**Budapesti Corvinus Egyetem**

Közgazdasági intézet

**Lakásépítés járási szinten – Van a CSOK iránti érdeklődésnek hatása?**

Dittrich Levente

Alkalmazott közgazdaságtan szak

Témavezető: Kovács László

**2024.**

Tartalomjegyzék

[Bevezetés 2](#__RefHeading___Toc995_2628363530)

[Adatok és módszer 3](#__RefHeading___Toc997_2628363530)

[Eredmények 9](#__RefHeading___Toc999_2628363530)

[Diszkusszió 13](#__RefHeading___Toc1001_2628363530)

[Irodalomjegyzék 14](#__RefHeading___Toc1003_2628363530)

## Bevezetés

Magyarországon 2015 óta létezik a családi otthonteremtési kedvezmény (továbbiakban CSOK), amely egy, a családokat saját otthonhoz segítő támogatási forma. A 2016-tól induló Családvédelmi Akcióterv részeként a CSOK támogatja a házas egy, vagy több gyermeket vállalókat új lakás építésében, vásárlásában, meglévő felújításában, bővítésében. Ezen lehetőségek széles spektrumán a támogatás és támogatott hitel is változó, mind a felhasználást, mind a gyerekek számát illetően. Dolgozatomban azonban a leglényegesebb az, hogy az új ingatlan vásárlására vagy építésére 10 millió forint támogatás igényelhető 15 millió forint államilag támogatott hitellel, abban az esetben, ha a párnak van három vagy több gyermeke, illetve ha még nincsen három vagy több, de vállal minimum hármat. A lakásvásárlás és építés fellendítésére a kormány szintén 2016-ban csökkentette 5%-ra a lakásvásárlás áfáját. A későbbiekben ez ki lett vezetve, majd pedig 5 millió forintig lehetett visszaigényelni. Ezen intézkedések után még a babaváró támogatást is bevezették.

Szakdolgozatomban azt kívánom megvizsgálni, hogy a CSOK iránti érdeklődés az interneten (Google keresések képében) és több más változókkal egyetemben milyen hatással van a lakásépítésekre járási szinten a 2012-2021 közötti tíz évben, így nézve a CSOK előtti éveket is. Különös figyelmet szentelek a területi autokorreláció szűrésére, a standard modellfeltételeknek eleget téve. Ezen felül módomban áll a CSOK mellé bevont kontrollváltozók hatását megvizsgálni a lakásépítésre és egymásra.

**Téma relevanciájának indoklása és szakirodalmi források**

A családi otthonteremtési kedvezmény a magyar hitelezés és ingatlanpiac egy maghatározó fogaskereke, nem is beszélve a családokról, akik ezáltal jutnak új otthonhoz vagy esetleg egy használt ingatlant vásárolnak, újítanak fel. A közpolitikai intézkezdésről bevezetése óta több kutatás, publikáció született. *Dancsik és szerzőtársai* (2022) a kamattámogatott hitelek túlárazottságára hívják fel a figyelmet a piaci hitelekkel szemben, ami adódhat a megfelelő ösztönzők hiányából fakadóan. Egy, a Statisztikai Szemlében megjelent tanulmányban a CSOK költségvetési terheit becsülik meg a jövőre nézve és ennek fenntarthatóságát változatlan makrogazdasági pálya esetén (*Tatay et. Al, 2019*).

A magyar tudományos életben a területi panel és keresztmetszeti modellek, elemzések is rendre gyakoriak, Varga Attila 2002-es Statisztikai Szemlében megjelent cikke nagy hatással volt a hazai ökonometriával foglalkozó közgazdászokra. A magyar térökonometria meghatározó alakja a fent említett cikkében magyar nyelven mutatta be a területi autokorreláció fogalmát, a szomszédsági mátrixokat, Moran I-jét, a térbeli autokorrelációt lineáris és ökonometriai modellekben. Békés és szerzőtársai (2016) járási és települési szinten vizsgálták az átlagos lakóingatlanárakat. Kutatásukban azonosítottak gócpontokat és területi jellemzőket is, azonban a területi autokorreláció jelenségével nem foglalkoztak.

Az utóbbi években a magyar tudományos életben megszaporodtak a Structured Equations Model-eket használó publikációk, ami megfigyelt és nem megfigyelt változók közötti feltételezett kapcsolatrendszer ellenőrzésére szolgáló módszertan. Vékás és Vakhal (2021) az KKE (Kelet-Közép-Európa) régió országainak árnyékgazdaságának méretét becsülte meg SEM segítségével, Kovács és szerzőtársai (2021) a fent említett modellel olyan egyenletrendszert hoztak létre, amivel a Corvinus gazdaságinformatikus hallgatók egyes tárgyakból szerzett érdemjegyei, neme és évfolyama közötti összefüggéseket tárták fel. Szintén sikeresen alkalmazta Aranyossy és Recskó (2021), akik a kriptovaluták elfogadottságát vizsgálzák UTAUT fogalmi keretben. Gosztonyi (2021) azon kutatásában is fontos szerepe van a SEM-nek, amiben a hazai középvállalkozások fenntarthatóságát tanulmányozta.

A forráskutatás során nem találtam olyan tanulmányt, amely területi alapon vette volna a CSOK hatását a lakásépítésre más változókkal egyetemben, ahogy olyan forrást sem találtam, ahol a lakásépítéseket magyarázó változók közötti kapcsolatokat térképezték volna fel.

A XXI. században a tájékozódás egyik legfontosabb eszköze az internet, jogszabályok, hírek és bejelentések tömkelegével nap mint nap bővül a rendelkezésre álló információ. Jogos lehet így a feltételezés, hogy a magánszemélyek jelentős része első körben az internetről tájékozódik az állami támogatásokról, közpolitikai intézkedésekről, így a családi otthonteremtési kedvezményről is. Úgy vélem, hogy fontos lehet megvizsgálni, hogy a lakásépítésekhez hogyan kapcsolódik a CSOK iránti internetes érdeklődés, illetve a járás szintű adatokon ez hogyan vetül le a többi változó mellett. A modell építése során a fellelhető területi változók felhasználásával és szelektálásával igyekeztem olyan modellt alkotni, amely a CSOK iránti internetes érdeklődést elhagyva is logikus, magas magyarázó erővel bíró legyen.

Kutatási kérdéseim tehát a következők:

- Van-e hatása a CSOK iránti internetes érdeklődésnek a lakásépítésekre járási szinten?

- Milyen kapcsolat van a CSOK mellett használt lakásépítésekre ható változók között?

