

Székelyhidi Katalin,
a Szegedi Tudományegyetem
PhD-hallgatója
E-mail: kata.szekelyhidi@gmail.com

A munkaerőpiac előrejelzésére használt modellek és az előrejelzés lehetőségei

DOI: 10.20311/stat2016.03.hu0300

A munkaerőpiac előrejelzése, modellezése ugyan régóta kutatások tárgya, ennek ellenére még nem született olyan modell, amely általánosan elfogadott módszerként használható lenne az előrejelzésére. Az ilyen modellek többnyire a foglalkoztatás egyszerű, más tényezőket figyelembe nem vevő extrapolációjától a bonyolultabb, dinamikus strukturális modellekig terjednek, amelyek számos befolyásoló tényezőt vesznek figyelembe és igyekeznek megmutatni a munkaerő-piac komplexitását. Az ezen a területen végzett előrejelzések egyik nehézségét tehát a jelenség összetett és dinamikus tulajdonsága adja, azaz számos tényező befolyásolja alakulását. Időben is változékony, hiszen nem tudhatjuk, hogy például évek múlva egy képzési szerkezet-változás hogyan befolyásolja a munkaerőpiacot. Éppen ezért kulcsfontosságú a munkaerőpiacot megfelelően leíró modell felépítése. A munkaerő-piaci előrejelzések egy másik kritikus területe az előrejelzésre használt statisztikai technika. Ugyanis a különböző modellek más-más hibával dolgoznak, így az előrejelzések pontossága nagymértékben függ a használt előrejelzési modellektől.

1. Előrejelzési modellek és módszerek

A munkaerőpiac előrejelzésére létrehozott modellek hosszú múltra tekintenek vissza. Már az 1960-as években is folytattak tervezési céllal ilyen irányú kutatásokat. Az első ilyen előrejelzést az OECD-országok munkaerőpiacának vizsgálatára hajtották végre. Ebben az időszakban a munkaerőpiac előrejelzésének célja a tervezés

alátámasztása volt, azaz adott gazdasági cél, például növekedési célok mellett vetítették előre a munkaerő szükségletet számos foglalkozás és képzettség viszonylatában. A megközelítés célja az volt, hogy oktatási, képzési stratégiát állítsanak fel meghatározott gazdasági cél elérése érdekében. Ezeket az előrejelzéseket számos kritika érte. A modellek legnagyobb hibájaként azt vetették fel, hogy:

- az adatok megfelelősége kérdőjelezhető volt,
- módszertanilag kevésbé voltak kidolgozottak ezek a modellek, hosszú távú előrejelzésre nem voltak alkalmasak,
- túlzottan mechanikusan működtek és nem vették figyelembe a munkaerőpiac jellemzőit, dinamikáját, a szektorok és foglalkozások vándorlási folyamatait,
- rögzített együtthatókkal kapcsolták össze az ágazati növekedést a munka keresletével, így ez nem adott megbízható képet a kereslet jövőjét illetően,
- nem vették figyelembe a kereslet és kínálat együttmozgását,
- egy-egy foglalkozás munkaerő igényének meghatározása nem ad elégséges információt oktatási programok kialakításához.

A kritikákra adott válaszként a modelleket hosszú távú helyett középtávú előrejelzésre hangolták át, és az oktatási program kialakítása helyett sokkal általánosabb célokat határoztak meg úgy, mint foglalkoztatási trendek vizsgálata, a döntéshozók és tervezők támogatása a munkaerő-piaci trendek bemutatásával (*Neugart–Schömann* [2002]).

Az előbbieknél megfelelően az előrejelzésekkel szemben támasztott követelmények *Raupeliené* [2011] megfogalmazása szerint a következők:

- az előrejelzéseknek középtávra kell készülniük (ötéves periódusokban), továbbá figyelembe kell venni az időeltolódást is az előrejelzésekben, ami abból ered, hogy a legtöbb előrejelzés alapját éves adatok képezik, ami azt jelenti, hogy az adatok szükségszerűen legalább egy év késleltetéssel érhetők el,
- a foglalkoztatás szerkezetének vizsgálatában magyarázó modelleket kell alkalmazni a jelenség pontosabb megértése érdekében,
- az előrejelzéseket meghatározott időszakonként ismételni kell,
- az előrejelzési modellekben figyelembe kell venni a regionális különbségeket és ennek megfelelően kell kalibrálni a modelleket,
- más gazdasági szektorok folyamatait is figyelembe kell venni a munkaerő-piac minél pontosabb megértéséhez (*Raupeliené* [2011], *Neugart–Schömann* [2002]).

A munkaerőpiac előrejelzésére számos módszer használatos, ezek közül néhány jelentősebb és feltörekvő technika a következő:

– A leggyakrabban használt idősorelemzés, amely az idősorokban található rendszerességet használja fel előrejelzésre. A legalapvetőbbként említett idősor-elemzési technika a trendelemzés és -simítás, míg az összetettebb modellek egyike az autoregresszív modell és a mozgóátlagolás módszerének kombinálását végzi el, amelyet ARIMA- (autoregressive integrated moving average – integrált autoregresszív mozgóátlag) modellként is neveznek. Az autoregresszív modell a vizsgált változó értékeit annak korábbi időszakhoz tartozó értékei alapján jelzi előre. Továbbá még számos autoregresszív modellezésre épülő technika használatos, amelyeket a későbbiekben mutatok be.

– A kereslet- és kínálatorientált, makroökonómiai modellek egyrészt a kereslet és kínálat egyenlőtlenségeit vizsgálja, másrészt az egyes iparágak és képzések foglalkoztatási sajátosságait is figyelembe veszi,

– a lineáris programozás, amelynek lényege, hogy bizonyos feltételek mellett megkeressük egy függvény minimumát vagy maximumát. A modell erősségét az adja, hogy az előrejelzés szempontjából fontos információkat használjuk fel, továbbá lehetőséget nyújt érzékenység elemzésekre is,

– a neurális hálók módszere, amelynek használata munkaerő-piaci kutatások során kevésbé elterjedt. Lényege, hogy a regresszióval ellentétben nem szükséges a függő- és független-változók közötti kapcsolatok meghatározása, hanem az előrejelzés bemeneteként változók közötti nemlineáris kapcsolatokat modellez, ahol az inputokat transzformációnak veti alá és egy vagy több nemlineáris indikátorral cseréli fel (*Hampel et al.* [2008]).

1.1. Idősoros modellek

Az idősoros elemzések között megkülönböztetünk egyszerűbb technikákat, mint például a lineáris extrapoláció, az exponenciális simítás és a naiv előrejelzési módszerek. Ugyanakkor ezek sok esetben nem képesek az idősorok dinamikus szerkezetét is belefoglalni a modellbe. Ezen igény kielégítésére később ennek fontosságát hangsúlyozó modellek is születtek. A szofisztikáltabb módszerek közé tartozik a korábban már említett ARIMA-modell, amely az autoregressziós modellt és a moz-

góátlagolású trendillesztést kapcsolja össze (Ashenfelter–Card [1982], Wong–Chang–Chiang [2005], KSH [2005]).

Wong és szerzőtársai megfogalmazása szerint e modell népszerűsége abban rejlik, hogy viszonylag egyszerű az alkalmazása és megbízhatók az eredményei. A módszertan általános megközelítését Box és Jenkins dolgozta ki 1976-ban, és azóta is széles körben elterjedt idősor-elemzési technikaként használatos. Alkalmazható stacionárius és nem stacionárius, szezonális hatással rendelkező, illetve általa nem befolyásolt idősorok esetében is. Abban az esetben adja a legjobb eredményt, ha olyan környezetben használják, ahol a magyarázóváltozók nem változnak nagymértékben az idővel. A modell alkalmazása egy négy lépcsőből álló ismétlődő folyamatként fogható fel (Wong–Chang–Chiang [2005]).

