot ezekben az esetekben az jelenti, hogy a súlyokat háztartásokhoz rendelik, míg az illeszkedés peremfeltételeit egyéni jellemzők adják.

4. A munkapiaci állapotok korrekciójára választott módszer és alkalmazása

Az egyszerű munkapiaci állapotok korrekciójára a *Frazis és szerzőtársai* [2005] által alkalmazott eseménytér némileg egyszerűsített formáját választottuk, amit a 3. táblázat mutat be, és amelyre a gereblyezés módszerét alkalmaztuk. A táblázatban és később is rövidítéseket használtunk, ahol F a foglalkoztatott, M a munkanélküli és I az inaktív státust jelöli, ezek kombinációja pedig az átmeneteket és az áramlás nagyságát egyszerűre. Az FM jelölés ennek megfelelően a foglalkoztatásból munkanélküliségbe áramlásra utal. Látható, hogy az áramlások vastag kerettel kiemelt 3×3 -as mátrixa mellett további állapotként jelenik meg a korcsoportba beáramló és az onnan kiáramló népesség nagysága. Az ez után fennmaradó népességszám-eltérést az egyéb be- és kiáramlás második oszlopa, illetve sora jeleníti meg pozitív számként aszerint, hogy a népesség száma nő (beáramlás), vagy csökken (kiáramlás). Az arányokat a kapcsolódó

időszak (rendre: tárgy- és bázis-időszak) állományi arányainak megfelelően osztjuk szét az állapotok között, amivel egyfajta átlagos viselkedést szándékozunk megjeleníteni. Két korlát szerepel a mátrixban: egyrészt nincs olyan 15 éves kort elhagyó személy, aki éppen 65 éves lesz, és olyan sem, aki „egyéb beáramló” státusából azonnal „egyéb kiáramló” lesz. Ezekkel a korrekciókkal a kibővített áramlási mátrix már konzisztens a demográfiai ki- és belépésekkel bővített referencia-állományokkal.

A mátrixot a MEF korábban kiválasztott két hullámának, 2009 negyedik, és 2010 első negyedévének egyéni adatait felhasználva töltöttük fel számokkal. A kapcsolt adatbázisban a második időszakban 15–64 éveseken kívül azokat is megtartottuk, akik a második időszakban 65 évesek, de a korábbi időszakban 64 évesek, illetve azokat, akik az első időszakban 14 évesek voltak, de a másodikban már 15 évesek voltak. A tárgyidőszakot véve alapul, ennek keresztmetszeti súlyszámait használjuk kiindulásként, majd ennek a keresztmetszeteiből becsült összlétszámához igazítjuk a súlyokat. Elsőként kiszámoljuk az életkort váltók számát, amihez – tekintve, hogy külső felhasználók számára a keresztmetszeti adatbázisban napra pontos életkor nem érhető el – már a paneladatokat kell használni. A népességszámnak a bázisidőszakhoz képest tapasztalt eltérését a korábban leírtaknak megfelelően az egyéb ki- és beáramlás állapottal korrigáljuk.

3. táblázat

*A főbb munkapiaci állapotok közötti átmenetek sémája:
a KSH munkaerő-felmérés felhasználásával történő számításhoz használt eseménytér*

| Megnevezés | | | Tárgyidőszak | | | | | Állomány |
|-------------------|----------------------|-----------|-------------------|-----------|------|-----------------|----------|----------|
| | | | Munkapiaci státus | | | Egyéb kiáramlás | | |
| | | | F | M | I | 65 lett | Egyéb ki | |
| Bázis- időszak | Munkapiaci státus | F | FF | FM | FI | FL65 | Fek | F |
| | | M | MF | MM | MI | ML65 | MEk | M |
| | | I | IF | IM | II | IL65 | IEk | I |
| | Egyéb beáramlás | 15 volt | V15F | V15M | V15I | 0 | V15Ek | V15 |
| | | Egyéb be | EbF | EbM | EbI | EbL65 | 0 | Eb |
| Állomány | | F' | M' | I' | L65 | Ek | | |

Az így kialakított táblázatra alkalmaztuk a gereblyezés módszerét, az állományi referencia- és az áramlásokból következő vektorok különbségének négyzetösszegét minimalizálva, megállási szabályként 0-hoz közeli pozitív számot használva. Az iterációk számát 500-ban korlátoztuk, amit a rendszer a gyakorlatban egyszer ért csak el, de a hiba akkor is mérsékelt maradt. Az eredményt a 4. táblázat mutatja be.