## Adatok és módszer

Dolgozatom során főleg a KSH-tól szerzett adatokkal dolgozom, pontosabban a KSH Területi atlasz – Térképes Interaktív Megjelenítő Alkalmazás nevezetű, röviden TIMEA (KSH, 2023) online programon, bárki által elérhető járási és megyei szintű adatokkal. A másik adatforrás, amivel dolgoztam a Google Trends (Google, 2023) szolgáltatása, amin a Google keresések trendjének volumenének adatai érhetők el, akár megyei szinten is.

Eredményváltozóm tehát a tízezer főre jutó lakásépítések száma (LAKAS), ami azt jelenti, hogy az adott járásban az adott évben tízezer főre ennyi használatbavételi engedélyt bocsátottak ki lakásokra és házakra. Azzal, hogy ez arányosítva van a járás lakosainak számával, remekül összehasonlíthatóak a járások eltérő lakosszám esetén is. Magyarázó változóim az egy állandó lakosra jutó személyi jövedelemadó alapot képző jövedelem járási szinten (SZJA), az adott járásban az átlagos ingatlanár millió forintban (ATLAGAR), a százezer lakosra jutó regisztrált sértettek (bűncselekményből adódóan) száma (SERTETT), egy pedagógusra jutó tanulók száma a járásban (PEDAGOGUS) a foglalkoztatottsági ráta megyei szinten (MUNKA), beruházási teljesítményérték egy lakosra vetítve megyei szinten (BEUHAZAS) és a “CSOK” szó keresési trendje Google-n szintén megyei szinten (CSOK ill. ez abből eredő eltérő névpermutációk a különböző modellek között). Magyarázó változóim közé bevettem egészségügyi változókat a KSH *Éves településstatisztikai adatok 2022-es településszerkezetben nevű platformjáról*, szintén arányosítva a lakossághoz.

Az adatok egy részéből főkomponenseket képeztem főkomponens-elemzés segítségével, mivel erősen korrelálnak egymással. Az első főkomponens a háziorvosok rendelőn belüli forgalmát (HO\_FORG\_RB), összes forgalmát (HO\_FORG\_OSSZ), felnőtt és gyermek részére szervezett háziorvosi szolgálatok számát (FGYHO\_SZOLG\_SZAM), háziorvosok szolgálatának számát (HO\_SZOLG\_SZAM), házi gyermekorvosok számát (HGYO\_SZAM) és az ápolók számát (HO\_APOLO\_SZAM). Ezek az adatok tízezer főre vannak vetítve, hogy jobban összehasonlíthatók az eltérő lakosszámú megyék. A második főkomponensben, szintén tízezer főre vetítve a házi gyermekorvosok száma (HGYO\_SZAM), ápolók száma (HO\_APOLO\_SZAM) és a háziorvosi szolgálatok száma (HO\_SZOLG\_SZAM) szerepel. Egyik főkomponensbe sem került bele a Védőnői dolgozók betöltött állásainak száma 10 ezer főre vetítve (VEDONO), mivel a modellépítés és a főkomponens elemzés során arra jutottam, hogy külön változóként érdemes kezelni.

A vizsgált tíz év egy viszonylagos prosperitási időszak, gazdasági értelemben konjunktúra, életszínvonal növekedés volt tapasztalható (Nagy et al., 2016). A CSOK bejelentése után egyre nagyobb arányban szerepelnek a lakáshitelek között a támogatott lakáshitelek (Dancsik et al. 2022), viszont nagy részüket használt lakás vásárlására vették fel. (MNB, 2022)

*1.táblázat: a lakásépítések leíró statisztikai mutatói*

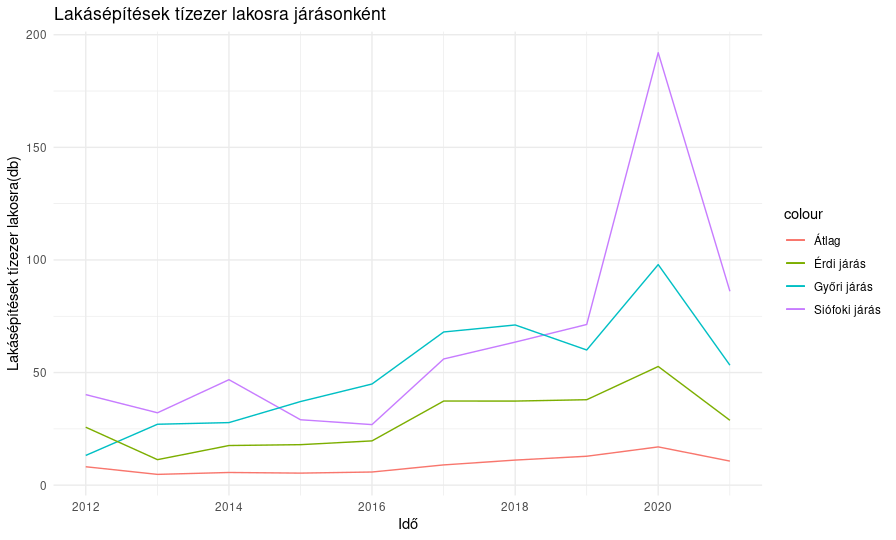
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Átlag | Medián | Min | Max | σ | α3 | α4 |
| 2012 | 8,18 | 5,27 | 0 | 62,2 | 9,47 | 2,38 | 7,18 |
| 2013 | 4,79 | 2,67 | 0 | 32,89 | 5,97 | 2,63 | 8,23 |
| 2014 | 5,63 | 3,2 | 0 | 46,84 | 7,02 | 2,53 | 8,12 |
| 2015 | 5,35 | 3,22 | 0 | 37,13 | 6,47 | 2,16 | 5,03 |
| 2016 | 5,84 | 2,86 | 0 | 46,78 | 8,28 | 2,59 | 7,45 |
| 2017 | 9 | 3,7 | 0 | 85,7 | 13,28 | 2,85 | 9,89 |
| 2018 | 11,14 | 4,94 | 0 | 95,52 | 16,11 | 2,43 | 6,25 |
| 2019 | 13,62 | 7,33 | 0,31 | 76,85 | 16,48 | 1,89 | 3,22 |
| 2020 | 18,13 | 8,18 | 0,2 | 192,03 | 25,26 | 3,03 | 13,74 |
| 2021 | 11,2 | 4,32 | 0,2 | 86,09 | 15,68 | 2,2 | 4,88 |

*Forrás: KSH, saját számítás*

A tízezer főre jutó lakásépítések valóban emelkedtek az évek során 2020-ig, ahogy az *1. ábra* is mutatja. 2020-ban tetőzött a használatbavételi engedélyek száma az országban, majd 2021-ben már sokkal alacsonyabb volt a lakásépítések száma, a 2019-es érték alá esett több helyen és az átlagos lakásépítésekben is.

