

OTDK-dolgozat

2023

BUDAPESTI CORVINUS EGYETEM



**HATÁSSAL VOLTAK-E A KORONAVÍRUS-JÁRVÁNY NYOMÁN
BEVEZETETT SZIGORÍTÁSI INTÉZKEDÉSEK
A MUNKANÉLKÜLISÉGRE EURÓPÁBAN?**

**DID THE RESTRICTIVE POLICY INTERVENTION
IN RESPONSE TO THE CORONAVIRUS PANDEMIC
HAVE AN EFFECT ON UNEMPLOYMENT IN EUROPE?**

Kavas Katalin

Konzulens: Dr. Vincze János

Kézirat lezárásának dátuma: 2021. 04. 12.

Absztrakt

Dolgozatom a koronavírus-járvány nyomán bevezetett szigorítási intézkedések munkaerőpiaci hatására fókuszál. A kutatás során az Amerikai Egyesült Államok államaira vonatkozó eredményekből kiindulva 31 európai ország esetén megvizsgáltam a világjárvány kezelésére irányuló intézkedések 2020-as munkanélküliségi rátára kifejtett hatását. A vizsgálatot az Oxfordi Egyetem korlátozási indexének segítségével, regressziós módszereket alkalmazva végeztem el.

A szakirodalomban gyakran alkalmazott éves adatokra épülő egyszerű regressziós módszer intuícióellenes eredményre vezetett, ezért vizsgálatomat kibővítettem a megbetegedések előrejelzésében gyakran alkalmazott panelregressziós becsléssel. A havi gyakoriságú adatokon végzett panelregresszió eredménye már megerősítette a korábbi eredményeket, miszerint a korlátozások szigorításával növekszik a munkanélküliség is. A paneladatok azonban szignifikáns országhatásokat tükröztek, amelyek az európai országok eltérő járványkezeléséből és a fiskális intézkedések eltérő szintjéből eredhettek.

Mivel a fiskális támogatások és a korlátozási index között valószínűsíthető a pozitív korreláció, így ezt követően megvizsgáltam a korlátozási indexet alkotó mutatókat a hagyományos statisztikai elemzésektől eltérő, gépi tanulási szemléletű módszer segítségével. A gépi tanulás módszerének alkalmazása lehetővé tette a szigorítási intézkedéseket komponenseinek mélyrehatóbb vizsgálatát és szelekcióját is, amely arra az eredményre vezetett, hogy a meglehetősen országspecifikus és a korábban nemigen vizsgált „nyilvános tájékoztató kampányok szintje” mutató segítségével jelezhető elő a legjobban a munkanélküliségi ráta alakulása a korlátozási index mutatói közül.

Kulcsszavak: koronavírus, munkanélküliség, korlátozási index, fiskális intézkedések, regressziós becslés, gépi tanulás

JEL klasszifikáció: C51, J11, J21

Tartalomjegyzék

Táblázatjegyzék.....	iii
Ábrajegyzék.....	iv
I. Bevezetés.....	1
II. Szakirodalmi áttekintés	1
II.1. A koronavírus-járvány gazdasági következményei	1
II.2. A koronavírus-járvány hatása a munkaerőpiacra.....	3
II.3. A téma aktualitása.....	5
III. Adatok és adattisztítás	8
III.1. Az éves modellek elkészítéséhez felhasznált adatok bemutatása	8
III.2. Adattisztítás az éves modellekben.....	11
III.3. A havi gyakoriságú adatok bemutatása	12
III.4. Adattisztítás a havi gyakoriságú adatokon	12
IV. Az éves lineáris regressziós modell.....	13
IV.1. A kiinduló éves modell bemutatása	13
IV.2. Az első modell eredményei.....	14
IV.3. Turizmus-vendéglátásban foglalkoztatottak arányának kihagyása melletti modell	15
V. Havi adatok elemzése panelregresszióval	20
V.1. Panelregressziós módszer bemutatása	20
V.2. A második modell eredményei	25
VI. Gépi tanulási módszerek.....	28
VI.1. Gépi tanulási mechanizmusra épülő módszerek bemutatása	28
VI.2. Véletlenerdő-eljárás	29
VI.3. A véletlenerdő-eljárás eredménye	31
Hivatkozásjegyzék.....	35
Mellékletek.....	I
1. számú melléklet: A vizsgált európai országok listája.....	II
2. számú melléklet: A korlátozási indexet alkotó mutatók bemutatása.....	III
3. számú melléklet: A lineáris modell futtatásához szükséges programkód	VII
4. számú melléklet: A panelmodell futtatásához szükséges programkód.....	X
5. számú melléklet: A véletlenerdő-modell futtatásához szükséges programkód	XIV

Táblázatjegyzék

1. táblázat: A korlátozási mutató meghatározásánál figyelembe vett mutatószámok, saját ábra az Oxford korlátozási index adatai alapján	8
2. táblázat: (1) Regressziós modell illesztésének eredményei	14
3. táblázat: A (2) egyenlettel készült lineáris becslés eredményeit összefoglaló táblázat	15
4. táblázat: A koronavírus-válságot kísérő gazdasági intézkedések eltérően alakultak az egyes európai országok esetében.....	17
5. táblázat:A within modell eredményei a korlátozási indexhez tartozó együttthatóra.....	25
6. táblázat: A pooling modell eredményei.....	26
7. táblázat: A between modell eredményei	26
8. táblázat: A korlátozási indexet alkotó mutatók véletlenerdő eljárással nyert fontossági sorrendje, csökkenő sorrendbe rendezve.....	31

Ábrajegyzék

1. ábra: A munkanélküliség szóra világszerte az interneten rákeresők alakulása az elmúlt 5 évben, a legmagasabb értékhez viszonyítva, százalékban.	6
2. ábra: Az elmúlt 5 évben a munkanélküliség szóra interneten rákeresők száma az Amerikai Egyesült Államokban, legmagasabb értékhez viszonyítva, százalékban	6
3. ábra: a 4. táblázatban bemutatott európai országok fiskális támogatásra fordított GDP arányos kiadásai a Bruegel Intézet mutatói szerinti bontásban és az éves átlagos korlátozási index ...	18
4. ábra: A 31 vizsgált országban mért havi átlagos korlátozási indexek átlagos értékei	23
5. ábra: A 2020-as havi átlagos korlátozási indexeket szemléltető doboz ábrák	24

I. Bevezetés

Napjainkban a minket körülvevő vírushelyzet nemcsak az emberek egyéni, mindennapi életvitelében okozott változásokat, de azokat a komoly kihívásokat is elhozta, amelyekkel az államoknak szembe kell nézniük. 2021 tavaszán, - amikor ezt a dolgozatot készítettem - a járvány hosszútávú következményeiről még nem kaphattunk teljeskörű képet, azonban a rendelkezésre álló információkból már levonhattunk következtetéseket. Az egészségügyi-, gazdasági- és társadalmi következmények mérlegelésében hasznos segítség lehet egy költség-haszon elemzés készítése, azonban azt a környezetünkben is láthatjuk és tapasztalhatjuk, hogy a munkaerőpiac a 2020-as év elejéhez képest mára már jelentősen átrendeződött. Munkahelyek semmisültek meg és teremtek a vírus, valamint a vírust kordában tartani igyekvő intézkedések hatására. A kérdés aktualitása, és relevanciája miatt a dolgozatom elkészítésének célja a koronavírus megjelenése nyomán bevezetett szigorítási intézkedések munkanélküliségi rátára kifejtett hatásának vizsgálata.

II. Szakirodalmi áttekintés

II.1. A koronavírus-járvány gazdasági következményei

A koronavírus nyomán kibontakozó egészségügyi válságot kísérő gazdasági válság nem újkeletű következmény, és ennek munkaerő piacra kifejtett hatásainak vizsgálata indokolt (Moreno et al. 2021). Példaként említve Maruthappu és szerzőtársai (2015) 74 országra kiterjedő elemzésükben megmutatták, hogy a 2009-es gazdasági válság nyomán megnövekedett munkanélküliség a HIV fertőzöttek halálozási arányának emelkedéséhez vezetett.

Petersen és szerzőtársai (2020) szerint a koronavírus olyan megbetegedési hullámot okozott, amelyhez hasonló gyorsasággal terjedő járvánnyal a fejlett világnak az 1918-as spanyolnátha járvány óta nem kellett szembenéznie.