Az illeszkedés vizsgálatok azt várjuk, hogy az adott állapotból induló-, illetve az oda tartó áramlások összege és a demográfiai változással együtt adódó, származtatott állomány az egyéb áramlásokkal korrigált referenciaállománnyal egyezzen meg. A táblázat e két állományt mutató soraiból és oszlopaiból látható, hogy az illeszkedés a foglalkoztatás esetében tapasztalt néhány százast – a multiplikatív igazításból és a kerekítésből adódó – eltéréstől eltekintve tökéletes. Elsőként érdemes a vizsgálat fő tárgykörébe nem tartozó áramlásokat szemügyre venni. Tízezres nagyságrendű áramlásokról lévén szó, nem hanyagolható el az a tény, hogy a korcsoportba és -ből történő áramlások nagyságrendje is hasonló: mintegy 20 ezer fő volt a vizsgált negyedévek között. A népesség időbeli különbségét mégsem ez adja, hanem a máshova nem sorolt, mintegy 7 és fél ezres „egyéb” csoport, amelynek több, mint felét a foglalkoztatáshoz rendeltük.

4. táblázat

Kibővített munkapiaci átmenetmátrix a 2009 negyedik és 2010 első negyedéve között, gereblyézással konzisztenssé téve, keresztmetszeti súlyozásból indulva (fő)

| Megnevezés | | | Tárgyidőszak | | | | | | |
|---------------------|------------------------|----------|-------------------|-----------|-----------|-----------------|----------|------------------------|---------------------|
| | | | Munkapiaci státus | | | Egyéb kiáramlás | | Származtatott állomány | Referencia-állomány |
| | | | F | M | I | 65 lett | Egyéb ki | | |
| Bázis-időszak | Munkapiaci státus | F | 3 672 965 | 89 301 | 83 014 | 1 152 | 4 255 | 3 846 432 | 3 846 433 |
| | | M | 30 643 | 284 763 | 20 856 | 0 | 372 | 336 262 | 336 261 |
| | | I | 32 411 | 28 736 | 2 519 115 | 19 339 | 2 876 | 2 599 601 | 2 599 601 |
| | Egyéb beáramlás | 15 lett | 0 | 0 | 21 175 | – | 23 | 21 175 | 21 175 |
| | | Egyéb be | 0 | 0 | 0 | 0 | – | | 0 |
| | Származtatott állomány | | | 3 736 019 | 402 800 | 2 644 160 | 20 491 | | |
| Referencia-állomány | | | 3 736 019 | 402 800 | 2 644 160 | 20 491 | 7 526 | | 6 803 470 |

A 4. táblázat kiemelten fontos, sötéttel keretezett területét a 2. táblázat megfelelő területével érdemes hasonlítani. Korábban azt találtuk, hogy az áramlások alapján a foglalkoztatás változásának mértékét korrekció nélkül jelentősen alulbecsülnénk, ennek eredményeként itt a foglalkoztatásból mindkét irányba mért kiáramlásra nagyobb becslést adunk. A munkanélküliség változása által emelt korlát érvényesül: a növekmény sokkal nagyobb mértékben jelentkezik a foglalkoztatásból munkanélküliségbe áramlás esetében, mint az inaktivitásba áramlásnál. Hasonló gondolatmenet

alapján lefelé korrigáltuk a munkanélküliségből foglalkoztatásba és inaktivitásba, az inaktivitásból munkába lépők, valamint a stabil inaktivitásban levők számát is, felfelé pedig a többi értéket.

Az átlépési esélyek az áramlások abszolút nagyságával együtt változnak, a korrekció hatásának megítéléséhez mégis érdemes ezeket külön is kiszámítani. Az 5. táblázat bemutatja az átlépési esélyeket korrekció nélkül és korrekcióval, valamint azt is, hogy a korrekció hány százalékos változást idézett elő az érintett átmenet nagyságának becslésében (ez egyúttal az abszolút számokban megfigyelhető változást is mutatja).