1. Az első lépésben meg kell bizonyosodni az idősor stacionárius voltáról, hogy az idősor miképpen viselkedik hosszabb időszakon keresztül, azaz az idősor az időbeli eltolásra érzékenyen reagál-e. Egy idősor akkor stacionárius, ha a vizsgált jelenséget valamilyen sokk éri, és ennek hatása az idővel nem jelentkezik, nem fejt ki hatást a folyamatra, azaz hosszú távon az átlag és a variancia állandó és értelmezhető. Ezzel ellentétben, ha a hatás később is érvényesül, akkor nem stacionárius idősről van szó, és ebben az esetben stacionáriussá kell transzformálni az idősort (Mák [2011]). A stacionaritás a Dickey–Fuller-próbával tesztelhető, amely az egységgyököket vizsgálja idősorok esetében. A próba null hipotézise szerint az idősor nem tekinthető stacionáriusnak, azaz transzformációra van szükség.

2. Az idősor stacionaritásának megállapítása, vagy azzá történő transzformálása után következik az ARIMA-modell felírása, azaz az idősor paramétereinek és a leírására alkalmas modellnek a meghatározása (Wong–Chang–Chiang [2005]). A modell egy autoregresszív és mozgóátlagosú trendtagból épül fel. Az AR-tag (autoregressive – autoregresszív) a vizsgált változó értékeit annak korábbi időszakhoz tartozó értékei alapján jelzi előre, míg a mozgó átlagosú tag a vizsgált változó értékei és a korábbi időszakhoz tartozó fehér zaj közötti kapcsolatot ragadja meg és jelzi előre. A modell fő célja, hogy az idősorban levő véletlen folyamatot beazonosítsa és előre jelezze az AR- és MA-modell (moving average – mozgóátlag) segítségével (Bowerman [1987], Wei [1990], Nau [2015]). A modell azonosítása az autokorreláció és a parciális autokorreláció függvénye alapján történik. Ennek során meg kell vizsgálni, hogy az idősor milyen periodicitású, azaz az adatok például negyedéves vagy havi bontásban értelmezhetők, hogy az idősor jelenlegi értékét a múltbéli értékek hány időszakra visszamenőleg befolyásolják (a modell p , d és q paramétereinek megválasztása), továbbá vizsgálni kell, hogy érvényesül-e szezonális hatás az idősor esetében (Balan [2012]). Ha a vizsgált jelenség szezonalitást mutat, akkor a SARIMA- (seasonal autoregressive integrated moving average – szezonalitást figyelembe vevő, integrált autoregresszív mozgóátlag) modell használatos, amely figyelembe veszi a

szezonális hatást is. A szezonalitással bővített modell általános egyenlete a következő:

$$\varphi(b)\Phi(B)\Delta^d\Delta_s^D E_t = \theta(b)\Theta(B)\varepsilon_t, \quad /1/$$

ahol

- b és a B a szezonális és nem szezonális késleltetési operátorok,
- $\varphi(b)$ és $\Phi(B)$ a szezonális és nem szezonális autoregresszív polinomok, illetve
- $\theta(b)$ és $\Theta(B)$ szezonális és nem szezonális mozgóátlagolású polinomok,
- ε_t jelöli a fehérzaj-folyamatot (*Sarantis–Swales* [1999], *KSH* [2005]).

3. A modell illeszkedésének tesztelése, és ha szükséges, a modell javítása. A modell illeszkedésének tesztelésére tipikus eszközök az MSE (mean squared error – átlagos négyzetes hiba) és az információs kritériumok: az Akaike-féle információ kritérium és a bayesi információs kritérium (*Sarantis–Swales* [1999]). Ha a modell javításra szorul, akkor a folyamatot előlről kell kezdeni, azaz meg kell vizsgálni az idősor stacionárius voltát és a szezonális jelenlétét továbbá meg kell határozni a paramétereket és újra fel kell írni a modellt.

4. Az utolsó lépés a végső modell felhasználásával előrejelzés készítése (*Wong–Chang–Chiang* [2005]).

Az említett szerzők Hongkong építőiparának munkaerőpiacára készítettek előrejelzést az ARIMA-modell segítségével. Az előrejelzett változók a foglalkoztatási szint, a termelékenység, a reálbérek nagysága, munkanélküliségi ráta és az alulfoglalkoztatási ráta voltak. Megállapításaik szerint a módszer megfelelő eredményt hozott, ugyanakkor olyan modellek használata célszerűbb lehet a foglalkoztatás előrejelzésében, amelyek alkalmasak nem stacionárius és nemlineáris idősorokból történő előrejelzésre is. Ezt indokolja az is, hogy a foglalkoztatási szint előrejelzésére nem volt alkalmas a modell, amelyet a szerzők két terület robbanásszerű bővülést okozó építési hullámának tulajdonítottak (*Wong–Chang–Chiang* [2005]).

A szakirodalomban számos más kutató is alkalmazta a módszert a foglalkoztatás előrejelzésére. *Balan* [2012] Románia foglalkoztatásának előrejelzésére használta 2015 első félévére a 2002 és 2010 periódus adatait felhasználva. A munka során két modellt hozott létre. Az első egy egyszerű autoregresszív modell volt, amely nem vette figyelembe a szezonalitást, illetve egy SARIMA-modellt is létrehozott, amely számításba vette a szezonális hatást is és integrálta az autoregresszív modellt a moz-

góátlagolásával. A két modell eredményeit összehasonlítva az információs kritériumok alapján arra jutottak, hogy utóbbi modell jobb eredményt ad.

Trívez–Mur [1997] rövid távú előrejelzés készítésére használta az ARIMA-modellt Spanyolország Aragónia régiójára. A kutatás újszerűségét az adta, hogy a modellt regionális foglalkoztatás előrejelzésére használták, azaz a modellt szükség-szerűen össze kellett kapcsolni a nemzeti szinttel. Viszont a nemzeti szintű adatok a regionális adatok aggregálásával keletkeznek, így a regionális és nemzeti szintű változók között ok-okozati kapcsolat lehetséges, azaz a regionális változó meghatározza a nemzeti szintűt. Ugyanakkor a szerzők ennek lehetőségét kizárták, mivel a vizsgált régió szerkezete hasonló a nemzeti piac szerkezetéhez és teljesítményének csak kis hányadát (3,5%) teszi ki. Éppen ezért kizárták annak a lehetőségét, hogy a régió hatással van a nemzeti szintre. Így az egyenletek felírása során a magyarázóváltozók ugyanazok regionális szinten, mint nemzeti szinten.

Az ARIMA- és a SARIMA-modellek mellett a szakirodalom más autoregresszív modelleket is használt a munkaerő-piaci folyamatok előrejelzésére, mint például a BVAR-modellt (Bayesian vector autoregression – bayesi vektor autoregresszív).

A bayesi modellezést leginkább olyan jelenségek előrejelzésére használják, amelyek kimenetele nagyon bizonytalan. A módszer lényege, hogy véletlen változónak tekinti azokat a változókat, amelyek nem ismertek, és ezeket a paramétereket az eloszlásfüggvény alapján határozza meg. A bayesi módszer alkalmazásakor előzetesen meg kell adni a paraméterek eloszlásfüggvényét, ki kell választani a valószínűségi függvényt és végül ismételt meg kell adni a paraméterek korrigált eloszlásfüggvényét (*Kovács–Balogh* [2009]).