A vírus megjelenése óriási kihívásokat okozott Európa szerte is, annak ellenére, hogy az egészségügyi felkészültséget mérő Global Health Security Index¹ 2019-es adatai szerint az Európai országok a prevenciós skálán a két legjobban felkészült kategóriában szerepeltek (Váradi, 2020). A járvány terjedése azonban az egészségügyi intézkedések mellett életre hívott számos gazdasági beavatkozást is. Anderson és szerzőtársai (2020) a koronavírus következményeként felmerült gazdasági visszaesésekre adott fiskális válaszlépések elemzésére törekedtek tanulmányukban. Munkájukban az Európai Unió országaiban, az Egyesült Királyságban, illetve az Amerikai Egyesült Államokban tett intézkedéseket vizsgálva 3 csoportot különböztettek meg, amelyek jól tükrözik a vírus nyomán alkalmazott jellemző gazdaságpolitikai intézkedéseket:

- Az azonnali fiskális ösztönzések csoportjába tartoztak az olyan felmerült extra költséget jelentő kormányzati kiadások, amelyeket az egészségügyi kapacitások bővítésére vagy a munkahelyek megőrzésére fordítottak, és azonnali terhet jelentettek a költségvetés számára. Utóbbi tulajdonság miatt az adócsökkentés, illetve – elengedés is ebben a kategóriába került.
- A halasztások olyan később visszafizetendő, magánszemélyek vagy vállalatok fizetőképességének javítását célzó intézkedéseket tömörítettek, amely egyaránt tartalmaztak kamat- és tőkevisszafizetési engedményeket is, és amelyek csak átmenetileg terhelik a költségvetést, később bevételnövekedéshez vezetnek.
- Végül az egyéb likviditási tartalékok és garanciák csoportja hasonlóan épül fel, mint az előbb említett halasztás mutató, de azzal ellentétben nem automatikusan vehető igénybe, és nem jelöl meg konkrét célcsoportokat. Ide tartoznak többek között a különböző exportgaranciák, vagy fejlesztési hitelkeretek.

Ezt követően a mutatószámok értéke olyan módon került meghatározásra, hogy az egyes célokra fordított pénzösszegeket kategóriánként aggregálták, majd az adott ország 2019-es GDP-jével leosztották. Az így kapott mértékek már összemérhetők voltak.

Az elsőként említett fiskális mutató szerepe kiemelkedően fontos a koronavírus-járvány negatív gazdasági következményeinek enyhítésében Apeti és szerzőtársai (2021) eredményei szerint. Az általuk végzett kutatás előrejelző értéke kiemelkedő: 125 ország esetében vizsgálták meg - az IMF Koronavírussal kapcsolatos fiskális válaszlépések mutatójának segítségével - a fiskális kiadások alakulását, amelynek becslésére az adóbevételek államadóssághoz viszonyított aránya

¹ <https://www.ghsindex.org/> , letöltés dátuma: 2021. 02. 23.

bizonyult a legalkalmasabbnak (Apeti et al., 2021). Ezzel szemben foglal állást Hosny (2021), az Oxfordi korlátozási index adataival végzett elemzése szerint ugyanis éppen a kisebb mértékű fiskális reakciók vezetnek a vírus leghatékonyabb megfékezéséhez, és következményeinek kezeléséhez.

Ezen csoportosítási szempontokkal találkozhatunk az OECD Foglalkoztatottsági Előrejelzésében (2020) is, ahol arra az eredményre jutottak, hogy a COVID-19 hatására elvesztett munkahelyek számának növekedése, és a heti átlagos ledolgozott munkaórák csökkenése közel tízszeres volt a 2008-as válság során tapasztaltak.

Azonban a fenti mutatók Köllő (2020) vizsgálata szerint hiányos adatokra épülnek, ugyanis nem tartalmazzák a 2020-ban bevezetett összes intézkedést, illetve a Bruegel Intézet oldalán² közzétett információk szerint mindössze 11 országra terjednek ki, így inkább csak az első hullám során végzett vizsgálatokra voltak alkalmasak. Mindezen okok alapján dolgozatomban a szélesebb körű Hale et al. (2020) által kidolgozott Oxford stringency index adataival dolgoztam. Az oxfordi index előnye, hogy globális, napi szintű adatokat tartalmaz 2020. január 1.-től kezdődően, és a gazdasági mutatókon felül a kijárási-, bezárási- illetve utazási korlátozások szintjét is számszerűsíti, valamint a járvány egészségügyi kezelésére tett intézkedésekre is kiterjed. Ezt összefoglalóan a továbbiakban korlátozási indexnek nevezem, és a későbbiekben részletesen bemutatom.

II.2. A koronavírus-járvány hatása a munkaerőpiacra

Zádori et al. (2020) szerint a globálissá váló világban az eltérő fejlettségű régiók hosszú ideig humán-erőforrás problémákkal küzdöttek, amelyek hatására már a vírus megjelenését megelőző évtizedekben is érzékelhetők voltak a deglobalizációs folyamatok, vagyis a bezárkózó nemzetgazdasági irányzatok erősödése. Az előzőek hatására megjelenő negatív foglalkoztatási tendenciák mellett a koronavírus-járvány súlyosan érintette a munkaerőpiacot, és a munkanélküliségi ráták érzékelhető emelkedéshez vezetett (Guaitoli és Tochev, 2021).

Fairlie és szerzőtársai (2020) szerint a kijárási korlátozások bevezetése két hónap alatt közel 15 százalékos emelkedéshez vezetett a munkanélküliségi ráta területén, és különösen súlyosan érintette a kisebbségi csoportokat. Hershbein és Holzer (2021) OLS regresszióval nyert

² <https://www.bruegel.org/publications/datasets/covid-national-dataset/>, letöltve: 2021. 02. 23.

eredménye is megerősítette, hogy a kisebbségi csoportok, és alacsony fizetéssel járó szakmákat végzők, illetve a legkevésbé képzett rétegek munkaerőpiaci helyzetét érintette a legsúlyosabban a járvány. Rámutattak arra is, hogy a munkanélküliség változása nem egyenletes az éven belül sem: márciusban, a szigorítások szintjének drasztikus emelkedése alatt, illetve az egyes hullámok tetőzésének idején növekedett meg a leginkább, míg a hullámok közötti enyhítések alatt részlegesen talpra állt a piac (Hershbein és Holzer, 2021). Ország szintű vizsgálataik szerint ezek a negatív hatások még hosszú ideig érezhetők lesznek az egyes nemzetgazdaságokban, és az érintettek célzott támogatása (lásd fent, a halasztások pontban) egy lehetséges megoldás lehet ennek kezelésében (Hershbein és Holzer, 2021). Nelson (2021) nem állami tulajdonú vállalatokon végzett elemzéseiben rámutatott arra is, hogy a tartós munkavégzést igénylő szakmák kevésbé váltak érintetté, mint az időszakos foglalkoztatást folytatók. Utóbbira különösen nagy hatással volt, hogy a vizsgált országban milyen szigorú intézkedések érvényesültek az iskolák bezárására vonatkozóan (Nelson, 2021).

Guaitoli és Tochev (2021) USA államainak helyzetét vizsgáló munkájában kiemelte, hogy a koronavírus munkanélküliségre gyakorolt hatásában erős tovaryűrűző hatás figyelhető meg. Egy-egy vizsgált területen ugyanis akkor is nagyon magas növekedést figyeltek meg akkor is, ha az állam adott megyéjében ugyan alacsonyabb volt a korlátozási index értéke, de a szomszédos megyékben szigorú lezárásokat alkalmaztak (Guaitoli és Tochev, 2021). Eredményeik szerint a munkanélküliségi ráta változásának nagyjából 15-21%-át magyarázza meg a szomszéd-hatás, amelynek hálózatielméleti alapon végzett vizsgálata az Európai országok között a jelen tanulmány egy lehetséges folytatását képezhetné.

Baek et al. (2020) a lezárással töltött munkanapok számát a vizsgálati időszak teljes hosszához viszonyította, és rámutatott, hogy az Amerikai Egyesült Államokban 1 további hét teljes lezárás és kijárási korlátozás akár 1,9%-kal is emelheti az állam munkanélküliségi rátáját.

Zádori et al. (2020) koronavírus-járvány után a munkaerőpiaci verseny élesedését prognosztizálja, és állami beavatkozást látnak szükségesnek az egyenlőtlenségek mérséklésének céljából.

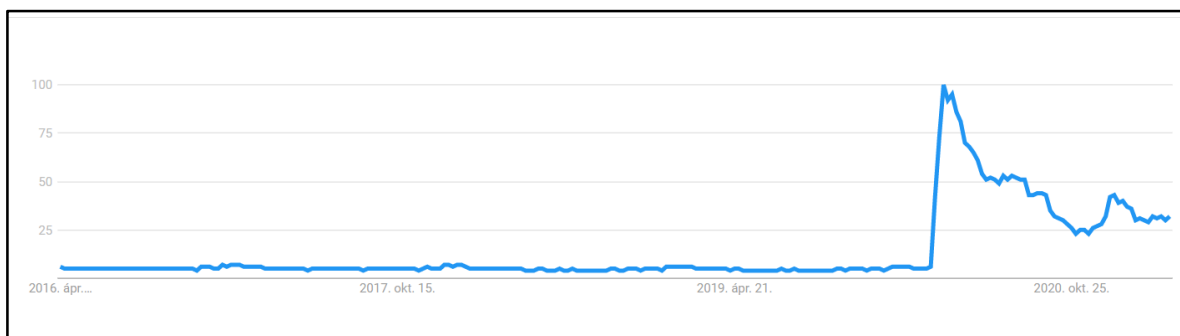
Del Roca et al. (2021) szintén a korlátozási indexet alapul véve vizsgálta az első és második hullám munkaerőpiacra és háztartási munkamegosztásra gyakorolt hatását Olaszországban.