Az *1. ábra* azonban mást is mutat: az átlaghoz viszonyítva egy általam választott Balaton melletti, egy észak-nyugat magyarországi és Budapest agglomerációjába tartozó járásban sokkal kiemelkedőbb a lakásépítések száma. Megvizsgálva az adatokat térben egy adott időpontban, mint például 2018-ban valami érdekes jelenség fedi fel magát. A *2. ábrán* láthatóak úgy nevezett “hotspotok”, amik a területi adatoknál fordulhatnak elő.

Az országban három nagyobb gócpontot tudtam azonosítani, az első Budapest és agglomerációja, amivel az M7-es autópálya nyomvonalán összeér a második gócponttal, Balaton-melletti járásokkal. Magyarországon regionális szinten eltérőek az egy főre jutó bruttó GDP adatok, például 2019-et nézve Budapesten 10 048 Ft bruttó hazai termék jutott egy főre, Győr-Moson-Sopron vármegyében 5 525 Ft, Szabolcsban pedig 2 857 Ft. (KSH, 2022) Látszik, hogy óriási különbségek vannak régiónként és Budapest a legkiemelkedőbb. Az átlagkeresetek is tükrözik ezeket a szakadékokat, Budapesten ugyanebben az évben a bruttó átlagkereset 460, Győr-Moson-Sopron vármegyében 389, Szabolcs-Szatmár-Bereg vármegyében 252 ezer forint volt. (KSH, 2022) Észszerű döntés lehet tehát a munkavállalók részéről Budapesten munkát vállalni és ott vagy az agglomerációban házat vagy lakást venni, építeni.

*1.ábra: A lakásépítések trendje*

*Forrás: KSH, saját számítás*

A második gócpont a Balaton melletti, a Székesfehérvári és Gárdonyi járások. A Balaton mellett az utóbbi évtizedben jelentősen fellendültek a nyaralóépítések, beruházások, és az ottani, állandó lakosok száma nem olyan nagy, ám lehet, hogy sokan szeretnének ott nyaralót, ezért a tízezer főre jutó lakásépítések száma hatalmasra is ugorhat. Székesfehérvári és Gárdonyi járások esete egy picit másabb, több hatás is közre játszhat. Közel vannak a Velencei tóhoz, valamint ott halad el az M7-es autópálya, így viszonylag gyorsan el lehet jutni Budapestre vagy a Balatonhoz.

A harmadik “hotspot” eggyel különlegesebb, mivel nem egy paca-szerű megjelenése van, a Győri, Mosonmagyaróvári, Soproni, Kapuvári, Kőszegi, Szombathelyi, Sárvári és Szentgotthárdi járásokat foglalja magába. Az legtöbbjük közelsége az osztrák határhoz igencsak szembeötlő, illetve Győr pedig egy nagyon fejlett, prosperáló város saját egyetemmel, rengeteg beruházással.

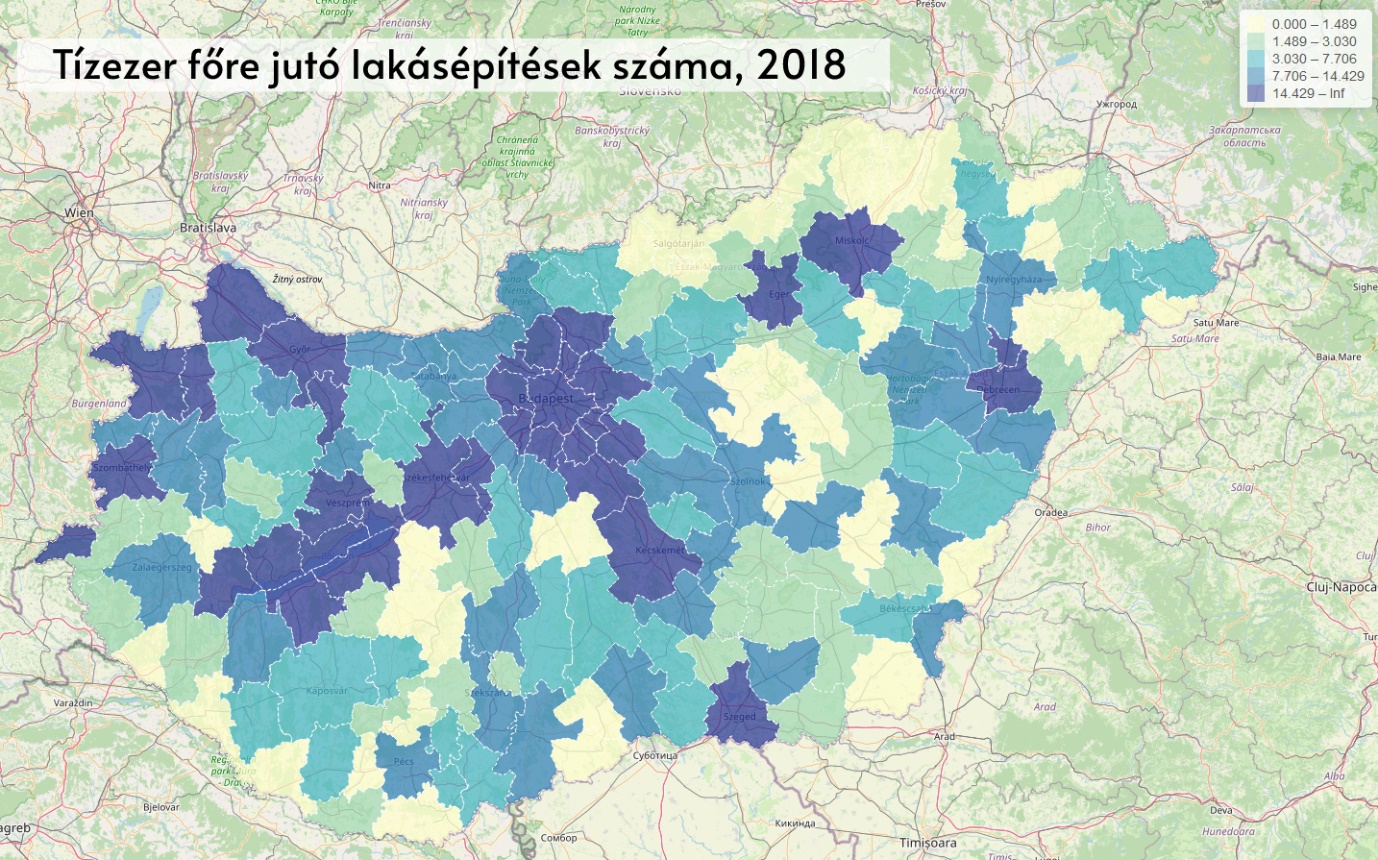
Vannak kisebb kiemelkedő pontok, ezek megyei jogú városok járásai, mint Miskolc vagy Debrecen, de ezek egyedi, sokkal kisebb pontok.