A VAR- (vector autoregression – vektor autoregresszív) modell az előrejelezni kívánt változók korábbi időszakokhoz tartozó értékeinek szabályszerűségein alapszik. Az ARIMA-modell többváltozós kiterjesztése a VAR-modell. A VAR- és ARIMA-modell problémája a túl sok paraméter használata és előrejelzési hibák jelentkezése, amelyeket a statisztikailag nem szignifikáns késleltetési operátorok kizárásával oldanak meg (*Puri–Soydemir* [2000]). A VAR-moddellel szembeni legnagyobb kritika, hogy ahogy növekszik az endogén változók száma, úgy a modell a szabadságfokok problémájával küzd, illetve a nem stacionárius idősorok esetében több hiba keletkezik a becslésben (*Seung–Ahn* [2010]). Ennek elkerülésére nyújt megoldást a BVAR-modell, amely a változó egyes késleltetési operátorainak kizárása helyett a kevésbé fontos együtthatók esetében megszorításokat tesz. Ez akkor fordulhat elő, ha adott esetben egy hosszabb késleltetési operátor (lag) statisztikailag bizonyítottan kevésbé fontos egy rövidebb késleltetési operátorhoz képest. A szerzők kutatási eredményei szerint a BVAR-modell jobb előrejelzéseket ad, mint a vizsgált másik két modell, mivel kiküszöböli azok hátrányait, nevezetesen azt, hogy az ARIMA-modell a változók múltbéli értékeiből ad előrejelzést a jövőre vonatkozóan, illetve a tiszta VAR-modell megkötések hiányából adódó tulajdonságát. A BVAR-

modell a változások irányának előrejelzésére is jobb eredményt kínál, ezért alkalmas olyan gazdasági folyamatok előrejelzésére, ahol fontos a jelenségek jövőbeli irányvonalának meghatározása (Puri–Soydemir [2000]).

Az idősoros modellek egy további lehetősége az ARDL- (autoregressive distributed lag – autoregresszív osztott késleltetésű) modell, szintén egy AR-modell, amely az idősorokban a jelenlegi időszak értékét a magyarázóváltozó jelenlegi és előző periódushoz tartozó értékei alapján jelzi előre. A modell segítségével nem csak az előrejelezni kívánt változó múltbéli értékeinek extrapolációjára kerül sor, hanem bevonható egy másik befolyásoló változó is az előrejelzésbe. Ez kifejezetten hasznos, hiszen korábban említettem, hogy a munkaerőpiacot sok tényező alakítja, így bevonható más tényező is a célváltozó előrejelzésébe. Ugyanakkor tény, hogy nem elegendő csak egy-egy változóval magyarázni. Az ARDL-modell használatának továbbfejlesztett lehetőségeit kutatta Rapach–Strauss [2008], akik az Egyesült Államok foglalkoztatásának előrejelzésére alkalmazták az ARDL-modellt oly módon, hogy az ezzel a modellel kapott előrejelzéseket különböző technikákkal fonta össze és hasonlította össze az eredményeiket. Harminc változó felhasználásával készítettek ARDL-alapú előrejelzéseket, majd ezek eredményeit különböző módszerekkel kombinálták. Tehát elkészítettek harminc egyedi előrejelzést különböző magyarázó tényezők felhasználásával, majd ezek eredményeit különféle technikákkal összefűzték egy végső, minden magyarázóváltozót tartalmazó előrejelzéssé. A különböző technikákkal egyesített végső előrejelzéseket összehasonlították. Az összehasonlítás az MSFE (mean squared forecast error – átlagos négyzetes előrejelzési hiba) mutatóval történt. A módszerek eredményeit vizsgálták az egyszerű AR-modell összevetésével is, amely a változó értékét a korábbi időszakokhoz tartozó értékei alapján becsüli meg, azaz magyarázóváltozókkal nem számol. Az egyedi ARDL-előrejelzések során olyan magyarázóváltozók voltak használatosak, mint például a heti átlagos munka-idő, munkaerő-piaci részvételi arányok korcsoportonként, nemenként stb.

Az előrejelzés eredmények kombinálásának jogosultságát a foglalkoztatás korábban már említett komplex tulajdonsága adja, azaz nem készíthető megfelelően illeszkedő előrejelzés egyetlen változó segítségével, továbbá a magyarázóváltozók köre időben is rendkívül változékony. Az egyedi előrejelzések alapját képező ARDL-modell képlete:

$$y_{t+h}^h = \alpha \sum_{j=0}^{q_1-1} \beta_j \Delta y_{t-j} + \sum_{j=0}^{q_2-1} \gamma_j x_{i,t-j} + \varepsilon_{t+h}^h, \quad /2/$$

ahol

– $x_{i,t}$ az adott modellben használt előrejelző, magyarázóváltozót jelöli,

- h az előrejelzés időtávja,
- ε_{t+h}^h a hibatag és
- α, β, λ a legkisebb négyzetek módszerével becsült paraméterek.

Az egyedi ARDL-modellek különböző technikákkal végzett kombinálása az alábbi képlet szerinti modellel történik:

$$\hat{y}_{c,t+h|t}^h = w_{0,t} + \sum_{i=1}^n w_{i,t} \hat{y}_{i,t+h|t}^h, \quad /3/$$

ahol $y_{i,t+h|t}^h$ az adott előrejelzési módszer eredménye $\{w_{i,t}\}_{i=0}^n$ súlyokkal.

A kombinált előrejelzéseket a következő technikákkal készítették el:

- egyszerű módszerekkel úgy, mint az eredmények átlagolásával, trimmelt átlag és medián használatával,
- az OLS (ordinary least squares – legkisebb négyzetek) módszerével összevetve az egyes előrejelzések eredményét,
- WLS (weighted least squares – súlyozott legkisebb négyzetek) módszerével, amely annyiban különbözik az előző technikától, hogy a meghatározott súlyok az időhöz kapcsolódó változékonyságot próbálják megragadni, amely strukturális változások jelenlétekor hasznos,
- inverz MFSE számításával, amely az egyenletben az előrejelzések súlyait a múltbéli előrejelzési képességük alapján határozza meg,
- bayesi információ-tömörítő technikákkal, ahol λ a tömörítés mértékének súlyát jelenti (*Rapach–Strauss* [2008]),
- a GETS (general-to-specific – általánostól az egyedi esetekig) modellel történő következtetéssel, amelynek során az egyedi esetekhez vezető úton a nem szignifikáns változók kizárásra kerülnek (*Rapach–Strauss* [2012]),
- főkomponens elemzéssel, amely az egyedi előrejelzések m faktorát határozza meg, majd ezeket regressziós modell segítségével kombinálja össze,
- közelítő bayesi átlagolással, amely az egyedi modellek súlyainak utólagos közelítésén alapszik,
- exponenciális újrásúlyozással, amely az egyes előrejelzések súlyainak exponenciális módon történő meghatározását jelenti,
- klaszteranalízis: az ARDL-előrejelzések eredményeit klaszterelemzés alá vetik. A klaszterezés az átlagos négyzetes hibák alapján történik, azaz a legkisebb hibával dolgozók kerülnek egy klaszterbe,

további klasztert alkotnak a közepes nagyságú hibával dolgozók és így tovább egészen a legnagyobb hibával dolgozó előrejelzések csoportjáig (*Rapach–Strauss* [2008]).