Korábbi eredményeik szerint az első hullámban a háztartási munkák egyenlőtlen megoszlása a nőket súlytotta jobban az első hullám idején, különösen ha távmunkában dolgoztak a járvány alatt (Del Boca et al., 2020). Ez a munkaerőpiacról való kieséshez és inaktivitáshoz vezetett

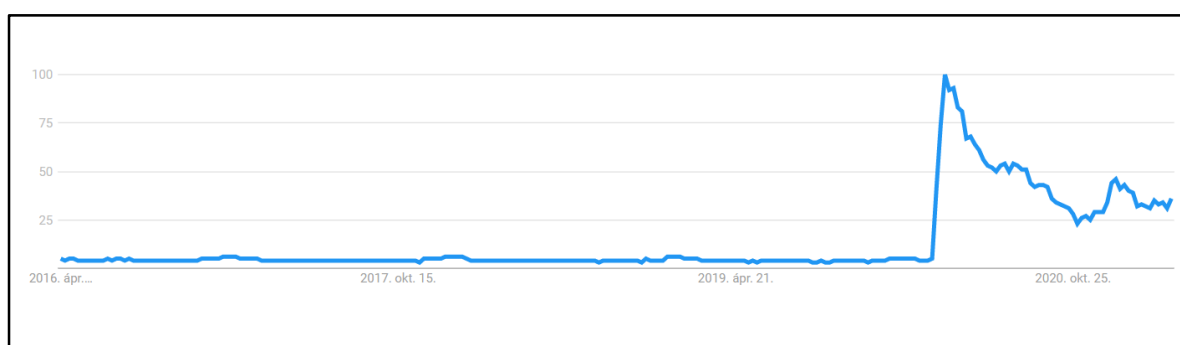
számos esetben (Del Boca et al., 2020). A 2021-es tanulmány megerősítette az eredményeket, sőt előre vetítette, hogy a tendencia a nemek közötti fizetésbeli egyenlőtlenségek növekedéséhez vezethet (Del Roca et al., 2021). Fabrizio et al. (2021) eredményei megerősítik, hogy a koronavírus-járványt kísérő lezárások első 9 hónapja alatt az alacsonyabb képzettségű kisgyermekes anyukák voltak a legsúlyosabban érintett munkavállalói csoport az Amerikai Egyesült Államokban: eredményeik szerint a 12 év alatti gyermekkel rendelkezők nők munkaszerzési esélye a 3%-kal csökkent. További tény, hogy az iskola bezárási korlátozások nemek közötti fizetésbeli egyenlőtlenségekre ható egyik legfontosabb tényezővé vált (Fabrizio et al., 2021). A nemek közti eltérő munkaerőpiaci következményekre hívja fel a figyelmet Nelson (2021) is.

II.3. A téma aktualitása

Kong és Prinz (2020) ökonometria módszerekkel vizsgálta a munkaügyi tárgyú Google-keresések mintázat, amelyet függő változóként felhasználva, igyekeztek a közpolitikai intézkedések munkaerőre gyakorolt hatását azonosítani. A Google Trends adatbázisában elérhető annak a vizuális ábrázolása, hogy egy-egy kifejezésre hányan keresnek rá az interneten az általunk kiválasztott időintervallumban. A grafikon segítségével így képet kaphatunk arról, hogy mekkora az érdeklődés az általunk választott téma iránt. Az alábbi ábrákon látható, hogy a munkanélküliség kifejezés iránt jelentősen megugrott az érdeklődés a koronavírus elterjedésének idejében. Az 1. ábra azt mutatja be, hogy az elmúlt 5 év során hogyan alakult a napi keresések száma: a legnagyobb forgalomhoz tartozik a 100%-os érték, a többi napon pedig ehhez viszonyítottan láthatjuk a százalékos értékeket. Az adatbázis adatai szerint a kifejezés az Amerikai Egyesült államokban volt a legnépszerűbb, így az országhoz tartozó trendet külön kiemeltam a 2. ábrán. Mindkét esetben elmondhatjuk, hogy az első 4 évben a legmagasabb tapasztalható értékek is alig érték el a 2020-as csúcs 10%-át. Azt az állítást, hogy a fokozott érdeklődés a járványhoz köthető, megerősíti az az eredmény, hogy az elmúlt 5 évben a „munkanélküliség” szóhoz leggyakrabban kapcsolt kifejezés a „pandémia” volt, mind globálisan, mind az Amerikai egyesült Államokban. Ezek az adatok és eredmények megtekinthetők az ábrákhoz tartozó leírásokban található hivatkozásokon keresztül.



1. ábra: A munkanélküliség szóra világszerte az interneten rákeresők alakulása az elmúlt 5 évben, a legmagasabb értékhez viszonyítva, százalékban. Forrás: <https://trends.google.com/trends/explore?date=today%205-y&q=unemployment>



2. ábra: Az elmúlt 5 évben a munkanélküliség szóra interneten rákeresők száma az Amerikai Egyesült Államokban, legmagasabb értékhez viszonyítva, százalékban. A keresések során a leggyakrabban társított kifejezés a pandémia volt a vizsgált időintervallumban. Forrás: <https://trends.google.com/trends/explore?date=today%205-y&geo=US&q=unemployment>

Összefoglalva az eddigieket azt mondhatjuk, hogy napjainkban releváns kérdés a 2020-as munkanélküliségi ráta alakulásának vizsgálata. A rendelkezésre álló adatok korlátai miatt még nem vonhatunk le mélyreható, vagy hosszútávú következtetéseket a világjárvány kezelésére irányuló stratégiák eredményességét illetően, azonban lehetőségünk van a már tapasztalt hatások mérésére. A munkaerőpiac a gazdaság egy, a járvány által erősen érintett szegmense, és a munkanélküliség alakulásának fontos következményei lehetnek mind gazdasági, mind társadalmi értelemben az elkövetkező években.

Dolgozatomban így a munkanélküliségi ráták alakulásában tapasztalható változásokra koncentrálok, egészen pontosan arra a kérdésre, hogy hatással vannak-e rá a jelenleg alkalmazott korlátozási intézkedések. Várakozásaim szerint a korlátozási index szignifikáns hatással bír a munkanélküliségi ráta alakulására. A szakirodalomban eddig elterjedt éves

lineáris regressziós modell mellett, a következőkben havi gyakoriságú adatokra épülő panelregressziós modell segítségével vizsgálom meg, hogy lehetséges-e az éves modell illeszkedésének finomítása.

Végezetül gépi tanulási módszerek alkalmazásával keresem arra a választ, hogy az Oxfordi 9 korlátozási indexet alkotó mutatók közötti szelekcióval javítható-e a modell magyarázó ereje, illetve kiegészíthető-e az index olyan egészségügyi- vagy gazdasági intézkedéseket számszerűsítő mutatóval, amelyek segítségével pontosabb képet kaphatunk a korlátozások munkanélküliségi rátára kifejtett hatásáról.

III. Adatok és adattisztítás

III.1. Az éves modellek elkészítéséhez felhasznált adatok bemutatása

A becslés során figyelembe vett korlátozási index a következő 9 komponenst tartalmazza:

1. táblázat: A korlátozási mutató meghatározásánál figyelembe vett mutatószámok, saját ábra az Oxford korlátozási index adatai alapján

Jelölés	Megnevezés	Lehetséges értékei
C1	Iskola bezárása	0; 1; 2; 3
C2	Munkahelyek bezárása	0; 1; 2; 3
C3	Nyilvános események lemondása	0; 1; 2
C4	Gyülekezési korlátozás	0; 1; 2; 3; 4
C5	Tömegközlekedés leállítása	0; 1; 2
C6	Lakóhely elhagyásának korlátozása,	0; 1; 2; 3
C7	Belföldi utazások korlátozása;	0; 1; 2
C8	Nemzetközi utazások korlátozása;	0; 1; 2; 3; 4
H1	Nyilvános tájékoztató kampányok szintje	0; 1; 2

A korlátozási index napi értékének meghatározása a fenti 9 mutató szerint képzett indikátor számtani átlagából adódik, az Oxfordi COVID-19 Kormányzati reakció követő (2020) által alkalmazott jelöléseket alkalmazva:

$$I = \frac{1}{9} \cdot \sum_{j=1}^9 I_j$$

Ahol:

I_j a fenti mutatók szerint képzett indikátorok, $j=1; 2; \dots; 9$

I pedig a napi korlátozási index értéke.