5. táblázat

*A munkapiaci állapotok között átlépési esélyek gereblyezés előtt és után,
2009 negyedik és 2010 első negyedéve között
(százalék)*

| Munkapiaci státus | F | M | I |
|-------------------|-----------|-------|-------|
| | Eredeti | | |
| F | 0,955 | 0,020 | 0,020 |
| M | 0,107 | 0,836 | 0,068 |
| I | 0,013 | 0,010 | 0,964 |
| | Igazított | | |
| F | 0,955 | 0,023 | 0,022 |
| M | 0,091 | 0,847 | 0,062 |
| I | 0,012 | 0,011 | 0,969 |
| | Változás | | |
| F | 0 | 19 | 7 |
| M | -15 | 1 | -9 |
| I | -6 | 12 | 1 |

Az átlépési esélyek főként a foglalkoztatás és a munkanélküli státus közötti átmenet esetében módosulnak jelentősen, 15, illetve 19 százalékkal. A változás nem tűnik jelentősnek, de közelebbről megvizsgálva látható, hogy mégis az nagyobb, mint azok a változások, amelyeket az átlépési arányban a válság előtti időszakban megfigyelhettünk, és közelíti azok nagyságrendjét, amelyek a válság hatására alakultak ki.

A korrekciós eljárás leírásakor magától értetődőnek kezeltük, hogy a kiinduló átmenetmátrix számítása során súlyoznunk kell, mégpedig keresztmetszeti súlyokkal. Ennek oka az, hogy az elméleti irodalom nem tárgyalja e kérdést, a napi gyakorlat-

ban pedig – legalábbis a BLS-en belül⁴ – keresztmetszeti súlyokat használnak kiindulópontként. A súlyozás szerepe bizonyos mennyiségek keresztmetszeti egyszerű pontbecslések készítése esetében világos, ha azonban ettől eltérünk, már koránt sem az – lásd erről *Johansson* [2007] tanulságos összehasonlító elemzését, amely azt vizsgálja, hogy a súlyozás vagy a torzulások modellezése mely esetekben vezet kedvezőbb eredményre. Megmaradva az egyszerű pontbecslés eseténél, a súlyozásnak akkor van létjogosultsága, ha annak eredményeként a becslés minősége (konzisztenciája és/vagy hibája) javul. A munkaerő-felmérés esetében a súlyozásra egyrészt azért kerül sor, hogy a mintaterv sajátosságait tükrözhesse – erre valók az előre ismert „design” súlyok –, másrészt mert a nem használható címek kiválasztódása és a válaszmegtagadás nem véletlenszerű. Míg a mintasúlyok a mintaterv alapján előre tudhatók, a második ok miatt szükséges a súlyokat, ahogy arról már szóltunk, becslési eljárással előállítani. A becslési eljárás bizonyos Z kulcsváltozók – jellemzően demográfiai tulajdonságok – együttes eloszlása tekintetében biztosítja a minta illeszkedését a kiválasztott referenciaadatokhoz. Amennyiben (például: regressziós) elemzésünk tárgya az X változó, amely teljesen nem azonos Z -vel, akkor a súlyozás hatása olyan mértékben érvényesül, amilyen erős az X és Z közötti összefüggés.

A keresztmetszeti súlyozás alapját a MEF-nél is demográfiai jellemzők adják. A foglalkoztatást jelző változót az elemzés során akkor érdemes súlyozni, ha ezek a jellemzők korreláltak vele (jelen helyzetben úgy véljük: igen), és ha a minta torzulásának explicit modellezése nem lenne célszerű: ilyen az egyszerű és rutinszerűen végzett pontbecslés esete. Az áramlások számításakor azonban a helyzet összetettebb. Itt a súlyozást az X -ek egy adott státussal való korrelációja még nem indokolja, sokkal inkább az, hogy magával az átmenetet jelző változóval (tehát az időben egymást követő státusok interakciójával) korrelált. Ennek eldöntése az érintett változók ismeretét és külön elemzést igényli. Sajnálatos módon a munkaerő-felmérés módszertanát ismertető kiadvány (*KSH* [2006]) nem szolgál e téren részletes információval, *Mihályffy* [1995] pedig a munkaerő-felmérés súlyozásának csak az elvét dokumentálja, a pontos eljárást azonban, ide értve a referenciaként felhasznált változók ismertetését, nem. Ezek ismerete nélkül nem vizsgálható a referenciaváltozók korrelációja a mintában maradással, ezáltal a súlyok alkalmazásának relevanciája sem. Támponként *Molnár* [2005], a szintén ELAR mintakeretre alapozott, a HKF-fel kapcsolatban szerzett tapasztalataira hivatkozhatunk, amely szerint a keresztmetszeti súlyok kalibrálása esetében a demográfiai szempontok szerepet kapnak. Mivel tudjuk, hogy ezen tulajdonságok közül több összefügg a munkapiaci sikerességgel, illetve sérülékenységgel, arra a következtetésre juthatunk, hogy a keresztmetszeti súlyok a becslést segítik, és nem hátráltatják. Ilyen megfontolásból azokat az itt közölt számításoknál felhasználtuk, így a kiindulásként alkalmazott átmenetmátrix súlyozott.