Az eredményeik alapján számos előbb említett ARDL-alapú technika jobban teljesít, mint az AR-modell. A kombinált előrejelzések közül a főkomponens elemzéssel kombinált ARDL-modell működött a legjobban. Az AR-modellhez képest a hagyományos átlagolás, az inverz MSFE, klasztertechnika, a közelítő bayesi átlagolás is jobb eredményt ad, mint a tiszta AR-modell. Az OLS-, illetve WLS-technikák az AR-modellhez képest is gyengébben teljesítettek. *Rapach–Strauss* [2012] ugyanezt a technikát alkalmazta tizenegy változó bevonásával a korábbi harminc helyett. Későbbi munkájukban a korábban már használt GETS-, inverz MSFE-módszereket és egy közelítő faktor modell struktúrát használták fel a foglalkoztatás növekedésének előrejelzésére az Egyesült Államokban. Utóbbi technika az államok és a nemzeti szintű foglalkoztatás kapcsolatára koncentrált. Meghatározza egy-egy állam *béta* együttthatóját a nemzeti szintű foglalkoztatás növekedését leíró regressziós modellben, majd egy-egy állam foglalkoztatásának előrejelzése során az előbb meghatározott nemzeti szintű növekedést beilleszti a modellbe.

A szerzők egyrészt összehasonlították a három modell eredményeit, egymáshoz viszonyítva és az AR-modellhez képest is az MSFE-vel, másrészt egyesítették a három modell eredményeit az előre jelzett növekedések egyszerű átlagolásával. Az inverz MSFE-módszerrel adott előrejelzés az összehasonlításként használt AR-modellhez képest, egyetlen állam kivételével, kisebb hibával dolgozott, mint az egyszerű AR-modell. A másik két modell még ennél is pontosabban dolgozott néhány államban, ugyanakkor némely esetben alulmaradt az AR-modellhez képest. A három modell eredményeinek összeolvasztása konzisztensebb eredményt hoz az utóbb említett két modell előrejelzéséhez képest, másrészt kisebb hibával dolgozik, mint az inverz MSFE-modell.

1.2. Hierarchikus modell

Frees [2003] a hagyományos többváltozós idősoros modellek helyett hierarchikus és longitudinális modelleket használt a munkaerőpiac előrejelzésére. A munkaerőpiaci részvételi arányokat demográfiai cellánként külön számította kor, nemek, családi állapot és gyermek jelenléte szerint, majd egymás függvényében jelezte előre a foglalkoztatást.

A hierarchikus, más néven többszintes előrejelzési modellek lineáris hatásmodellek, amelyeket leginkább oktatási és pszichológiai területen használnak. *Frees* [2003] megkísérelte meghatározni a munkaerő-piaci részvételi arányok legfontosabb tényezőit és

ezek hierarchikus struktúráját egy kovariancia-mátrix segítségével, amellyel a cellákra jellemző együtthatókat és modellt határozta meg. Később, a hierarchikus modell gyenge pontjait tovább fejlesztve, egy longitudinális modellt alkotott meg, amellyel a túlságosan sok dimenzióval rendelkező (101×101) kovariancia-struktúra hatékonyabbá tehető.

A rendelkezésre álló adatokból huszonkilenc korcsoportot képzett, amelyeken belül, ahol lehetséges volt, még több csoportosításokat hajtott végre további három kategória (családi állapot, gyermek jelenléte, nem) alapján. A korcsoportok nem egységesek, azaz a 16–17, illetve a 18–19 évesek mellett a 20 év felettieknél öt éves korcsoportokat képzett 54 éves korig, majd az egy éves korcsoportok következnek 74 éves korig. Ennek eredményeként 101 demográfiai cella jött létre. Az így létrehozott mátrix alapötlete, hogy a szomszédos cellák információt osztanak meg egymással, és erre alapul az egész technika. A korrelációkat a nagyobb megbízhatóságuk érdekében, a lakosság száma alapján súlyozta és azt vizsgálta, hogy az egyes cellák vagy nagyobb csoportok miként korrelálnak a munkaerő-piaci részvételi arányokkal. Így például azt is elemezte, hogy a családi állapot és a foglalkoztatás milyen kapcsolatban állnak egymással. A későbbi modellek érdekében a létrehozott cellákat aggregálta, és így a következő három csoport jött létre:

- 16 és 18, illetve 55 év feletti férfiak és nők vegyesen, családi állapot és gyermek lététől függetlenül,
- 20 és 24, illetve 50 és 54 év közötti férfiak csoportja, továbbá a 45 és 49, illetve az 50 és 54 év közötti nők csoportja, figyelembe véve a családi állapotot, de a gyermek létét nem,
- 20 és 24, illetve 40 és 44 év közötti nők csoportja, figyelembe véve mind a családi állapotot, mind a gyermek létét.

A hierarchikus technika alapját képező autoregresszív modell a következő:

$$y_{c,ms,s,a,t} = \beta_{g,0} + \beta_{g,1}y_{c,ms,s,a,t-1} + \varepsilon_{c,ms,s,a,t} / w_{c,ms,s,a,t} \quad /4/$$

A modellben az $y_{c,ms,s,a,t}$ a munkaerő-piaci részvételi arány értéke, ahol c legalább egy 6 éves vagy fiatalabb gyermek jelenlétére utal, ms a családi állapotra, s a nemre, a a korra és t az időtényezőre utal. $\varepsilon_{c,ms,s,a,t} / w_{c,ms,s,a,t}$ a zaj, ahol w a lakosság súlyának négyzete, amellyel a súlyozás történik. Az alapmodellt a metszéspont és a meredekség finomításának kiterjesztésével bővíti, azaz első körben csak a metszéspont ($\beta_{g,0}$) demográfiai cellákhoz való hozzárendelésével:

$$y_{c,ms,s,a,t} = \beta_{c,ms,s,a,0} + \beta_{g,1}y_{c,ms,s,a,t-1} + \varepsilon_{c,ms,s,a,t} / w_{c,ms,s,a,t} \quad /5/$$

Tovább bővítve, felírható az a modell, ahol a meredekség is demográfiai cellánként eltérő lesz:

$$y_{c,ms,s,a,t} = \beta_{c,ms,s,a,0} + \beta_{c,ms,s,a,1} y_{c,ms,s,a,t-1} + \varepsilon_{c,ms,s,a,t} / w_{c,ms,s,a,t} \quad /6/$$

Azon esetekben, amikor nincsenek időbeli szabályosságok, vagyis minták az idő-sorban, a meredekség nullához közelít, ezért a részvételi arányok változásainak előrejelzett értékei a metszéspontok lesznek. Mivel a metszéspontok cellánként változnak, mutatják az egyes cellák múltbéli alakulásait.

A következő lépcsőben a zajsint szétbontása következik aszerint, hogy az egyes, korábban említett aggregálási szinteken milyen tényezőket vesznek figyelembe. Tehát az egyes szinten a családi állapotot és a gyermek jelenlétét nem vesszük figyelembe, így a zaj:

$$\varepsilon_{s,a,t}^{(1)} = \varepsilon_{c,ms,s,a,t}, \quad /7/$$

ahol $\varepsilon_{s,a,t}^{(1)}$ minden aggregációs szinten közös, amelyhez hozzáadódik a reziduális tag (τ):

1. szint: $\varepsilon_{s,a,t}^{(1)} = \varepsilon_{a,t}^{(2)} + \tau_{s,a,t}^{(3)}$, ahol a családi állapot (ms) és gyermek jelenléte (c) nincs figyelembe véve,

2. szint: $\varepsilon_{ms,s,a,t}^{(2)} = \varepsilon_{a,t}^{(1)} + \tau_{ms,s,a,t}^{(4)}$, ahol a gyermek jelenléte (c) nem meghatározó tényező,

3. szint: $\varepsilon_{c,ms,s,a,t}^{(3)} = \varepsilon_{a,t}^{(1)} + \tau_{c,ms,s,a,t}^{(5)}$, ahol minden tényező részt vesz a csoportképzésben.