Mivel a C1 – C8 mutatók lehetséges értékei eltérők, ezért minden indikátor esetében egy 0 és 100 közötti értéket határozzunk meg a következő módszerrel (Oxfordi COVID-19 Kormányzati reakció követő, 2020 jelöléseivel):

$$I_j = 100 \cdot \left(C_j \cdot \frac{1 - w}{N_j} + w \cdot G_j \right)$$

ahol:

C_j a j-edik mutató által felvett tényleges érték;

w az egyes C_j mutatók súlya;

N_j a j-edik mutató által felvehető maximális érték;

G_j a j-edik mutató nagyságának meghatározásánál figyelembe vett intézkedések kiterjedését tükröző érték. Esetükben két lehetséges értéket vehet fel: célzott intézkedések esetén $G_j=0$, általános érvényű intézkedések mellett $G_j=1$.

Az C_j mutatókhoz tartozó w súly meghatározása a lehetséges maximális értékek szerint történik:

$$w = \frac{1}{8} \cdot \sum_{j=1}^8 \frac{1}{N_j + 1}$$

A korlátozási index meghatározásában $w = 0,29$.

Az utolsó H1 mutatóhoz tartozó indexérték:

$$I_9 = 100 \cdot \frac{H1}{N_9}$$

módon adódik.

A fenti mutatók részletes bemutatását, mérési módszerét és lehetséges értékeit az 1. számú melléklet tartalmazza. Dolgozatomban a korlátozási index vizsgált időszakra jellemző szintjét a vizsgált időperiódus alatti értékek számtani átlagaként határoztam meg, Doti (2021) módszertanát követve.

A becslésem során a turizmusban foglalkoztatottaknak az összes foglalkoztatotthoz viszonyított arányát az Eurostat statisztikájának megfelelően, az oldalon elérhető adatok segítségével

határoztam meg. A hiányosan rendelkezésre álló foglalkoztatottsági mutatók miatt, a turizmusban foglalkoztatottak összességét az Európai Unió gazdasági tevékenységeket besoroló NACE (European Classification of Economic Activities) kódok alapján határoztam meg. Két csoportot vettem figyelembe a becslés során:

- az I jelű „Szállás és Étkezési szolgáltatások” (Accommodation and food service activities) szektorokban foglalkoztatottak száma;
- valamint az N79 jelű „Utazási irodák, utazásszervező és más szállásfoglalási szolgáltatások és kapcsolódó tevékenységek” (Travel agency, tour operator and other reservation service and related activities) szektorokban foglalkoztatottak száma.

Így a becsült foglalkoztatottsági arány a következőképpen alakult:

$$j_{lh19_i} = \frac{emp_{I,i} + emp_{N79,i}}{emp_{total,i}} \cdot 100$$

ahol:

j_{lh19_i} az i-edik országban a turizmusban foglalkoztatottak aránya az összes foglalkoztatotthoz viszonyítva, százalékban;

$emp_{I,i}$ az i-edik országban az I jelű ágazatokban foglalkoztatottak száma 2019-ben, ezer főben mérve;

$emp_{N79,i}$ az i-edik országban az N79 jelű ágazatban foglalkoztatottak száma 2019-ben, ezer főben mérve;

$emp_{total,i}$ az i-edik országban az összes aktív korú foglalkoztatott száma 2019-ben, ezer főben mérve.

A munkanélküliségi mutatók meghatározása során a kiinduló adatokat az Eurostat 2019. január 1. és 2021. január 1. közötti időszakra vonatkozó, aktív korú népességhez viszonyított munkanélküliségi ráta havi bontású, szezonhatással kiigazított adatai jelentették. A regresszióhoz meghatároztam a 2019-ben és 2020-ban bekövetkezett relatív változásokat a következő év január elseji és az adott év január elseji munkanélküliségi ráta hányadosaként, mind a 31 vizsgált ország esetében (ezen országok listáját a countries változó tartalmazza). Az adatok használatát az összehasonlíthatóság indokolta a munkanélküliség konkrét értékeivel

szemben, ugyanis az egyes országok a kezdő időpillanatban eltérő nagyságú munkanélküliségi rátával bírtak. Következő lépésként a csak nemnegatív értékű változókat logaritmizáltam. A vizsgálat évét megelőző változó bevezetésével Doti (2021) a korlátozási index hatását szeretne volna izolálni, ugyanis feltételezése szerint a vizsgált államokban a gazdasági növekedési trendek miatt a 2020-ban bekövetkezett változások szintje megközelítőleg megegyezett volna a 2019-ben bekövetkezettekkel, ha nem áll fenn a vírushelyzet.

III.2. Adattisztítás az éves modellekben

A bemutatott változók előzőek szerint nyert értékeit adattáblába rendeztem, amelyet a becsléshez használtam fel.

Összefoglalva az eddigieket, az adattábla az alábbiakat tartalmazza:

- az ország neve;
- 2019. 01. 01 és 2020. 01. 01 között, a munkanélküliségi rátában bekövetkezett relatív változás természetes alapú logaritmusa;
- 2020. 01. 01 és 2021. 01. 01 között, a munkanélküliségi rátában bekövetkezett relatív változás természetes alapú logaritmusa;
- a korlátozási index éves átlagos értéke 2020. 01. 01. és 2020. 12. 31. között;
- a turizmus-vendéglátás szektorban foglalkoztatottak aránya az összes foglalkoztatotthoz viszonyítva 2019-ben.

Ezt követően az adatokat megtisztítottam az olyan megfigyelésektől, amelyek hibásak voltak. 3 fajta hiba fordult elő az adatban:

- a) hiányos volt a munkanélküliségi rátára vonatkozó megfigyelés a 2020-as évben;
- b) hiányos adatsorból származott az átlagos korlátozási mutató értéke;
- c) hiányzott az adott országhoz tartozó turizmusban foglalkoztatott aránya adat.

Ezek alapján összesen 12 megfigyeléstől kellett a vizsgált adatokat megtisztítani.

III.3. A havi gyakoriságú adatok bemutatása

A havi munkanélküliségi ráta változók meghatározásához használt kiinduló adattábla ismét az Eurostat 2019 január 19. és 2020. január 19. közötti időszakra vonatkozó aktív korú népességhez viszonyított munkanélküliségi ráta havi bontású, szezonhatással kiigazított adatait tartalmazó adatbázis volt. A regresszióhoz ezt követően meghatároztam a 2019-ben és 2020-ban bekövetkezett relatív változásokat, minden két egymást követő hónap esetében, az előző hónap utolsó napjához és az adott hónap utolsó napjához tartozó munkanélküliségi ráta hányadosaként, a 31 vizsgált országra vonatkozóan.

Következő lépésként ismét Doti (2021) módszertanát követve megbecsültem minden vizsgált ország esetében, minden hónapra a havi átlagos korlátozási indexet, a hónap napjaihoz tartozó korlátozási indexek átlagaként.

A modellben a turizmus-vendéglátásban dolgozók összes foglalkoztatotthoz viszonyított aránya nem szerepel, mivel a változóról havi gyakoriságú adat nem érhető el, illetve a lineáris modellben nem bizonyult szignifikánsnak a változó.

Mind az éves, mind a havi adatok esetében longitudinális formájú adatokkal dolgoztam.

III.4. Adattisztítás a havi gyakoriságú adatokon

A sortban ismét csak az adattisztítás lépése következett, ebben az esetben 2 típusú hiba merült fel:

- a 2020-as munkanélküliségi rátára vonatkozó megfigyelések hiányossága, ezek logaritmusai ugyanis mínusz végtelenként szerepelt volna az illesztésben;
- illetve a havi korlátozási index hiányossága.

A fentiek nyomán összesen 37 megfigyelést kellett figyelmen kívül hagynom, így a panelregressziós becsléshez 335 megfigyelés állt rendelkezésre.

IV. Az éves lineáris regressziós modell

IV.1. A kiinduló éves modell bemutatása

Doti (2021) nyomán a következő egyszerű regressziós modell segítségével vizsgálom meg a koronavírus-járvány kezelésével járó intézkedéseknek a munkanélküliségi rátára gyakorolt hatását, melyet az oxfordi egyetem stringency indexének (a továbbiakban: korlátozási index) értékeivel közelítem. A legkisebb négyzetes eltérések elve szerint vizsgálta a koronavírus következtében bevezetett lezárások munkanélküliségre kifejtett hatását Guaitoli és Tochev is (2021). Doti (2021) ezen felül még egy magyarázóváltozóval, a turizmus-vendéglátásban foglalkoztatottak összes foglalkoztatotthoz viszonyított arányával egészíti ki az általa vizsgált modellt. Így az általam vizsgált modellben a célváltozónak a munkanélküliségi ráta 2020-ban bekövetkezett változását választottam, magyarázó változókként pedig a korlátozási index éves átlagos értéke, a munkanélküliségi ráta 2019-ben bekövetkezett éves változása, illetve a turizmusban foglalkoztatottak aránya szolgált.