⁴ A BLS munkatársának, *Greg Erkensnek* személyes közlése alapján.

5. A munkanélküliség alakulása a válság alatt az átmenetek tükrében

Az áramlások korrekcióját nemcsak egy időszakra, hanem egy teljes idősorra is alkalmazhatjuk, hiszen az eljárás páronkénti megfelelést biztosítva az áramlások tetszőleges hosszúságú idősorát konzisztenssé teszi. A 6. táblázat az előbbieken ismertetett módszert alkalmazva az átmenetmátrixnak a munkanélküliséggel kapcsolatos minden komponensét, a 2. ábra pedig a munkanélküliség változását és csak azon komponenseit ábrázolja a 2007 első és 2010 első negyedéve közötti időszakra, amelyeket a rutinszerű elemzésben használni szoktunk. A munkanélküli-állomány időbeli változását mindkét esetben D.M-mel jelöltük, a gereblyezett áramlásokból származtatott változás pedig a D.Mi fejrovat alatt található a táblázat utolsó oszlopában a korrekció pontosságának megítélése érdekében. Láthatjuk: az értékek közel vannak egymáshoz, a korrigált áramlások tehát jól visszaadják a megfigyelt változásokat.

6. táblázat

Az ILO-definíciót követő munkanélküliség alakulása 2007 és 2010 első negyedéve között –
állományi és gereblyézéssel állománykonzisztenssé tett áramlási adatok
(fő)*

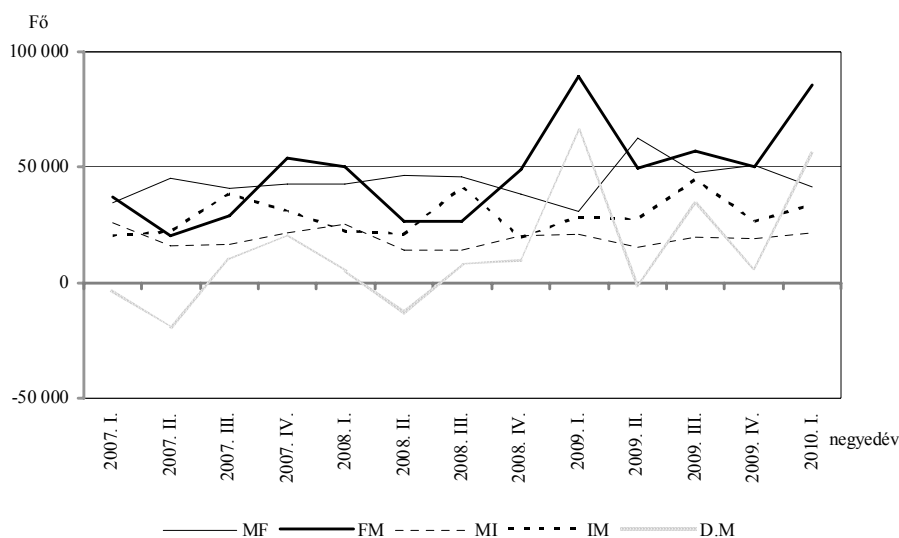
| Negyedév | M | MM | MF | FM | MI | IM | M65 | ME | 15M | EM | D.M | D.Mi |
|------------|---------|---------|--------|--------|--------|--------|-----|-----|-----|-----|---------|---------|
| 2007. I. | 316 044 | 258 661 | 34 604 | 36 933 | 25 726 | 20 450 | 0 | 465 | 0 | 0 | -3 411 | -3 412 |
| 2007. II. | 296 646 | 253 857 | 45 508 | 20 599 | 16 246 | 22 190 | 0 | 433 | 0 | 0 | -19 398 | -19 398 |
| 2007. III. | 306 662 | 239 152 | 40 847 | 29 183 | 16 375 | 38 327 | 0 | 272 | 0 | 0 | 10 016 | 10 016 |
| 2007. IV. | 327 322 | 242 263 | 42 874 | 53 951 | 21 525 | 31 108 | 0 | 0 | 0 | 167 | 20 660 | 20 827 |
| 2008. I. | 332 390 | 259 705 | 42 634 | 50 352 | 25 118 | 22 333 | 0 | 32 | 0 | 0 | 5 068 | 4 901 |
| 2008. II. | 318 970 | 271 615 | 46 443 | 26 254 | 13 981 | 21 100 | 0 | 350 | 0 | 0 | -13 420 | -13 420 |
| 2008. III. | 327 142 | 258 703 | 45 799 | 26 815 | 14 260 | 41 624 | 0 | 208 | 0 | 0 | 8 172 | 8 172 |
| 2008. IV. | 336 633 | 267 911 | 38 373 | 49 268 | 20 650 | 19 454 | 0 | 208 | 0 | 0 | 9 491 | 9 491 |
| 2009. I. | 402 800 | 284 763 | 30 643 | 89 301 | 20 856 | 28 736 | 0 | 372 | 0 | 0 | 66 167 | 66 166 |
| 2009. II. | 401 368 | 324 077 | 62 618 | 49 825 | 15 284 | 27 465 | 248 | 573 | 0 | 0 | -1 432 | -1 433 |
| 2009. III. | 435 361 | 333 637 | 47 664 | 57 120 | 19 969 | 44 603 | 0 | 98 | 0 | 0 | 33 993 | 33 992 |
| 2009. IV. | 441 744 | 365 033 | 50 702 | 50 001 | 19 234 | 26 710 | 0 | 392 | 0 | 0 | 6 383 | 6 383 |
| 2010. I. | 497 557 | 378 215 | 41 626 | 85 671 | 21 903 | 33 672 | 0 | 0 | 0 | 273 | 55 813 | 56 087 |