A csoportosítások alapján számolt öt variancia-paraméter közül kettő negatív eredményt hozott több módosítás és tisztítás után is, ami nem lehetséges. Ezt a szerző arra vezette vissza, hogy a modell nem képes magyarázni a karrierjük csúcán levő férfiak változékony munkaerő-piaci mozgásait, amely heteroszkeaszticitást eredményezett a modellben. Így a hierarchikus modell nem bizonyult hatékonynak a foglalkoztatás előrejelzésében, éppen ezért, az itt tapasztalt problémák orvoslására alkalmazta a szerző a következő alfejezetben bemutatott longitudinális modellt (Frees [2003]).

1.3. Longitudinális modell

Frees [2003] demográfiai cellák ötletét más modell keretében is hasznosította, amely az egyes cellákat nem csak külön-külön képes vizsgálni, hanem egymás függ-

vényében is, azaz a korábban felírt hierarchikus modellt megkísérelte továbbfejleszteni. Minden demográfiai cellára felírható egy regressziós egyenlet, amelyeket a SUR- (seemingly unrelated regression – látszólag független regressziós) modell segítségével fűz össze. A SUR lényege, hogy sok-sok egymástól függő regressziós egyenletet fűz össze az általános legkisebb négyzetek módszerével. Mivel a 101×101 mátrix esetében nehézkes lenne ilyen nagyszámú regressziós egyenletet kezelni, így a fent említett három aggregációs szint felhasználásával hozza létre az egyenletet annyi módosítással, hogy a nőket és férfiakat nem kezeli egy csoportként, tehát az 1-es és 2-es csoportot különválasztja férfiak és nők szegmensére. A SUR-modell alapján így létrehozott longitudinális modell a hierarchikus modellben leírt három AR-egyenletet használja fel:

$$y_{g,i,t} = E \quad y_{g,i,t} + \varepsilon_{g,i,t} / w_{g,i,t}, \quad /8/$$

ahol E $y_{g,i,t}$ a /4/, /5/, /6/ egyenletek értékei és i a csoportokban (g) fellépő reprodukciót jelöli, amely a csoportokon belüli kor, családi állapot, nem és gyermek jelenléte alapján kialakuló variációkat mutatja. A modell becslése két lépcsőben történik az általános legkisebb négyzetek módszerével. Elsőként a regressziós paraméterek és reziduálok becslése történik meg. Második lépésben a /8/ modellt írja fel a kovariancia-struktúra alapján. Az így létrehozott modellel a demográfiai cellák külön-külön és egymással összefüggésben is vizsgálhatók, kihasználva a szomszédos cellák közötti információt (Frees [2003]).

1.4. Makroökonómiai modellek

Ezek olyan dinamikus makroökonómiai modellek, amelyek egyrészt a kereslet és kínálat egyenlőtlenségeit vizsgálják, másrészt figyelembe veszik az ágazatok közötti dinamikát és iparági szinten adnak becslést a keresletre az ökonometria eszköztárának felhasználásával. Ezek a modellek az egyes iparágakon belül foglalkozásonként, majd ezen belül képzettség szerint is alkalmasak további előrejelzésre. Számos külföldi modell szolgál példaként a munkaerőpiac makroszintű modellezésére, úgy, mint például az ausztráliai Monash Egyetem, az amerikai BLS, az ILO, az OECD, a holland ROA, a német Ifo kutatóközpont modellje.

Az MTA KRTK KTI 2009-ben hozott létre hasonló modellt a munkaerőpiac előrejelzésére. A felépített modell egy több iparágat vizsgáló, több részből felépülő makrogazdasági modell, mely előrejelzi a GDP várható alakulását és ezt összekapcsolja a munkakeresleti és munkakínálati modellel, majd kiegészíti az így kapott adatokat nemzetközi és haza munkaerő-piaci trendekkel, illetve elemzi a kereslet és

kínálat egyensúlytalanságának okait is. Az MTA által létrehozott modell nemcsak statisztikai célú adatgyűjtések adatait tartalmazza, hanem adminisztratív adatokat is. A modell munkakeresletet előrejelző része egyrészt vállalkozási adatokra támaszkodik, amely a Nemzeti Adó- és Vámhivatal mérlegbeszámolóiból származik. A munkakereslet csoportonkénti (foglalkozások, nemek, régiók, iskolai végzettség) előrejelzéséhez a Nemzeti Foglalkoztatási Szolgálat bértarifa-felvétele szolgáltatja a részletes munkavállalói adatokat. A munkakínálati előrejelzéshez szükséges a munkakínálat meghatározása, amelyhez számos adatforrás használatos, úgymint a népszámlálás, a KSH Munkaerő Felmérése, az Életpálya-adatfelvétel, amely az általános iskola elejétől a középiskola végéig nyilvántartja a diákok pályafutását. Továbbá használatosak hatósági adatforrások is, amelyek összekapcsolhatók (Országos Egészségbiztosítási Pénztár, Országos Nyugdíjbiztosítási Főigazgatóság, Magyar Államkincstár). Az előrejelzés a tíz vizsgált szektor GDP-értékének meghatározásával kezdődik, illetve az ezen kibocsátás eléréséhez szükséges munkaerő nagyságának ágazati szinten történő meghatározásával. A munkakeresleti rész finomítja az előbb kiszámított ágazati munkaerő szükségletet vállalati adatok alapján. A munkakeresleti rész továbbbontja az ágazatokat alágazatokra, ami így még részletesebben mutatja az egyes ágazatok közötti különbségeket. A munkakeresleti oldal további feladata, hogy az előrejelzéseket megadja nemek, korcsoportok, foglalkozások, iskolai végzettség és régiók szerint is csakúgy, mint a munkakínálati oldal. A modellben külön blokkban szerepel a kereslet és kínálat eltérése változásainak előrejelzése, amely azért fontos, mert a kereslet és kínálat oldalán nyers előrejelzések szerepelnek, viszont a kiegészítő modulban számos olyan tényező van, amelyek nem elhanyagolható szerepet töltenek be a kereslet és kínálat nagyságának különbözőségében. A kiegészítő modul tartalmaz többek között nemzetközi trendeket, vállalkozások várakozásait, a vállalkozási méretek foglalkozásban betöltött szerepének vizsgálatát, atipikus és be nem jelentett foglalkozás vizsgálatát (*Bakó–Cseres–Gergely–Galasi [2013]*).

Az MTA KRTK KTI előrejelző rendszere volt az első olyan, hatalmas munkát igénylő hazai munkaerő-piaci előrejelző rendszer, amely adatforrások ilyen széles körét használta fel egy gazdasági összefüggéseket is figyelembe vevő modellben. A modell felépítése is kedvező, ugyanis modulszerű kialakítása révén a GDP, a kereslet és kínálat külön-külön történő előrejelzésére is alkalmas, ugyanakkor egymás kontrolljaiként is használhatók. Hátrányaként említhető, hogy sok megkötéssel és kompromisszummal él a modell, úgymint például a két szektor GDP előrejelzésére alkalmas modell használata a tíz szektor adatainak előrejelzésére. Előfeltevéssel él a modell arra vonatkozóan, hogy csak a súrlódásos munkanélküliség létezik, ugyanakkor fel is hívják a figyelmet ezen előfeltevés korlátaira. Továbbá a munkakínálat 2009 és 2020 közötti előrejelzése során azt feltételezi a modell, hogy a szakma szerinti megoszlás a korcsoport, nem és régió szerinti kategóriákban változatlan marad a vizsgált időszakban. Összességében a modell előnye, hogy adatok széles körét használja fel

és sok szempontot figyelembe vesz, néha túl sokat is, ugyanakkor sok feltételezéssel él, így kérdéses lehet, hogy ezek mennyire befolyásolják a modell megbízhatóságát.