A lineáris modell előnye, hogy egyszerűen értelmezhető eredményre vezet, amely a későbbi becslések viszonyítási alapját képezheti.

A fentiek értelmében a kiinduló modellem:

$$u_{i,t} = \beta_0 + \beta_s \cdot s_i + \beta_u \cdot u_{i,t-1} + \beta_{tv} \cdot jlh19_i \quad (1)$$

ahol:

$u_{i,t}$ a munkanélküliségi ráta 2020-ban bekövetkezett éves változásának természetes alapú logaritmusa az i -edik országban;

s_i a korlátozási index átlagos értéke 2020-ban az i -edik országban;

$u_{i,t}$ a munkanélküliségi ráta 2019-ben bekövetkezett éves változásának természetes alapú logaritmusa az i -edik országban;

jlh_i a turizmus-vendéglátásban foglalkoztatottak aránya az összes foglalkoztatotthoz viszonyítva 2019-ben, az i -edik országban;

β_0 , pedig a becsülni kívánt paraméterek.

A tisztított adatokkal ezt követően lineáris regressziót illesztettem a 2020-as munkanélküliségi adatokra az (1) egyenlet segítségével, a legkisebb négyzetes eltérések módszerét alkalmazva (továbbiakban: OLS).

Az elvégzett becslés során sejtésem szerint a 2020-as munkanélküliségi ráta alakulására hatással van:

- korlátozási index éves átlagos értéke;
- a munkanélküliségi ráta 2019-es év alatt bekövetkezett változása;
- a turizmus-vendéglátásban foglalkoztatottak összes foglalkoztatotthoz viszonyított arány;

és mindhárom változó értékének növekedése növeli a munkanélküliségi ráta 2020-ban bekövetkezett növekedésének nagyságát.

IV.2. Az első modell eredményei

Az adattisztítást követően elvégeztem regressziós becslést az alábbi (1) egyenlettel. Az eredmények a következő módon alakultak:

$$u_{t,i} = \beta_0 + \beta_s \cdot s_i + \beta_u \cdot u_{t-1,i} + \beta_{tv} \cdot jlh19_i \quad (1)$$

2. táblázat: (1) Regressziós modell illesztésének eredményei

Jelölés	Együttható értéke	t érték	Pr(> t)
β_0	0,8484	5,118	0,0001

β_s	-0,0136	-3,827	0,0017
β_u	0,4229	3,213	0,0058
β_{tv}	0,0019	0,212	0,8347

Akár $\alpha=1\%$, akár $\alpha= 5\%$ vagy éppen $\alpha=10\%$ -os szignifikancia szintet választunk, a becslés eredményei alapján azt mondhatjuk, hogy szignifikáns hatása van a korlátozási index értékének és a 2019-es éves munkanélküliségi ráta alakulásának a 2020-as munkanélküliségi ráta változására. Azonban a magas p érték miatt nem utasítjuk el, hogy $\beta_{tv} = 0$, tehát a becslés szerint a turizmusban foglalkoztatottak összes foglalkoztatotthoz való arányának nincs hatása a 2020-as munkanélküliségi ráta alakulására.

IV.3. Turizmus-vendéglátásban foglalkoztatottak arányának kihagyása melletti modell

A következőben megvizsgáltam, hogy javítható-e a modell illeszkedése a turizmusra vonatkozó változó kihagyásával. Az új modell:

$$u_{i,t} = \beta_0 + \beta_s \cdot s_i + \beta_u \cdot u_{i,t-1} \quad (2)$$

Eredmények:

3. táblázat: A (2) egyenlettel készült lineáris becslés eredményeit összefoglaló táblázat

Jelölés	Együttható értéke	t érték	Pr(> t)
β_0	0,8489	5,282	0,00007
β_s	-0,0134	-4,048	0,0009
β_u	0,4241	3,326	0,0043

A modellszelekció során látható, hogy a korrigált R négyzet 0,6328-re nőtt 0,6095-ről, tehát a turizmus változó elhagyásával a modell magyarázó ereje javult.

Mivel a t és F tesztek megkövetelik a hibatagok normalitását, ezért Jarque-Bera normalitási tesztet végeztem a reziduumokon, amelyek nem-normalitása jelzi a hibatagok nem-normalitását. A JB teszt eredményeként nem utasíthatjuk el, hogy a reziduumok eloszlása

normális. RESET teszt segítségével ellenőriztem, hogy a függvényforma valóban lineáris-e, és a p érték alapján 5%-os szignifikancia szinten nem utasíthatjuk el ezt a hipotézisünket. Végül a heteroszkedaszticitást teszteltem, Breusch-Pagan teszt segítségével arra az eredményre jutottam, miszerint homoszkedaszticitás feltételező nullhipotézisünket nem utasíthatjuk el, tehát az OLS becslés konzisztens marad, ugyanakkor a késleltetett tag miatt nem mondhatjuk, hogy az OLS torzítatlan lenne.

Az előzőek fényében, az együtthatók értékét behelyettesítve a munkanélküliségi ráta változásának becslésére szolgáló egyenlet a következő:

$$u_{t,i} = 0,8489 - 0,0134 \cdot s_i + 0,4241 \cdot u_{t-1,i}$$

Vagyis azt mondhatjuk, hogy éves adatokat elemezve, becsléseink szerint ha:

- Egy adott országban 2020 során magasabb volt a korlátozási index átlagos értéke, akkor a 2020-as év során a munkanélküliségi ráta kevésbé emelkedett. Megfigyelhető, hogy a korlátozási index emelkedése mellett a munkanélküliségi ráta csökkenő irányba mozdulna el, az előző évben tapasztalthoz képest.
- Egy adott országban, ahol magasabb volt 2019-ben a munkanélküliségi ráta változása, akkor a 2020-as évben is magasabb változást tapasztalhatunk.

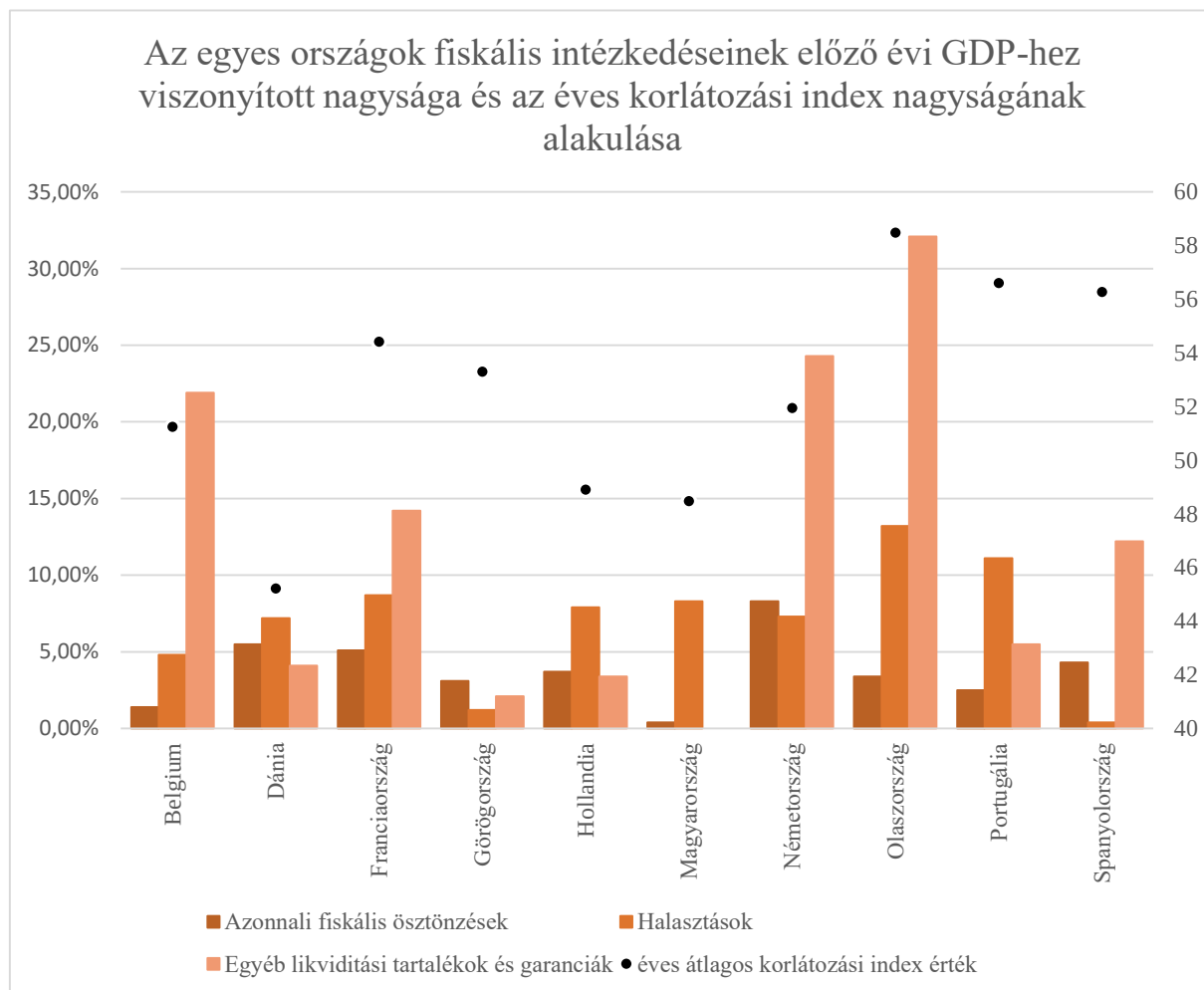
A szakirodalomban tapasztalható eredményekkel összevetve elmondható, hogy Doti (2021) eredményei szerint a turizmusban alkalmazottak aránya szignifikáns magyarázó változó a munkapiaci költségek vizsgálatában, ugyanis az elvesztett állások számára nagy hatással van. Ezzel szemben áll becslésemben a turizmus változó nem szignifikáns volta. Az előző éves adatokhoz tartozó paraméterek előjele megegyezik a Doti (2021) által kapott eredményekkel. Nelson (2021) munkája is megerősíti a korlátozási index foglalkoztatottságra történő hatását a tartósan foglalkoztatottak, illetve a női munkavállalók körében, míg nem bizonyult szignifikánsnak a változó az alkalmi munkát végzők esetében. A korlátozási index előjele éppen a várakozásinkkal ellentétesen alakult, és ellentmond Doti (2021) eredményeinek is, miszerint enyhébb szigorítási intézkedések mellett több munkahely őrizhető meg.