* MEF 60. és 73. hullám.

A 2. ábrából 2008 végéig egy viszonylag stabil MF és MI áramlással, illetve minimális szezonális szezonalitást mutató FM és IM áramlással jellemezhető időszakot azonosíthatunk.

A változást 2008–2009 telén egy rendkívül erős, a foglalkoztatásból a munkanélküliségbe irányuló áramlás hozza meg, ami a korábnál mintegy 30 ezer fővel magasabb szinten látszik stabilizálódni. Ez egyúttal a beáramlás mértékének megduplázását is jelenti. Mindeközben a munkanélküliségből a foglalkoztatásba áramlás is nő, de egyik negyedévben sem képes a munkanélküliséget nettó módon növelő áramlást megfordítani. Az inaktivitásból munkanélküliségbe áramlás ezalatt nem változik jelentősen, illetve nettó módon kismértékben nő. A válság során tehát a korábnál jóval nagyobb mértékben nőtt meg a munkanélküliségbe áramlás, de egyúttal fokozódott a munkanélküliségből a munkába áramlás intenzitása is. Mindez összeegyeztethető a válság tisztító hatásával, az azonban ebből nem nyilvánvaló, hogy az inaktivitásból a munkanélküliségbe áramlás háttérben a támogatott foglalkoztatásában bekövetkezett változások állnak. Erről a jelen cikk egy bővebb munkaváltozata ad számot (Cseres-Gergely [2010]).

2. ábra. A gereblyézéssel konzisztenssé tett munkanélküliség-változás és komponensei 2007 és 2010 első negyedéve között



6. Következtetések

Ebben az írásban a munkapiaci áramlások konzisztens számbavételének módszerét tárgyaltuk. Bemutattunk egy egyszerű eljárást, a gereblyézést, és azt, hogy miként lehet segítségével a munkapiaci állományok változását dekomponálni más munkapi-

aci állapotok felől érkező és oda tartó átmenetekre. Az eljárás a KSH munkaerő-felmérésére alkalmazva hatékonynak bizonyult, amennyiben használatával jelentős, a munkanélküliség és a foglalkoztatás közötti átmenetek esetében közel 20 százalékos inkonzisztens eltérést sikerült korrigálni. A számításokkal szerzett tapasztalatok alapján úgy gondoljuk, hogy az eljárás alkalmazható mind az aggregált áramlási idősorok, mind az ezekre alapozott átmenetmátrixok számítására. A tapasztalt eltérések nagysága miatt nemcsak ajánlott a korrekció – főként a kisebb állományokat érintő nagyobb változások esetében –, de szükséges is ahhoz, hogy az adatokból hiteles következtetést vonhassunk le.