1.5. Neurális hálóok

A neurális hálóok elmélete abból indul ki, hogy egy-egy jelenséget meghatározó tényezők közötti összefüggések nagyon nehezen térképezhetők fel, és folyamatosan változnak, illetve ahogy a változások bekövetkeznek, úgy a jelenség vizsgálatára addig felépített modellünket módosítjuk, azaz tanulunk. A neurális hálóok elméletének kidolgozása az emberi viselkedés, tanulási folyamatok, észlelés és az agy működésének tanulmányozásával kezdődtek és később terjedt el használatuk más tudományterületeken is, mint például a gazdasági elemzések területén. Az eddigi előrejelzési modellekben problémaként jelentkezik, hogy nagyszámú bemeneti adattal rendelkezünk, így nagy a hiba lehetősége. A neurális hálóok alkalmazásának elterjedése és népszerűsége ahhoz köthető, hogy az összefüggések feltárását a rendszerre bízunk (Benedek [2000]).

A neurális hálóok rendszerének bemenetei és kimenetei vannak, egy tanulási folyamat révén a bemeneti vektorok alakulnak át kimeneti vektorokká. A hálózat neuronokból áll, amelyek rétegeket alkotnak. Ezek a rétegek három részből épülnek fel. A bemenet és a kimenet között helyezkedik el a közbenső réteg (Kristóf [2005]). A közbenső, más néven rejtett réteg, a háló bonyolultságát és tanulási képességét mutatja meg, továbbá ez felelős az inputokból érkező információ dekódolásáért (Benedek [2000]).

A bemenet és a közbenső réteg között kapcsolat van, amelyet a bemenet fontossága szerint súlyoznak (Kristóf [2005]). A súlyok a tanulási folyamat eredményei és folyamatosan változnak (Patuelli et al. [2011]).

A neurális hálóok egyik kritikus problémája lehet a túltanulás, az az eset, amikor az adatbázist ismeri ki egyre jobban a neurális háló és nem pedig magát a vizsgált problémát (Kristóf [2005]).

A neurális hálóok a regionális foglalkoztatási előrejelzések során nyertek teret. Patuelli és szerzőtársai [2011] használták a technikát a németországi NUTS 3 szintű régiók munkaerő-piaci előrejelzésére két adatbázis felhasználásával, amelyekben az adatok 1987 és 2004, illetve 1993 és 2004 közötti időszakokra vonatkoztak. A vizsgálatban kétéves periódusban becsülték a növekedést. Mivel a neurális hálóok alapvetően nem kezelik az időbeli korrelációt, ezért kulcsfontosságú volt annak megválaszolása, hogy miként lehet megadni a rendszernek az időbeli sokkok és a német újraegyesítés folyamatos hatásainak kezelését. További nehézséget jelentett, hogy kétéves előrejelzést készítettek, ugyanakkor időbeli sokkok kezelése éves szinten volt szükséges. Erre két modellt használtak, amelyek különbözőképpen kezelik a jelensé-

get. Az egyik modell az éveket dummy változóként kezelte, azaz az 1 értéket kapta az adott év, az összes többi 0 értéket, így csak az adott év játszott szerepet a neurális háló tanulási folyamatában. A másik modellben minden év 0 és 1 közötti értéket kapott, a modell így minden időszakot beazonosított, mivel minden időszak más értéket kapott.

A második modell szerint kezelt időszakokkal további öt modellt vizsgáltak, amely újabb tényezőket vont be a modellbe. Az első modell egy kilenc fokozatból álló indexet tartalmazott az urbanizáció és agglomeráció vizsgálatára. A második modell a napi bérek alakulását vonta be a kereslet/kínálat és a bérek közötti kapcsolat vizsgálatára. A harmadik modell versenyképességi faktorról lett kiegészítve, amely a régiók versenyképességét szektoronként vizsgálta, összehasonlítva a szektorok nemzeti szintű tevékenységével (shift-share analízis). A negyedik modell a harmadiktól annyiban tért el, hogy a körzetek nem a nemzeti szintű teljesítménnyel, hanem a szomszédos régiókkal kerültek összevetésre. Az ötödik modell szintén a harmadik továbbfejlesztett változata volt, amely a harmadik modell szektoronkénti versenyképességi faktorait az általános, minden évre elvégzett shift-share regresszió együtthatóival felszorozta, hogy még tovább pontosítsa a létrehozott modellt. Minden modellt külön vizsgáltak Nyugat- és Kelet-Németországra. Az eredmények alapján azok a modellek működtek kisebb hibával, amelyek a shift-share analízist használták, vagyis azok, amelyek területi összehasonlításokat alkalmaztak (*Patuelli et al.* [2011]).

A szerzők szerint ezt a technikát az különbözteti meg a standard lineáris modellezési technikáktól, hogy alkamasabbak komplex, sok tényező által befolyásolt jelenségek vizsgálatára. Ilyen jelenség a munkaerőpiac is, amelynek előrejelzése igen összetett feladat, hiszen figyelembe kell venni a kereslet-kínálat alakulását, a szektorok viszonyait, jellemzőit, földrajzi tényezőket, intézmények szerepét. A neurális hálók további előnye például a korábban említett BVAR-technikával szemben is, hogy nincs szükség előzetes tudás meglétére, ami előny, hiszen sok esetben kevés előzetes információval rendelkezünk. Ugyanakkor hátrányai is vannak a modellnek. Az egyik az, hogy nehezen formalizálható, azaz a modell elméleti alapokon történő specifikációja nem lehetséges. Sok esetben új modell leírására van szükség, amely az adott vizsgálandó jelenség függvénye, így nehéz elméleti háttérrel kiépíteni a neurális hálók működésére (*Patuelli et al.* [2011]).

1.6. Lineáris programozás

A lineáris programozás a neurális hálók módszeréhez hasonlóan a területi foglalkoztatási adatok előrejelzésére kifejlesztett módszer, mivel a regionális fejlődés okozati kapcsolatát vizsgálja és a területi különbségekre koncentrálna. A modell az ENTROP- (entropy optimizing procedure – entrópiatimalizálási folyamat) mód-

szert alkalmazza, amely az entrópia, azaz a rendszer rendezetlenségének optimalizálására épül. A módszert kifejezetten a munkaerőpiac előrejelzésére fejlesztettek ki, és lényeges jellemzője, hogy rugalmasabban kezeli a vizsgált jelenséget, az elérhető információkat pedig szélesebb körben használja fel, mint a shift-share analízis vagy a regressziós modell (*Blien–Tassinopoulos* [2001]).

Az entrópia minimalizálására irányuló módszerek az input-output (I/O) táblázatok használatára épülnek, amelyeket sok esetben alkalmaznak a gazdasági politikák kidolgozásában és alátámasztásában, mert átláthatóan mutatják az erőforrásokat és felhasználásukat, illetve a termékek és a szolgáltatások áramlását. Az I/O táblázatból bayesi paraméter-bebecsléssel meghatározható egy I/O mátrix, amelyben a priori használjuk az előző időszak I/O tábláját (*Ahmed–Preckel* [2007]).