Az intuícióval ellentétes eredmény abból eredhet, hogy az USA-beli államok járványkezelési eljárásai homogénebbek, mint az európai országok által alkalmazottak. Gondoljunk csak a 2020. első járványhulláma alatt a nyájimmunitás elérésére törekvő svéd módszerre, és a hazánkban alkalmazott teljes lezárásra. A gazdasági intézkedések terén is jelentős eltéréseket tapasztalhatunk az egyes országok között. Az alábbi táblázat a Bruegel Intézet adataival szemlélteti, hogy a szakirodalmi áttekintésben bemutatott mérőszámok értéke szemmel láthatóan eltérően alakult a feltüntetett országok esetében. A táblázat adatait az Amerikai Egyesült Államok aggregált adataival is kiegészítettem, azonban az adat hiányossága miatt a többi vizsgált ország esetén nem áll rendelkezésre.

4. táblázat: A koronavírus-válságot kísérő gazdasági intézkedések eltérően alakultak az egyes európai országok esetében, A feltüntetett adatok a megfelelő ország 2019-es GDP százalékában vannak kifejezve.
Saját szerkesztés, az adatok forrása: Anderson et al., 2020

Ország	Azonnali fiskális ösztönzések	Halasztások	Egyéb likviditási tartalékok és garanciák
Amerikai Egyesült Államok	9,1%	2,6%	2,6%
Belgium	1,4%	4,8%	21,9%
Dánia	5,5%	7,2%	4,1%
Franciaország	5,1%	8,7%	14,2%
Görögország	3,1%	1,2%	2,1%
Hollandia	3,7%	7,9%	3,4%
Magyarország	0,4%	8,3%	0,0%
Németország	8,3%	7,3%	24,3%
Olaszország	3,4%	13,2%	32,1%
Portugália	2,5%	11,1%	5,5%
Spanyolország	4,3%	0,4%	12,2%

A fenti 4. táblázat adatait szemléletesen a következő hisztogramon ábrázoltam. Az elsődleges tengely mentén ábrázolt fiskális támogatások GDP-százalékban mért értékét olvashatjuk le, míg a másodlagos tengelyen az országos éves átlagos korlátozási index érték szerepel. A fiskális intézkedésekről itt is láthatjuk, hogy meglehetősen eltérőek és az országok más-más intézkedésre fordították a legtöbbet. Azt azonban elmondhatjuk, hogy a korlátozási indexek alakulása hasonló tendenciát követ mint a fiskális



3. ábra: a 4. táblázatban bemutatott európai országok fiskális támogatásra fordított GDP arányos kiadásai a Bruegel Intézet mutatói szerinti bontásban és az éves átlagos korlátozási index, saját ábra

A szakirodalomban elterjedt nézetek szerint azokban az országokban ahol a fiskális ösztönzések szintje magasabb volt, ott kevésbé emelkedett a munkanélküliség. Várakozásaink szerint éppen ezekben az országokban magasabb volt a lezárások mértéke is, vagyis a lezárások pozitívan korrelálnak a nagyobb mértékű fiskális támogatásokkal. A következőkben panelregresszió

segítségével megvizsgáltam, hogy valóban beszélhetünk fixhatások fellépésről a munkanélküliségi ráta alakulásának becslése során.

V. Havi adatok elemzése panelregresszióval

V.1. Panelregressziós módszer bemutatása

Az éves adatok elemzése után, megvizsgáltam, hogy havi gyakoriságú adatokra térve finomítható-e a becslés, és kimutathatók-e országhatások a becslés során.

A havi adatok elemzését lineáris panelregresszió segítségével végzem el. A havi megfigyelések ugyanis panel adatoknak tekinthetők Balázsi et al. (2014) definíciója alapján: az egyedek (itt: országok) egyes tulajdonságainak (munkanélküliségi ráták, korlátozási index) időbeli alakulásáról vannak megfigyeléseink. A megfigyelt adatunk így több dimenzióssá válik, Hildreth (1950) és Balázsi et al (2014) elnevezéseivel élve, a vizsgálni kívánt modell egyszerre tartalmazza a csoport- és idődimenziót. A panelregresszió további előnye, hogy képes kezelni a heterogenitást, ami abból ered, hogy az eltérő egyedek eleve oksági kapcsolatban állhatnak a megfigyelni kívánt változóinkkal, vagyis a paneladatok vizsgálatával jobban megragadható az oksági kapcsolat, mint keresztmetszeti vagy idősoros adatbázis elemzésével (Balázsi et al., 2014). Ez abból adódik, hogy a modellünk lehetővé teszi az aggregált trendekre, és az országokra egyedileg jellemző, időben állandó hatásokra (egyéni hatás) történő kontrolálást is.

A paneladatok vizsgálatának alapjait Airy (1861) fektette le, aki csillagászati mérésekkel belátta, hogy a napi mérései a korrekciók elvégzése után is eltérnek egymástól, és a mérési hibákban szabályszerűséget, összefüggéseket fedezett fel. Ez a tulajdonság a paneladatok egyik legfontosabb jellemzőjét jelenti mind a mai napig. A szabályszerűségek mélyre hatóbb vizsgálata során később elkülönítették a fix- és véletlen hatás fogalmát és problémaosztályát (Fisher, 1925; Nerlove, 2005, Eisenhart, 1947). Eisenhart (1947), Scheffé (1956) és Madaras (2009) a definíciói alapján:

- fixhatás lép fel, amikor az összes megfigyelés egy csoportjának tagjaira felmerül egy állandó, azonos és jellemző tulajdonság, amely a csoportok között eltér, de időben állandó;
- míg a véletlen hatás akkor lép fel, amikor a mutatók egymás között nagyon eltérők, de időben állandók, vagy időben nagy varianciát mutatnak, viszont egymás között állandók.

Hildreth (1950) és munkájában összehasonlította a két modellt, és eredményei, illetve Croissant és Millo (2008) nyomán dolgozatomban ezt követően a fixhatás modellekre koncentrálok, és ezzel a módszerrel végzem a becslést a havi adatokra.

A fixhatás modellünkben a β_{hs} , illetve a β_{hu} együtthatók becslésére két módszert is választhatunk (Hanck et al., 2020):

- a korábbiakhoz hasonlóan OLS becslést végzünk a dummy változókkal kiegészített regressziós modellünkben;
- OLS becslés végzése az úgynevezett „entity demeaned” adatokon, melynek lényege hogy a regressziós oldal mindkét oldalát átlagolja, majd az átlagot kivonja a megfigyelések értékéből.

Dolgozatomban az utóbbi segítségével elvégeztem a regresszió illesztését. A panelregresszió alkalmazásának további előnye, hogy ha valamely egyedre hiányos az adat (például egy ország valamely hónapjához nem tartozik érték adathiba vagy adathiány miatt), a módszer akkor is alkalmazható. Ha az adatunk minden változó megfigyelését tartalmazza minden egyedre és időszakra, akkor kiegyensúlyozott (*balanced*), ellenkező esetben kiegyensúlyozatlan (*unbalanced*) panelről beszélünk (Croissant és Millo, 2008; Hanck et al., 2020). Jelen esetben az utóbbi áll fenn. A panelregresszió alkalmazása a koronavírus hatásainak elemzésében leginkább a fertőzöttek számának elemzésében elterjedt. Chernozhukov et al., (2021) megbetegedések (korlátozási index összetevőjeként megjelenő) iskolák bezárási szintjével való kapcsolatának vizsgálja a modell segítségével, amellyel figyelembe tudja venni a betegség lappangási idejét is.