Irodalom

- ABOWD, J. M. – ZELLNER, A. [1985]: Estimating Gross Labor-force Flows. *Journal of Business & Economic Statistics*. 3. évf. 3. sz. 254–283. old.
- CSERES-GERGELY Zs. [2010]: *Munkapiaci áramlások, gereblyezés és a 2008 végén kibontakozó gazdasági válság foglalkoztatási hatásai*. BWP 2010/4. MTA Közgazdaságtudományi Intézet. Budapest.
- DEMING, W. E. – STEPHAN, F. F. [1940]: On a Least Squares Adjustment of a Sampled Frequency Table when the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics*. 11. évf. 4. sz. 427–444. old.
- ELSBY, M. W. – HOBIJN, B. – SAHIN, A. [2010]: *The Labor Market in the Great Recession*. National Bureau of Economic Research. Working Paper. Series 15979. Cambridge. <http://www.nber.org/papers/w15979>
- FAGAN, J. T. – GREENBERG, B. V. [1988]: *Algorithms for Making Tables Additive: Raking, Maximum Likelihood, and Minimum Chi-Square*. Bureau of the Census. Washington, D.C.
- FRALLER G. – HARCSA I. – JÓNÁS I. – KMETTY Z. [2007]: *A 2001. évi népszámlálás és a 2005. évi mikrocenzus összekapcsolása – Módszertani kísérlet*. Munkaanyag.
- FRAZIS, H. J. – ROBINSON, E. L. – EVANS, T. D. – DUFF M. A. [2005]: Estimating Gross Flows Consistent with Stocks in the CPS. *Monthly Labor Review*. 128. évf. 9. sz. 3–9. old.
- FSZH (FOGLALKOZTATÁSI ÉS SZOCIÁLIS HIVATAL) [2010]: *Az aktív foglalkoztatáspolitikai eszközök működése 2009-ben*. Budapest.
- HECKMAN, J. J. [1979]: Sample Selection Bias as a Specification Error. *Econometrica*. 47. évf. 1. sz. 152–161. old.
- JOHANSSON, F. [2007]: *How to Adjust for Nonignorable Nonresponse: Calibration, Heckit or FIML?* Uppsala University. Munkaanyag.
- KÁTAY, G. – NOBILIS, B. [2009]: *Driving Forces Behind Changes in the Aggregate Labour Force Participation in Hungary*. MNB Working Papers 2009/5. Budapest. http://english.mnb.hu/Root/Dokumentumtar/ENMNB/Kiadvanyok/mnben_mnbfuzetek/mnben_WP_2009_5/wp_2009_5.pdf
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2006]: *A munkaerő-felmérés módszertana*. Statisztikai Módszertani Füzetek. 46. Budapest. <http://portal.ksh.hu/pls/ksh/docs/hun/xftp/idoszaki/pdf/munkfelmmod.pdf>

- MIHÁLYFFY L. [1995]: Meghiúsulások kompenzálása lakossági felvételekben: egy speciális lineáris inverz probléma. *Sigma*. XXV. évf. 4. sz. 191–202. old.
- MOLNÁR GY. [2005]: Az adatállomány és a rotációs panel. In: *Kapitány Zs. – Molnár Gy. – Virág I.* (szerk.): *Háztartások a tudás- és munkapiacra*. MTA KTI. Budapest.
- PARIKH, A. [1979]: Forecasts of Input-Output Matrices Using the R.A.S. Method. *The Review of Economics and Statistics*. 61. évf. 3. sz. 477–481. old.
- RUBIN, D. B. [1976]: Inference and Missing Data. *Biometrika*. 63. évf. 3. sz. 581–592. old.

Summary

Analyses of the Hungarian employment situation are almost exclusively concerned with the change in the number of the employed, unemployed and inactive, not paying particular attention to the flows between these states. The present paper discusses a method of calculating labour market flows in a consistent way and the employment effects of the crisis unfolding after 2008, separating supported and not-supported employment. The presented method is easily applicable in the everyday work of analysts and researchers and supplies information not available from the analysis of stocks.