Blien–Tassinopoulos [2001] a statisztikai módszerek hátrányát abban látták, hogy az előrejelzések során figyelembe kellett venni, hogy az adatok legalább egy év késleltetéssel érhetők el. A szerzők által felállított előrejelzések három, munkaerőpiacot befolyásoló dimenzió mentén készültek. Ezek a dimenziók az iparági struktúra, a régió típusa és a tágabb értelemben vett terület voltak, amelyhez a körzet tartozik. Ezeken belül aldimenziók vannak, amelyek a dimenziót tovább bontják.

Az ENTROP-módszer lépései a következők:

1. a független regionális trendek számítása körzetekre vonatkozóan, azaz a kezdeti \mathbf{R} mátrix megadása: az első lépésben körzeti és iparági munkaerő-piaci előrejelzések készülnek két évre vonatkozóan a múltbeli növekedési arányok alapján. A mátrix elemei a következők:

$$r_{jk}^{99} = m_{jk}^{97} \left(1 + 0 \cdot 5 \ln \left(\frac{m_{jk}^{97}}{m_{jk}^{93}} \right) \right),$$

ahol

- m_{jk} a mátrix elemeit,
- j a régiót,
- k az iparágat jelöli.
- Az indexben szereplő számok az évszámokat jelölik.

2. Iparági trendek, régió és szövetségi államok beépítése, azaz a mátrix sorainak és oszlopainak összegzése. Ennek célja, hogy ne kizárólag a körzetek szintjére, hanem magasabb területi szintre is készüljenek előrejelzések, mivel a körzeti szint csak egy-egy területi részegység fejlődését mutatja, így a körzeti eredmények összevonása nem adhat megbízható képet nagyobb regionális egységek mutatóira vonatkozóan. Három kategóriára (iparágak, régiók, tágabb földrajzi területek) becsülik az alkalmat

zásban állók arányát egyszerű lineáris regresszió segítségével, majd ezek után folytatható le a foglalkoztatás előrejelzése a korábban említett dimenziókra.

3. Rendelkezésre álló információk bevonása: sok esetben elérhetőek iparág- vagy régióspecifikus fejlődési adatok, amelyek hozzáadhatók a modellhez.

4. A mátrix becslése és az eredmények tesztelése, amely \mathbf{X} mátrix \mathbf{R} kezdeti mátrix alapján történő becslését jelenti. A kezdeti mátrix az egyedi körzetek trendjeit mutatja, míg \mathbf{X} az előző két pontban létrehozott megszorításokkal, bővítésekkel rendelkező \mathbf{R} és \mathbf{X} közötti távolságok minimalizálásával jön létre. Ez az optimalizációs folyamat az ENTROP-algoritmus alapján működik, amely egy folyamatosan ismétlődő folyamat során számítja ki, korrigálja és frissíti a mátrix elemeit. Az előrejelzés megbízhatóságának tesztelése a súlyozott átlagos négyzetes hiba számításával történik (Blien–Tassinopoulos [2001]).

Az ENTROP-módszer a regionális kutatásokban ismert entrópia optimalizálását követi. A módszer előnye, hogy sok elemből álló mátrixokat is tud kezelni, a számítások gyorsan és rugalmasan végezhetőek. Blien és Tassinopoulos a módszer hátrányait nem említik, ugyanakkor érdekes lenne a módszer összehasonlítása más technikák által adott előrejelzések hibájával. Kiemelik továbbá, hogy ez regionális előrejelzések során alkalmazható, tehát ez a módszer regionális szintű tervezés esetén lehet hasznos. A leírt módszert Németországra fejlesztették ki, ahol a regionális szintű előrejelzések gyakoribbak, mint az látható volt a neurális hálók Németország munkaerő-piacának előrejelzésére alkalmazott elemzése során is, amely szintén olyan esetekben használatos, amikor a munkaerő-piacot régiók szintjén vizsgáljuk.

*

Mindezek alapján látható, hogy számos modell született a munkaerőpiac vizsgálatára, amelyek sok esetben tudományos kísérletekként foghatók fel a piacot minél pontosabban leíró módszer kidolgozására. Ezek a módszerek olyan próbálkozások, amelyek más modellek hátrányainak kiküszöbölésére törekednek, de újabb, további módosításokat igénylő hátrányokat hoznak létre. Ugyanakkor fontos tapasztalatok vonhatók le belőle. Legfontosabb az, hogy a munkaerőpiacot mindenképpen olyan modellben kell vizsgálni, amellyel figyelembe lehet venni számos befolyásoló tényező jelenségre gyakorolt hatását. Vannak olyan modellek, (mint például a lineáris programozás, neurális hálók), amelyeket eddig kevés kutatásban alkalmaztak, és nem születtek összehasonlító tanulmányok más, többször alkalmazott modellel. Ezek alkalmazása a kevés tapasztalat miatt kevésbé javasolt. Több kutató által használt technika az autoregresszív modell és mozgóátlagolású trend összekapcsolásával (ARIMA) történő előrejelzés, ugyanakkor ezek a modellek akkor vezetnek megbízható eredményre, ha a magyarázóváltozók időben nem változnak túlzottan. A munkaerőpiac rendkívül komplex, sok tényező által befolyásolt jelenség, amelyet erős dinamizmus jellemez. További jellegzetessége,

hogyan fog alakulni maga a piac és az azt befolyásoló tényezők, mint például a képzési szerkezet. Éppen ezért ez a módszer sem javasolt a munkaerőpiac előrejelzésére. A megoldás olyan modellek használata lehet, mint például a bayesi megközelítést alkalmazó vektor autoregresszív modellek, amelyek figyelembe veszik a nagyfokú bizonytalanságot. Továbbá alkalmazhatók lehetnek a munkaerőpiac előrejelzésére a kombinált modellek, úgymint a *Rapach* és *Strauss* által használt modell, amelyben ARDL-modellt használtak az előrejelzés alapjául, de ezek eredményeit különböző technikákkal kombinálták. Az így létrehozott modell figyelembe veszi a munkaerőpiac komplexitását és azt befolyásoló tényezőket.

2. Összefoglaló

A tanulmány a munkaerő-piaci előrejelzések nemzetközi és hazai gyakorlatának eddigi tapasztalatait és problematikáját dolgozza fel, kitérve az előrejelzések elméleti hátterének ismertetésére. A tanulmányban először bemutatásra került, hogy a szakmában eddig milyen tudományos igényű kutatásokat folytattak le az előrejelzés megvalósítására. A tapasztalatok azt mutatják, hogy nincs egységes módszertana a munkaerő-piaci előrejelzéseknek, számos próbálkozás van a jelenséget minél pontosabban leíró modell kialakítására. A legtöbb kutató nem éri be a munkaerőpiac egy kiválasztott indikátorának egyszerű extrapolációjával, hanem összetettebb modelleket vizsgál, amelyek számos tényező bevonásával magyarázzák a munkaerőpiac legfontosabb jellemzőit. A feltárt kutatások alapján megállapítható, hogy a munkaerőre irányuló előrejelzések területén a modell felállítása és magának az előrejelzésnek a technikája kulcsfontosságú kérdés, hiszen nagymértékben befolyásolja a megbízhatóságot, azt, hogy mely tényezőkkel magyarázzuk a munkaerőpiacot és milyen előrejelzési hibával dolgozó statisztikai eszközzel készítjük az előrejelzést. A munkaerőpiacot mindenképpen olyan modellel célszerű vizsgálni, amellyel figyelembe lehet venni több befolyásoló tényezőnek a jelenségre gyakorolt hatását. A megoldást olyan összetett modellek adhatják, amelyek a munkaerőpiac komplexitását és az azt befolyásoló tényezőket is számításba veszik, úgy, mint például a kibocsátást, a képzési szerkezetet, az atipikus foglalkoztatási formák elterjedését, a területi jellemzőket, a demográfiai tényezőket. Ezen kívül még más olyan tényezők is számításba vehetők, amelyek ismerete és modellben történő szerepeltetése kulcsfontosságú az előrejelzés megbízhatósága szempontjából.