Az idősoros adatoknál fordulhat elő az autokorreláció jelensége, ez azt jelenti, hogy az idősor elemei nem korrelálnak egymással. A korreláció és parciális korreláció is az idősor Yt és Yt-k elemeire nem tér el szignifikánsan 0-tól ∀k > 0-ra (t,k ϵ N). A területi adatok térben egymás mellett helyezkednek el, így ezek is korrelálhatnak térbeli dimenzióból nézve egymással, ez a térbeli autokorreláció. A térbeli autokorreláció mérésére vannak különböző mutatók, ám először meg kellene határozni, hogy mely járások szomszédosak. Észszerű megoldás lenne egyszerűen összeszedni, hogy mely járásoknak van szomszédos határa melyikkel. Alternatívaként azt is meg lehetne nézni, hogy a járások geometriai középpontjai egy megadott távolságon belül helyezkednek el egymáshoz képest. A szomszédság definiálásához én inkább egy inverz távolságalapú (euklideszi), sorstandardizált súlymátrixot fogok használni, mert logikusabbnak tűnik az, hogy a távolságoknak nagyobb jelentőségük van, mint az effektív határoknak. Magyarországon rengetegen ingázhatnak napi szinten munkába, és ha így lenne, akkor gyakorta járhatnak szomszédos vagy egyel messzibb járásokba is. Véleményem szerint a távolság, amitől egyes járásokat szomszédosnak lehet tekinteni 30-40 km között lehet, ez körülbelül elképzelhető az ország méretéből és a járások számából, nem is beszélve arról, hogy ez akár mindennapi ingázási távolságnak sem nagy. Ahhoz, hogy meg tudjam állapítani az optimális távolságot, hogy mit tekintek egy járás szomszédjának kellenek az autokorrelációs mutatószámok. Moran I-je és Geary C-je is bevett autokorrelációs mutató, amiket kiszámoltam több vizsgált évben és távolsággal.(Varga, 2002)

Moran I-je a [-1;1] tartományban vehet fel értéket, hasonlóan a Pearson-féle korrelációhoz, minél nagyobb az abszolút távolsága I-nek a nullától annál erősebb az autokorreláció, az előjelek pedig hasonlóan, ha I negatív, akkor negatív irányban van autokorreláció, ha pozitív, akkor pozitív irányban. Moran I-jének képlete, amennyiben sorstandardizált W súlymátrix áll rendelkezésre:

Geary C-je egy 0 és egy 1-nél nagyobb, nem meghatározott szám között lehet. C értéke 1 alatt pozitív autokorrelációt jelent, míg szignifikánsan felette negatív autokorrelációt jelent. (Varga, 2002)

Geary C-je az alábbi formulával áll elő:

*2. ábra: Tízezer főre jutó lakásépítések száma térképen, 2018*

*Forrás: KSH, saját számítás*

Számításaim során panel adatokkal dolgozom. Ez azt jelenti, hogy a megfigyeléseknek van keresztmetszeti és időbeli dimenziója is, azaz a 175 megfigyelt járásról tíz éven keresztül minden évben nagyrészt van adat. Ezzel a módszerrel nem csupán 175 megfigyelésem van, hanem egyenesen 1750, így a modellek és próbák szabadságfokok nagyobbak lehetnek.

Panel adatoknál felmerül a kiegyensúlyozottság kérdése, ez azt jelenti, hogy minden megfigyelt időpontban van megfigyelés az adott entitásról. Azonban az adataim között vannak hiányzó értékek, szám szerint 21 darab adat a 2019-2021 közötti időszakban. Ennek orvoslására egy Generalizált Additív Modellel (továbbiakban GAM) becslem meg a hiányzó adatokat. A modellben a lakásépítések számát egyedül az járás koordinátái (hosszúság, szélesség) és az adott évvel való interakciójuk alapján becslem meg és pótlom ki a hiányzó adatokat. A GAM modell egy olyan generalizált lineáris modell, amely bizonyos base-spline függvényekből tevődik össze szakaszosan. Így a modell fel tud venni olyan, nemlineáris formát is, amiben inflexiós pontok vannak. A b-splineok, lényegében polinomiális függvények, amik egyenként egy kis területet fednek le, így ezek sokaságából összeáll egy nagyobb tartomány. Ezek a b-splineok teszik lehetővé, hogy a modell térben ne egy sík legyen, hanem egy olyan felszín, melyen vannak kiemelkedések és süllyedések. (Wood, 2003) (Hastie-Tibshirani,1990)

A GAM általános egyenlete:

,

* Ahol Y egy olyan vektor, aminek a elemei egy exponenciális eloszláscsaládból származó valószínűségi változó megfigyelt értékei, E[Y] pedig Y várható értéke
* h(·) a GAM link függvénye
* [Ɛ a modell hibavektora](https://en.wikipedia.org/wiki/Latin_epsilon)
* Xj a *p*  darab megfigyelt magyarázó változó értékei
* fj(·) a j-edik magyarázó változóhoz tartozó transzformációs függvény

Mivel egyedül azért használtam a GAM modellt, hogy a hiányzó adatokat pótoljam, ezért nem tartom szükségszerűnek a modell további bemutatását. A kipótolt adatokkal azonban már kiegyensúlyozottnak tekinthető a paneladattáblánk, mivel minden megfigyelt egyedhez tartozik megfigyelés minden releváns időszakban, így 197 járás 10 évre visszamenőleg 1970 db adatpontot jelent.

A paneladatok elemzése többféle módon is történhet, gyakran használnak *Pooled OLS*, *fix hatású* és *random hatású* modelleket. A *Pooled OLS* egy olyan OLS becslés, amiben összeöntve vannak az adatok. Ekkor minden t időpontban ugyan az a modell érvényes. A *fix hatású* modellben van egyedhatás, vagyis a megfigyelt egyedeknek van egy egyed-specifikus alap konstans hatása. A *random effect* modellek abban különböznek a *fix hatasú* modellektől, hogy az egyéni konstansokat egy normális eloszlású valószínűségi változó megvalósult értékeinek tekintjük. Felmerülhet még *dinamikus panelmodellek* használata is, amik olyan fix vagy random effect modellek, amikben plusz változóként az előző időszakbeli eredményváltozó értéke is szerepel, azaz lényegében egy tetszőleges késleltetésű AR(i), vagyis autoregresszív tag lesz az egyenletben. Azonban ezt nem tartom feltétlenül szükségesnek, hiszen csupán 10 darab időszakot vizsgálok, így habár lehet valamiféle trend az eredményváltozóban, a tíz időszak azonban túl kicsi időszak, arról nem is beszélve, hogy az elején nem volt CSOK, a végén pedig már a covid-válság évei vannak.