Az első pont értelmében Hanck et al (2020) jelöléseit követve a vizsgált egyenlet, ahol a fixhatásokat dummy változóként kezeljük:

$$u_{h_{i,t}} = \beta_{h0} + \beta_{hs} \cdot sh_{i,t} + \beta_{hu} \cdot u_{h19_{i,t}} + CFE_i + MFE_t + \varepsilon_{i,t} \quad (3)$$

ahol:

$u_{h_{i,t}}$ a munkanélküliségi ráta havi relatív változásai 2020-ban az i-edik országban, a t-edik időszakban;

$sh_{i,t}$	a korlátozási index átlagos havi értéke 2020-ban az i-edik országban, a t-edik időszakban;
$u_{h19_{i,t}}$	a munkanélküliségi ráta havi relatív változásai 2019-ban a megfelelő t-edik hónapban, az i-edik országban;
CFE_i	Az i-edik ország egyedi jellemzőiből adódó fix hatás;
MFE_t	A t-edik hónap jellemzőiből adódó fix hatás;
$\beta_{h0}, \quad \beta_{hs}, \quad \beta_{hu}$	a becsülni kívánt paraméterek.

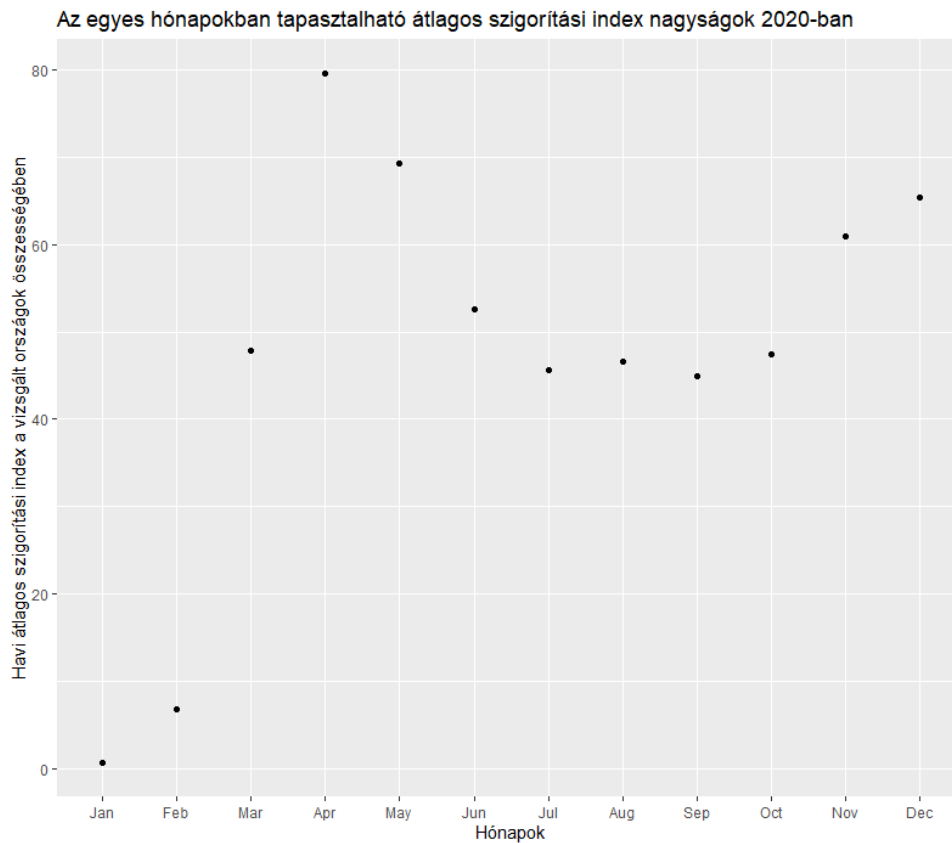
Az elvégzett becslés során hipotézisem szerint a 2020-as munkanélküliségi ráta havi alakulására hatással van:

- korlátozási index adott havi átlagos értéke;
- a munkanélküliségi ráta 2019-ben az adott hónapban bekövetkezett változása.

Ez utóbbi változó értékének növekedése mellett az $u_{h_{t,i}}$ változó növekedését is tapasztaljuk.

A panelregresszió illesztése során a 31 országra és 12 hónapra dummy változókat vezetünk be, összesen 42 darabot. Az alábbi táblázat tartalmazza a β_{hs} , β_{hu} együtthatók értékét és a hozzájuk tartozó t-statisztikát közlöm.

Megvizsgálva azonban azt, hogy a 31 országban átlagosan mekkora volt a korlátozási index mértéke, az a következtetés vonható le, hogy január, illetve február hónapokban messze elmaradt az index értéke az év folyamán később tapasztalhatóktól. A havi átlagos korlátozási indexnek a vizsgált 31 országban mért átlagos értéket az alábbi ábra mutatja be szemléletesen:

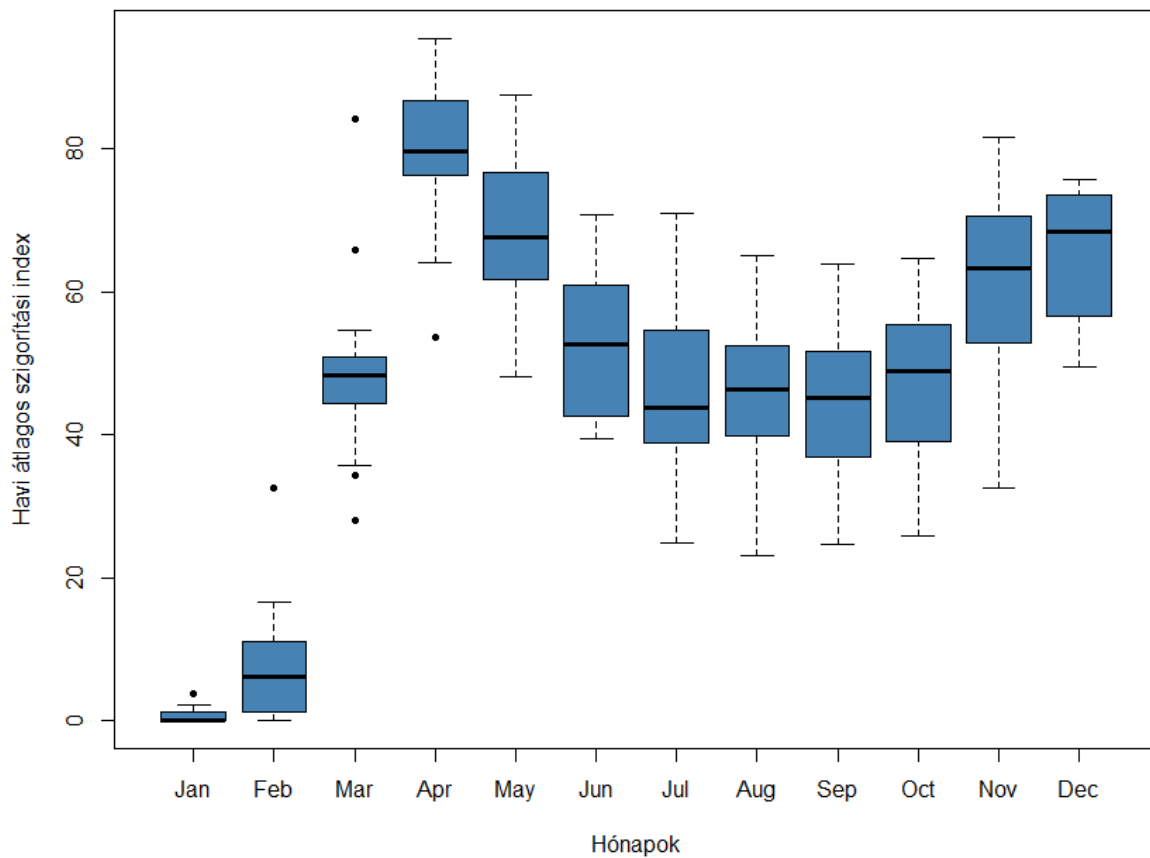


4. ábra: A 31 vizsgált országban mért havi átlagos korlátozási indexek átlagos értékei, saját ábra

Az ábrán hasonló tendenciát figyelhetünk meg, mint a munkanélküliség kifejezésre rákeresők számának esetében 2020-ban.