A tanulmányban a munkaerő-piaci előrejelzések hazai vonatkozásait is vizsgáltam. Az első komolyabb, összetettebb, gazdasági folyamatok közötti kölcsönhatásokat is vizsgáló modell az MTA KRTK KTI által létrehozott makrogazdasági modell

volt, amely a hazánkban rendelkezésre álló sokféle adatforrást kapcsolta össze. A modell előnye, hogy adatok széles körét használja fel, és sok szempontot vesz figyelembe, néha túl sokat is, ugyanakkor sok feltételezéssel él, így kérdéses lehet, hogy ezek mennyire befolyásolják a modell megbízhatóságát. Ugyanakkor ez a kutatás fontos mérföldkő volt a hazai munkaerő-piaci kutatások elindításában és számos tapasztalat vonható le eredményeiből, mint például az adatforrások összeegyeztethetőségének nehézsége és azok a kompromisszumok, amelyeket az előrejelezhetőség érdekében meg kell kötni.

Irodalom

- AHMED, S. A. – PRECKEL, P. V. [2007]: *A Comparison of RAS and Entropy Methods in Updating IO Tables*. Economics Association Annual Meeting. 29 July–1 August. Portland.
- ASHENFELTER, O. – CARD, D. [1982]: Time series representation of economic variables and alternative models of the labour market. *The Review of Economic Studies*. Vol. 49. No. 5. pp. 761–781. <http://dx.doi.org/10.2307/2297188>
- BAKÓ T. – CSERES-GERGELY ZS. – GALASI P. [2013]: Az MTA KRTK KTI munkaerő-piaci előrejelző rendszere. In: *Fazekas K. – Varga J. (szerk.): Trendek és előrejelzések: Munkaerő-piaci prognózisok készítése, szerkezetváltás a munkaerő-piacon*. MTA-KRTK-KTI. <http://www.krtk.mta.hu/hirek/Megjelent-Trendek-es-elorejelzesek-Munkaerő-piaci-prognózisok-készítése-szerkezetváltás-a-munkaerő-piacon/91/>
- BALAN, A. M. [2012]: Realities and perspectives concerning employment in Romania: Trends and forecast based on the Box & Jenkins methodology. *Economic Insights – Trends and Challenges*. Vol. 64. No. 1. pp. 94–103.
- BENEDEK G. [2000]: Evolúciós alkalmazások előrejelzési modellekben I. *Közgazdasági Szemle*. XLVII. évf. 12. sz. 988–1007. old.
- BLIEN, U. – TASSINOPOULOS, A. [2001]: Forecasting regional employment with the ENTROP method. *Regional Studies*. Vol. 35. No. 2. pp. 113–124.
- BOWERMAN, B. L. [1987]: *Forecasting and Time Series*. Wadsworth Publishing Company. Belmont.
- FREES, E. W. [2003]: Stochastic forecasting of labor force participation rates. *Insurance: Mathematics and Economics*. Vol. 33. No. 2. pp. 317–336. [http://dx.doi.org/10.1016/S0167-6687\(03\)00156-2](http://dx.doi.org/10.1016/S0167-6687(03)00156-2)
- HAMPEL, K. – KUNZ, M. – SCHANNE, N. – WAPLER, R. – WEYH, A. [2008]: Regional employment forecasts with spatial interdependencies. In: *Knobel, C. – Kriechel, B. – Schmidt, A. (eds.): Regional Forecasting on Labour Markets*. Rainer Hampp Verlag. München. pp. 68–88.
- KOVÁCS S. – BALOGH P. [2009]: Bayesi statisztikával becsült nem stacionárius idősorok a sertésárak előrejelzésében. *Statisztikai Szemle*. 87. évf. 10–11. sz. 1059–1077. old.
- KRISTÓF T. [2005]: A csödelőrejelzés sokváltozós statisztikai módszerei és empirikus vizsgálata. *Statisztikai Szemle*. 83. évf. 9. sz. 841–963. old.
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2005]: *Szezonális kiigazítás*. Statisztikai Módszertani Füzetek. 43. köt. Budapest.

- MÁK F. [2011]: Egységgyök-tesztek alkalmazása strukturális törések mellett a hazai benzinár példáján. *Statisztikai Szemle*. 89. évf. 5. sz. 545–573. old.
- NAU, R. [2015]: *Statistical Forecasting: Notes on Regression and Time Series Analysis*. <http://people.duke.edu/~rnau/411home.htm>
- NEUGART, M. – SCHÖMANN, K. [2002]: *Employment Outlooks: Why Forecast the Labour Market and for Whom?* Wissenschaft zentrum Berlin für Socialforschung. Discussion Paper. FS I 02-206. Berlin.
- PATUELLI, R. – REGGIANI, A. – NIJKAMP, P. – SCHANNE, N. [2011]: Neural networks for regional employment forecasts: Are the parameters relevant? *Journal of Geographical Systems*. Vol. 13. No. 1. pp. 67–85. <http://dx.doi.org/10.1007/s10109-010-0133-5>
- PURI, A. – SOYDEMIR, G. [2000]: Forecasting industrial employment figures in Southern California: A Bayesian vector autoregressive model. *The Annals of Regional Science*. Vol. 34. No. 4. pp. 503–514. <http://dx.doi.org/10.1007/s001680000030>
- RAPACH, D. E. – STRAUSS, J. K. [2008]: Forecasting US employment growth using forecast combining methods. *Journal of Forecasting*. Vol. 27. No. 1. pp. 75–93. <http://dx.doi.org/10.1002/for.1051>
- RAPACH, D. E. – STRAUSS, J. K. [2012]: Forecasting US state-level employment growth: An amalgamation approach. *International Journal of Forecasting*, Vol. 28. No. 2. pp. 315–327. <http://dx.doi.org/10.1016/j.ijforecast.2011.08.004>
- RAUPELIENĖ, A. [2011]: *Regional Foresights for Rural Labour Market in EU 27*. Proceedings of the International Scientific Conference: Rural Development. Vol. 5. 24–25 November. Akademija, Kaunas region, Lithuania, pp. 216–221.
- SARANTIS, N. – SWALES, C. [1999]: Modelling and forecasting regional service employment in Great Britain. *Economic Modelling*. Vol. 16. No. 3. pp. 429–453. [http://dx.doi.org/10.1016/S0264-9993\(99\)00009-7](http://dx.doi.org/10.1016/S0264-9993(99)00009-7)
- SEUNG, C. K. – AHN, S. K. [2010]: Forecasting industry employment for a resource-based economy using Bayesian vector autoregressive models. *The Review of Regional Studies*. Vol. 40. No. 2. pp. 181–196.
- TRÍVEZ, F. J. – MUR, J. [1999]: A short-term forecasting model for sectoral regional unemployment. *The Annals of Regional Science*. Vol. 33. No. 1. pp. 69–91.
- WONG, J. M. W. – CHAN, A. P. C. – CHIANG, Y. H. [2005]: Time series forecasts of the construction labour market in Hong Kong: The Box-Jenkins approach. *Construction Management and Economics*. Vol. 23. No. 9. pp. 979–991. <http://dx.doi.org/10.1080/01446190500204911>
- WEI, W. W. [1990]: *Time Series Analysis*. Addison-Wesley Publishing Company. Boston.