Ahhoz, hogy tudjam melyik modell válasszam, két tesztet is el fogok végezni, a Hausman-teszt, ami a *fix* és *random* *hatású* modell között segít, valamint egy F teszt a *Pooled OLS* és a *fix hatású* modell között segít dönteni. (Wooldridge, 2010)

A CSOK modellbe illesztésénél felmerül az a probléma, hogy a vizsgált időszak első négy évében végig nullának vettem a CSOK keresési trendjét), mert 2016 januárjában hozták meg a kormányrendeletet a CSOK-ról (a 2016-2021-es időszakban minden adat nagyobb volt nullánál). Ez esetben úgy gondolom érdemes lehet nem számként, hanem felbontani az egyes kvantilisek mentén és a nulla értékek egy új “nincs” nevű kategóriát kapnának, mert akkor lényegében nem volt CSOK. A modellszelekció során megpróbálom meghatározni az optimális CSOK használatot, ám mivel szakdolgozatomban azt igyekszem kideríteni, hogy szignifikáns változó-e a “CSOK” szó Google keresési trendje, ezért akár az is lehet, hogy egyáltalán nem szignifikáns semmilyen formában.

A modell illesztésénél felmerült az a probléma is, hogy egyes változók korrelálnak egymással, a panelmodellek után így készítettem *Structured Equation Model*-eket (továbbiakban SEM). Ezek a modellek egyenletrendszereket hoznak létre, így jobban vizualizálhatóak a kapcsolatok az egyes változók között amellett, hogy a szignifikáns és nem szignifikáns közvetlen és közvetett hatásokat be lehet azonosítani, éppen az egyenletrendszer mivolta miatt. Ez egy olyan módszertan, amivel a megfigyelt és nem megfigyelt változók között feltételezett kapcsolatrendszert lehet ellenőrizni. (Lomax, 2013) A modellben létre lehet hozni látensváltozókat, így megfigyelt változókból egy vagy több kvázi főkomponenst lehet létrehozni (faktormodellel), ami ez esetben az egészségügyi fejlettségét jelenti a járásoknak. A látensváltozóba az első főkomponens elemeit tettem a haznos infromáció összesűrítése érdekében. Ezt a látens változót az egészségügy fejlettségének tekintem az adott járásban az alapján, hogy a főkomponens elemzésben milyen következtetésekre jutottam. Három, különböző időszakra illesztettem SEM modellt: egy, a CSOK előtti évben, a CSOK kezdetén, illetve a CSOK csúcsán, de a COVID időszak előtt.

## Eredmények

Érdemes lehet a korrelációk alapján megnézni, hogy milyen mértékben és irányban mozognak együtt a változók. Az eredményváltozó és mind az SZJA, mind a munka változók között pozitív, közepes kapcsolat áll fenn. a beruházás, valamint a CSOK (számként) korrelációja a lakásépítéssel gyenge, pozitív irányú.

A GAM modell egyedül az adatok pótlására szolgált, így azon nem végeztem teszteket, sem modellszelekciót, így nem tartom szükségesnek tárgyalni a modell együtthatóinak értékeit és egyéb jellemzőket.

Miután a hiányzó értékeket megbecsültem a GAM modellel, az eredeti értékeket felhasználva meg kellett állapítanom, hogy van-e területi autokorreláció és mekkora távolságban határozzam meg a szomszédságot.

*2.táblázat: térbeli autokorreláció a lakásépítések körében*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 32 km | | 36 km | | | 40 km | |
|  | Moran I | Geary C | | Moran I | Geary C | Moran I | Geary C |
| 2016 | 0,237 | 0,709 | | 0,295 | 0,683 | 0,247 | 0,731 |
| 2017 | 0,188 | 0,729 | | 0,280 | 0,652 | 0,249 | 0,698 |
| 2018 | 0,361 | 0,633 | | 0,381 | 0,605 | 0,351 | 0,637 |
| 2019 | 0,392 | 0,593 | | 0,412 | 0,560 | 0,367 | 0,630 |
| 2020 | 0,300 | 0,679 | | 0,319 | 0,642 | 0,290 | 0,701 |

*Forrás: KSH, saját számítás*

A *2. táblában* minden évben látszik, hogy az I értéke a 36 kilométeres sugarú körben vett inverz, sorstandardizált súlymátrix esetében a legnagyobb. Ez nem csak azt jelenti, hogy van pozitív területi autokorreláció, hanem azt is, hogy az autokorreláció szempontjából legjobb, ha 36 kilométeres sugárban tekinthetők szomszédosnak a járások. Ahogy Moran I-jénél, úgy a C-nél is a legnagyobb pozitív autokorreláció a 36 kilométeres sugárban van minden évben. Ezek alapján a modellemben az autokorreláció kezelésére az optimális inverz távolságalapú súlymátrix 36 km-es sugárban tekinti szomszédoknak a járásokat. Ez a távolság az ország méreteihez és a járások számához képest teljesen elképzelhetőnek tűnik.

Térbeli autokorreláció tehát van, mégpedig pozitív irányban. A panel modellekben súlyként a megállapított 36 km-es szomszédsággal meghatározott sorstandarizált inverz távolságmátrixot fogom használni, hogy a területi autokorrelációt korrigáljam az előállítandó modellben, de még hátra van a főkomponens elemzés és el kell döntenem, hogy milyen típusú panelmodellt használjak a pótolt adatokra.

A főkomponens-elemzés során két főkomponens sikerült azonosítanom. Az első főkomponens az információtartalom 70,15%-kát hordozza magában, a második főkomponens pedig a 14,53%-kát, azaz összesen együtt az információtartalom 84,68%-kát. Az adatok és módszertan részben kifejtett módon az első főkomponens tízezer főre vetített háziorvosok rendelői forgalmát, összes forgalmát, felnőtt és gyermek részére szervezett háziorvosi szolgálatok számát, háziorvosok szolgálatának számát és az ápolók számát tartalmazza. Ezt a főkomponenst lehet úgy tekinteni, mint az adott járás egészségügyi felszereltségét, hatékonyságát, észszerű feltételezés lehet, hogy minél több orvosi szolgálati óra jut tízezer főre egy járásban, annál több könnyebb bekerülni egy orvoshoz a rendelésén, illetve minél nagyobb a forgalom a rendelőben, annál nagyobb hatékonysággal, gyorsabban tudják ellátni a betegeket, feltételezve, hogy az ellátás minősége minden régióban homogén. Ez az ingatlanvásárlás és bérlés esetében egy igen fontos szempont lehet a fogyasztó részéről. A második főkomponensben, szintén tízezer főre vetítve a házi gyermekorvosok száma, ápolók száma és a szolgálatok száma szerepel. Ezt a főkomponenst véleményem szerint lehetne a gyermekek számára elérhető egészségügyi ellátottságnak nevezni. Úgy gondolom, hogy a lakásépítéseknél, főleg amikor CSOK-ról is beszélünk fontos szempont a szülők részéről, hogy olyan helyen vásároljanak vagy vágjanak építkezésbe, ahol a megszületett, illetve születendő gyermekek megfelelő egészségügyi ellátásban részesüljenek.

A tesztek során 5%-os szignifikancia szinten fogok döntést hozni. A Hausman-teszt nullhipotézise szerint a *random modell* a megfelelő, alternatív hipotézis pedig az, hogy a *fix hatású modell* a megfelelő. A tesztben az eredményváltozó a lakásépítés és a magyarázó változók a foglalkoztatottsági ráta, egy állandó lakosra jutó SZJA alapot képező jövedelem, a foglalkoztatottsági ráta, beruházási teljesítményérték egy lakosra vetítve és legvégül a “CSOK” szó keresési trendje Google-n. A CSOK trendet többféle módon fogom próbálni a modellbe illeszteni, itt három részre bontva, minőségi változóként kezeltem, aminek a szintjei a “nincs”, vagyis még nem vezették be a CSOK-ot, majd a bevezetés után a mediánnál elvágva alacsony és magas érdeklődésre bontottam. A Hausman-teszt p-értéke 2,2×1016, ez az jelenti, hogy a nullhipotézis valószínűsége gyakorlatilag 0%, ezért H0-t elutasítom a kiválasztott 5%-os szignifikancia szinten.

A *random modellnél* tehát jobb *a fix hatású*, azonban még a *fix hatású* és *a Pooled OLS* közötti döntéshez még egy F tesztet kell elvégezzek. Az F-teszt nullhipotézise szerint a *Pooled OLS* a megfelelő, alternatív hipotézise pedig az, hogy a *fix hatású modell* a megfelelő. A tesztből kapott p-érték 2,2✕1016, ami azt jelenti, hogy a nullhipotézis valószínűsége közel 0%. Ez alapján 5%-on elutasítom H0-t, *a fix modellt* preferálom mind a *Pooled OLS*-sel és a *random hatásúval* szemben.

A fix hatású területi panelmodell általános egyenlete a következő:

Ahol:

- Yit az i-edik egyed eredményváltozója a t-edik időpontban

- μi az i-edik egyedhez tartozó fix hatás

- βj a j-edik magyarázó változó koefficiense

-xijt az i-edik egyedhez tartozó j-edik magyarázó változó a t-edik időpontban

- λ a területi autokorrelációs koefficiens

-Ỹt az eredményváltozó súlyozott átlaga az i-edik egyed szomszédos területekein a t-edik időpontban

-εit az i-edik egyedhez tartozó becslési hiba a t-edik időpontban

A kezdeti modell esetében a Yit a tízezer főre jutó lakásépítések száma az adott járásban, a μi az adott járás fix egyedhatása, amit a modell becsül meg, ahogyan az együtthatókat is a modell becsli. Az együtthatókhoz tartozó magyarázó változók a kezdeti modellben minden magyarázó változót tartalmaznak, azonban a modellszelekció során sok változót elhagytam. A modellszelekciót minden CSOK-érdeklődés-interpretáció esetében hasonló elvek mentén, egyesével végeztem, mivel többféle *fix hatású* modellt illesztettem, azonban szerintem ezek közül négyet érdemes kiemelni.

Az első modellben a CSOK változót 4 részre bontottam: nincsen, alacsony érdeklődés, közepes érdeklődés, magas érdeklődés. Ezeket a csoportokat úgy alakítottam ki, hogy 2016-ig nincsen CSOK érdeklődés, majd onnantól az adatokat sorba rendezve harmadoltam.

A második modell az elsőhöz hasonló, azonban CSOK érdeklődésnél magas és alacsony van, a mediánnál elvágva.

A harmadik modellben a CSOK érdeklődés számként szerepel.

A negyedik modellben pedig egyáltalán nem szerepel.

Megfelelő modellszűkítés után a *3. táblázatban* szereplő modellekre jutottam.

*3. táblázat: A fix hatású modellek eredményei*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y: LAKAS | Modellek | | | |
|  | (1) | (2) | (3) | (4) |
| λ | 0,3368\*\*\* | 0,3424\*\*\* | 0.3471\*\*\* | 0.3523\*\*\* |
| SZJA | -0,0043 | -0.0039 | - | - |
| MUNKA | - | - | - | - |
| BERUHÁZÁS | -0,0043 | -0.0043 | -0.0057 | -0.0057 |
| ÁTLAGÁR | 1,0397 | 1.0154 | 0.9497 | 0.9688 |
| SÉRTETT | - | -0.0007 | -0.0006 | -0.0008 |
| PC1 | -1,5550 | -1.5884 | -1.8309 | -1.7868 |
| PC2 | -0,9547 | -0.9721 | -0.9176 | - |
| CSOK - alacsony | 2,3126 | 2.4211 | - | - |
| CSOK - közepes | 2,8423 | - | - | - |
| CSOK - magas | 0,5985 | 1.2783 | - | - |
| CSOK - számmal | - | - | 0.0062 | - |
| Megjegyzés | \*p < 0,05; \*\*p < 0,01; \*\*\*p < 0,001 | | | |

*Forrás: KSH, saját számítás*

A modellek együtthatói között minden esetben 1 százalékon is szignifikáns az összes modellben szereplő az ÁTLAGÁR változó, illetve a területi autoregresszív koefficiens. Ez egyrészt utal arra, hogy az adott járás lakásainak átlagára minden esetben nagyon fontos, pozitív irányban ható koefficiens. Egymillió forintnyi ingatlan átlagár növekedés az adott járásban ceteris paribus 1,0397-tel több tízezer főre jutó lakásépítést eredményez. Úgyszint nagyon fontos a lambda paraméter szignifikanciája is, ami a lakásépítés területi autokorrelációját ilyen formában igazolja, a szomszédos járások eredményváltozóinak súlyozott átlaga szignifikánsan növeli a tízezer főre jutó lakásépítéseket.

Az első modellben a CSOK trendnek négyféle értéke lehet: nincsen CSOK, alacsony érdeklődés, közepes érdeklődés, magas érdeklődés. Mivel kategorikus adatként van kezelve, ezért dummy kódolással adtam meg a modellnek, azaz van egy alapnak vett kategória és a többi kategória ahhoz viszonyítva értelmezendő. Az alap kategória pedig nem más, mint a „nincsen CSOK”, vagyis amennyiben alacsony p-értékűek a többi koefficiensek, az azt jelenti, hogy szignifikánsan eltérnek attól, hogy nincsen CSOK. Ideális esetben, amennyiben feltételezésem szerint a CSOK iránti érdeklődés kategóriái szignifikánsan eltérnek attól, mintha nem lenne CSOK, akkor ezek a koefficiensei a magasabb kategóriával nőnek.

Érdekes megfigyelni, hogy a legnagyobb növekedés a lakásépítésben az alacsony érdeklődés mellett van ahhoz képest, mintha nem lenne CSOK. A magas érdeklődés pedig semmilyen szokványos szignifikanciaszinten sem szignifikáns, nagy az esélye, hogy a sokaságban annak a bétának az értéke nulla. A modellszelekciót az Akaike és a Bayes-i információs kritériumok segítségével végzem el.

*4.táblázat: Az Akaike és Bayes-i információs kritériumok a modelleken*

|  |  |  |
| --- | --- | --- |
|  | AIC | BIC |
| 1. modell | 11927.62 | 11987.76 |
| 1. modell | 11933.74 | 11988.41 |
| 1. modell | 11937.51 | 11981.24 |
| 1. modell | 11937.27 | 11970.08 |

*Forrás: KSH, saját számítás*

Míg az AIC szerint a négyfelé bontott CSOK keresés a BIC szerint a CSOK nélküli modell a jobb. Én a BIC szerint döntök, így végső modellnek a CSOK nélküli modellt választom. A végső modell együtthatóinak bétája is érdekes.

A végső modellben ceteris paribus amennyiben növekszik egy egységgel az egy főre jutó beruházások teljesítményértéke, akkor a tízezer főre jutó lakásépítések száma 0,0057-tel csökken. C.p. az átlagár egy millió forintnyi növekedése a tízezer főre jutó lakásépítések száma 0,9688-cal nő. Amennyiben minden más változatlansága mellett nő a 100 lakosra jutó sértettek száma (bűncselekményekből) eggyel nő, 0,0008-cal csökken a tízezer főre jutó lakásépítések száma. Az első főkomponens c.p. egy egységnyi növekedése a tízezer főre jutó lakásépítések száma 1,7868-cal csökken a tízezer főre jutó lakásépítések száma.

Az első modellben a magas CSOK iránti érdeklődés valószínűleg egyáltalán nem szignifikáns, valamint a végső modellbe nem került bele a CSOK. Ez lehet akár azért is, mert a támogatott lakáshitelek nagy része használt ingatlan vásárlására lett felhasználva. (MNB, 2022). Egy másik tényező is hozzájárulhatott ehhez, a Balatonnál feltehetően nyaralók épültek nagy számban, amit nem az ott lakók vettek birtokba. A CSOK egyik feltétele, hogy az újdonsült ingatlanba át kell jelenteni a szülők és gyerekek lakcímét, ami a nyaraló esetben nem kivitelezhető, ezért valószínűsíthetően ezek nem CSOK-os hitelből és támogatásból épültek.

A szignifikáns, ám logikailag nem teljesen érthető, negatív bétájú beruházás és foglalkoztatottság is egy felettébb különös eredmény. Véleményem szerint ez akár azért is lehet, mert ezen változók egymással is nagymértékben korrelálnak és egyfajta mediáció léphet fel. Ennek a kezelését a SEM modellekkel fogom végezni. A CSOK keresési trend egyáltalán nem volt szignifikáns számként és minőségi változóként is csak az „alacsony érdeklődés” tért el 1%-on szignifikánsan attól, hogy nincsen CSOK. Így a dolgozatom első kutatási kérdésére, minthogy van-e CSOK iránti érdeklődésnek hatása a lakásépítésekben a válaszom az, hogy a panelmodellek eredményei alapján nincsen.

A SEM modellbe ezek alapján nem teszem be a CSOK keresési trend egyik változatát sem. Mivel paneladataim vannak, viszont a SEM alapvetően egy keresztmetszeti adatokra használatos modell, ezért három különálló évet vizsgálok meg: 2012, 2016, 2019. Azért esett a választásom ezekre az évekre, mivel meg szerettem volna vizsgálni a CSOK előtti, CSOK korai időszakait, valamit a pandémia előtti utolsó évet.

Egyes változóim túlságosan nagy szórást mutattak, így normalizáltan kerültek a modellekbe mind az eredmény, mind a magyarázó változók. A SEM modellek egy látens komponensből és 3 db egyenletből állnak. A látens komponenst a PC1 nevű főkomponensem elemei alkotják. A három egyenlet közül az elsőben a tízezer főre jutó lakásépítések számát magyarázom az SZJA, beruházás, foglalkoztatottság, átlagár, és az egészségügyi (látens) változókkal. A második egyenletben az átlagárat az SZJA-val és az egészségügyi látensváltozóval. A harmadik egyenletben pedig az SZJA-t magyarázom a foglalkoztatottsággal, beruházással és az egészségüggyel.

A modellek illeszkedésének jóságára az RMSEA-jukat (Root Mean Square Error of Approximation)-jukat néztem meg,

## 

## Diszkusszió

## Irodalomjegyzék

Anselin, L. (1988*). Spatial econometrics: Methods and models*.Dordrecht: Kluwer Academic Publishers.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93-115. Forrás: <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International regional science review*, 26(2), 153-166. Forrás: <https://doi.org/10.1177/0160017602250972>

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence*. Regional Science and Urban Economics*, 26(1), 77-104. Forrás: <https://doi.org/10.1016/0166-0462(95)02111-6>

Anselin, L., Gallo, J. L., & Jayet, H. (2008). Spatial Panel Econometrics. In Mátyás, L. & Sevestre,P. (szerk.), *The Econometrics of Panel Data: fundamentals and recent developments in theory and practice (Third edition)*(old.: 625-660). Berlin: Springer. Forrás: https://doi.org/10.1007/978-3-540-75892-1\_19

Békés, G., Horváth, Á., & Sápi, Z. (2016). Lakóingatlanárak és települési különbségek*. Közgazdasági Szemle,* 63(12), 1289–1323. Forrás: <https://doi.org/10.18414/ksz.2016.12.1289>

Dancsik, B., Marosi, A., & Szabó, B. (2022). Túl drága az olcsó hitel – a családi otthonteremtési kedvezmény támogatott hitelkamatainak vizsgálata. *Közgazdasági Szemle*, 69(12), 1493–1506. Forrás: https://doi.org/10.18414/ksz.2022.12.1493

Fellner, Z., Marosi, A., & Szabó, B. (2021). A babaváró kölcsön hitelpiaci és reálgazdasági hatásai. *Közgazdasági Szemle,* 68(2), 150–177. Forrás: <https://doi.org/10.18414/ksz.2021.2.150>

Magyar Nemzeti Bank. (2022). *Lakáspiaci jelentés (2022 május)*. Forrás: <https://www.mnb.hu/letoltes/laka-spiaci-jelente-s-2022-ma-jus-hun.pdf>

Moran, P. A. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2), 17-23. Forrás: https://doi.org/10.2307/2332142

Nagy, K., Palócz, É., & Vakhal, P. (2016). Risks and challenges to international and Hungarian econo-my in the middle of the decade. *Statisztikai Szemle,* 93(Special number 20), 20–43*.* Forrás: https://doi.org/10.20311/stat2016.k20.en020

Plöchl, K., & Obádovics, C. (2021). A CSOK-támogatást igénylők vizsgálata az előzetes gyermekvállalás és az ingatlanszerzés mértéke szempontjából egy hitelintézet adatai alapján. Forrás: *Hitelintézeti Szemle*. Forrás: https://doi.org/10.25201/hsz.20.3.80109

Tóth, G. (2003). Területi autokorrelációs vizsgálat a Local Moran I módszerével. *Tér és Társadalom*, 17(4), 39-49.Forrás: http://real.mtak.hu/115938/1/EPA02251\_Ter\_es\_tarsadalom1825.pdf

Varga, A. (2002). Térökonometria. Statisztikai Szemle, 80(4), 354-370. Forrás: https://www.ksh.hu/statszemle\_archive/2002/2002\_04/2002\_04\_354.pdf

Varga, A. (2002). Térökonometria. *Statisztikai Szemle*, 80(4), 354-370. Forrás: <https://www.ksh.hu/statszemle_archive/2002/2002_04/2002_04_354.pdf>

Vasárus, G. (2022). Városhatáron belüli szuburbanizáció Magyarországon – egy paradoxon feltárása. *Területi Statisztika,* 62(4), 379–404. Forrás: https://doi.org/10.15196/ts620401

Wood, S. N. (2003). Thin Plate Regression Splines. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 65*(1), 95–114. http://www.jstor.org/stable/3088828

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Internetes források:

Google Trends (n.d.). „CSOK”. https://trends.google.com/trends/explore?date=all&geo=HU&q=CSOK&hl=hu, Letöltés dátuma: 2023.04.08.

Központi Statisztika Hivatal. (2023). *Térképes Interaktív Megjelenítő Alkalmazás*. <https://map.ksh.hu/timea/?locale=hu>, Letöltés dátuma: 2023.04.08.

Hivatkozott jogszabályok:

16/2016. (II. 10.) Korm. rendelet az új lakások építéséhez, vásárlásához kapcsolódó lakáscélú támogatásról

17/2016. (II. 10.) Korm. rendelet a használt lakás vásárlásához, bővítéséhez igényelhető

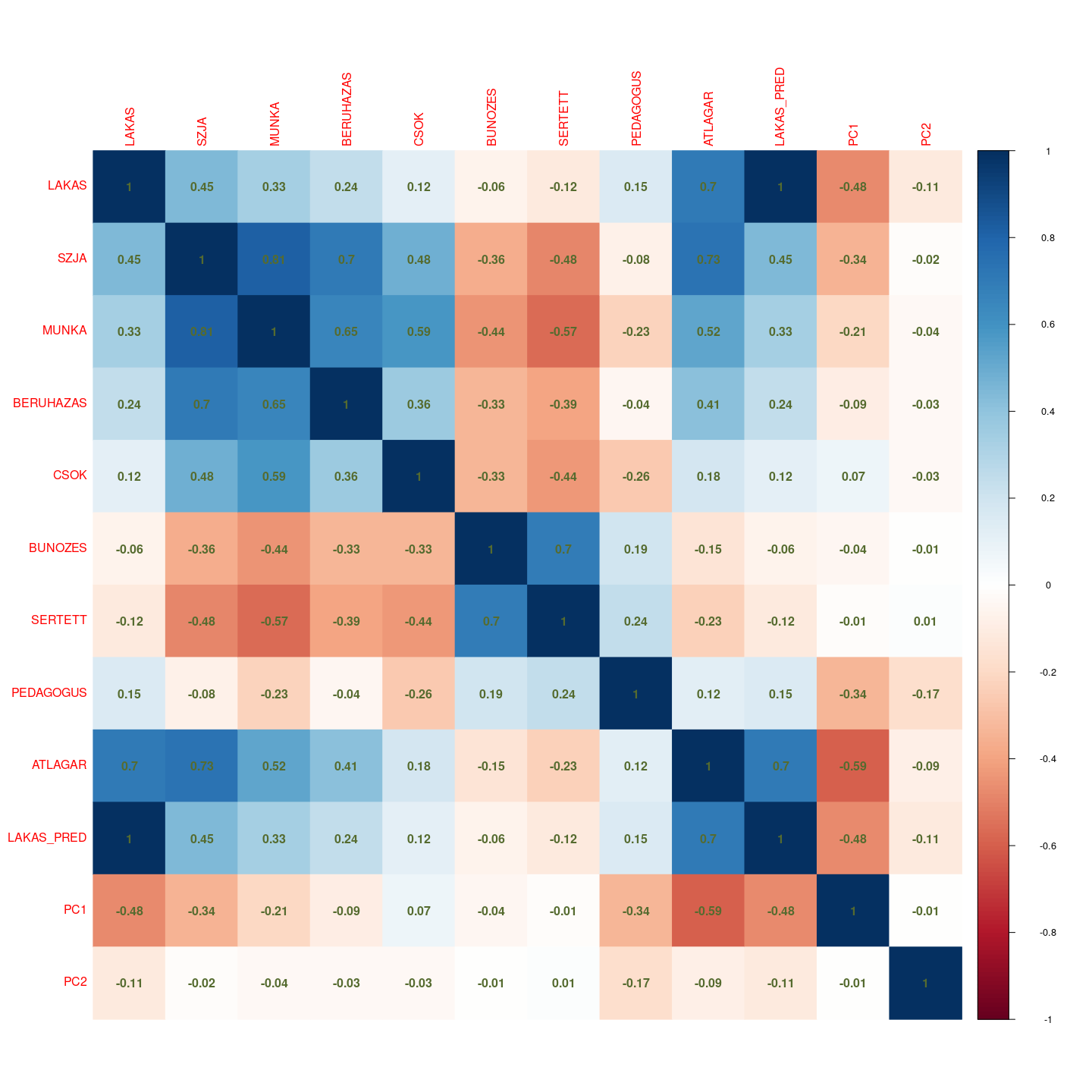
családi otthonteremtési kedvezményről

44/2019. (III. 12.) Korm. rendelet a babaváró támogatásról

2012. évi XCIII. törvény a járások kialakításáról, valamint egyes ezzel összefüggő törvények

módosításáról

*3. ábra: korrelációs mátrix*

*Forrás: KSH, saját számítás*

További adatvizualizációkhoz a készítettem egy R Shiny webappot a dolgozatban használt adatforrásokból, ami az alábbi linken érhető el: https://dleves.shinyapps.io/SzakdolgozatMapPlot/