A havi értékek országok között tapasztalt eltérő értékeit szemlélteti az alábbi (4.) doboz ábra, ahol látható a havi átlagos korlátozási indexek szóródása a 31 vizsgált ország esetén:

2020 egyes hónapjaiban tapasztalható átlagos szigorítási index nagyságok dobozábrái



5. ábra: A 2020-as havi átlagos korlátozási indexeket szemléltető doboz ábrák, saját ábra Hosny (2021) nyomán

A fenti ábrák alapján látható, hogy az egyes hónapokban a korlátozási index átlagos nagysága eltérő lehet, illetve jól elkülöníthető az első és második hullám ideje is, így indokolt lehet a havi rendszerességű adatok vizsgálata is.

5.2. Paneladatok elemzése a késleltetett tag elhagyása mellett

Elsőként a késleltetett tag elhagyásával végzek lineáris panelregressziót, ugyanis az itt alkalmazott fixhatás modell nem lenne konzisztens az $u_{h19_{i,t}}$ változóval való becslés mellett.

A plm programcsomag előnye, hogy figyelembe veszi a heteroszkedaszticitást, és a megfelelően korrigált kovarianciamátrix segítségével végzi el a becslést, így helyes sztenderd hibákat ad eredményül (Croissant és Millo, 2008). A program segítségével lineáris regressziós becslést hajtok végre a paneladatokon, 3 különböző modellt alkalmazva:

- A *within modell* n-1 dummy változót tartalmaz, amelyek először az ország- és hónapváltozók lehetséges értékeire kerültek bevezetésre, majd csak az országváltozókra.
- A *pooling modell* a klasszikus lineáris regressziós becslésnek feleltethető meg (Croissant és Millo, 2008), amelyet úgy kezelünk, mintha nem lépnének fel fixhatások, tehát minden dummy együtthatója megegyezik.
- A *between-effect modell* azzal a feltételezéssel él, hogy a változók egymás közti különbsége állandó, és csak időben változó (Madaras, 2009), így az országokra nézve átlagolja a magyarázó- és célváltozókat. A kapott eredmény a nem panel adatokat felhasználó éves egyszerű modell outputjával hasonlítható össze.

V.2. A második modell eredményei

5. táblázat: A *within modell* eredményei a korlátozási indexhez tartozó együtthatóra

β_s együttható	Együttható értéke	Sztenderd hiba	t érték	Pr(> t)
ország és hónap dummy mellett	0,00054	0,00037	1,4322	0,11531
csak ország dummy (individuális változó) mellett	0,00047	0,00016	2,8866	0,0042

6. táblázat: A pooling modell eredményei

Együttható	Együttható értéke	Sztenderd hiba	t érték	Pr(> t)
β_0	1,00165	0,00806	124,30	0,0000
β_s	0,00036	0,00015	2,3796	0,0018

7. táblázat: A between modell eredményei

Együttható	Együttható értéke	Sztenderd hiba	t érték	Pr(> t)
β_0	1,0699	0,01306	81,9493	0,0000
β_s	-0,00109	0,00027	-3,9550	0,0004

A within és pooling modell egymással szembeni teszteléséből arra a következtetésre jutunk, hogy mivel a teszthez tartozó p érték nagyobb, mint 0,05, így nem fogadjuk el, hogy a regresszióban minden együttható (a konstans tagot is beleértve) megegyező lenne. Tehát elmondhatjuk, hogy van fixhatás az adatunkban.

A Wooldridge-teszt eredményei alapján nem utasíthatjuk el, hogy az adatunkban autokorreláció lép fel (Wooldridge, 2002), ugyanis a nem megfigyelt egyedi hatás tesztet végezve, a kapott p érték kicsi (0,0037), így elutasítjuk azt, hogy a nincsenek egyedi hatások. Ez az eredmény a within modell használatának kedvez, azonban még nem jelenti azt, hogy nem merülhet fel korreláció az időszakok között. A Pesaran (2004) által kidolgozott keresztmetszeti függőségi teszt eredményei szintén megerősítik a within modell használatának előnyét a pooling modellel szemben.

Mind ezek alapján azt mondhatjuk el, hogy a paneladatokon végzett becslésünkben erős országhatásokat tapasztaltunk, és a magyarázóváltozók előjele már az intuíciónak megfelelően alakult. A between-effect modell a becslések ellenőrzéseként pedig valóban visszaadta az intuícióellenes lineáris regresszióval nyert hatást.

Az első *within* becslés eredményei nyomán érdekes kérdésnek tartottam, hogy a korlátozási indexet alkotó mutatók egyformán fontosak-e a munkanélküliség előrejelzésében, vagy szerepelnek az indexben olyan komponensek amelyek hangsúlyosabbak. Az eddigi eredményeim alapján a várakozásom az, hogy a korlátozási index valamely meglehetősen országspecifikus mutatója kerül túlsúlyba a többihez képest. A korlátozási index komponenseinek vizsgálatához gépi tanulási módszerre épülő modellt alkalmaztam.

VI. Gépi tanulási módszerek

VI.1. Gépi tanulási mechanizmusra épülő módszerek bemutatása

A következőkben bemutatott modellek a gépi tanulási mechanizmusra épülnek. Az ilyen módszeren alapuló modellek egy jellegzetes tulajdonsága, hogy nem élnek előzetes feltevással a vizsgált változók eloszlását, valamint a függő és független változók kapcsolatát illetően (Nyitrai, 2021). A gépi tanulás alatt a dolgozatban azon módszerek és mechanizmusok összességét értem, amelyek az inputként felhasznált adatokon egy, következtetések és predikciók levonására alkalmas, modellt tanulnak.

A gépi tanulásra épülő modellek alkalmazása során nem hagyható figyelmen kívül, hogy az illeszkedés hatékonyan javítható a hiperparaméterek változókhoz való hangolásával, ugyanakkor fennáll a túltanulás veszélye is: akár hiba nélküli eredmény is elérhető az előrejelzésében, ám nem előnyös ha a valós adatokhoz hasonlóan hibákat és zajt is tartalmaznak előrejelzett adataink (Hayashi és Oishi, 2018; Nyitrai, 2021).

A modell alapját képező döntési fák előnye, hogy egyszerre használhatók regressziós és klasszifikációs problémák megoldására (James et al., 2013). Amikor regressziós fát illesztünk a vizsgálni kívánt adatainkra, a fa gyökerétől indulva elágazások mentén haladunk a kimenetek felé. Az elágazási pontok definiálása során diszjunkt csoportokba soroljuk be a változóinkat, az általuk felvehető lehetséges értékek szerint. Az elágazások kialakításának egy lehetséges és általános módja a rekurzív bináris felosztás, amely az elágazási pontokban az RSS minimalizálása alapján alakítja ki a felosztást a következők szerint:

$$RSS = \sum_{j=1}^J \sum_{i \in (R_j, S_j)} (y_i - \hat{y}_{R_j})^2$$

ahol:

$\hat{y}_{R_j}$ a J-edik csoportba eső y értékek átlaga;

s_j az előre becsült elágazási érték az adott R_j csoportban;

illetve $J = 4$ esetünkben, a később részletezettek szerint.

Ezzel ugyanis az összes döntési pontra vizsgáljuk a legkedvezőbb elágazási értéket, ami az egész fára minimalizálja az RSS értéket.

A módszer alkalmazásához azonban, - hogy lehetséges legyen az elágazási pontok kialakítása - a folytonos változóinkat is ordinálissá szükséges transzformálni.

A klasszifikációs és regressziós fák (CART) modellje is egy rekurzív algoritmusra épülő gépi tanulási módszer, amely a rendelkezésre álló adatokon határoz meg könnyen vizualizálható döntési szabályokat (Choubin et al., 2018). A CART modell, amely Breiman és szerzőtársaitól (1984) származik, a Gini-index minimalizálására törekszik, a vágási pontokat az alapján határozza meg, hogy a magyarázóváltozók varianciáját minimalizálja. A Gini-index meghatározása az alábbiak szerint történik (Choubin et al., 2018) :

$$G = 1 - \sum_{j=0}^k p_j^2$$

ahol:

k a kialakított kategóriák száma;

p_j az egyes kategóriák relatív gyakorisága.

A magyarázó változókra számított Gini-indexek aggregálása a bontás relatív gyakoriságával súlyozva történik. A jól megválasztott CART modellben a kialakított fa Gini-index értéke a lehető legkisebb a vágásokat nem tartalmazó esethez viszonyítva. A rendelkezésre álló kis mintában azonban nem vezetett látványos eredményre az egyetlen döntési fára építő CART modell, így a mintanagysághoz viszonyítottan sok változó kezelésére alkalmas véletlenerdő-eljárással pontosítottam a becslésemet, amely már látványos eredményre vezetett.

VI.2. Véletlenerdő-eljárás

A véletlenerdő-eljárás (RF), lényegében egy olyan klasszifikációs algoritmus, amely Breiman (2001) nevéhez köthető és sajátossága, hogy egyszerre van klasszifikációs és regressziós célja (James et al., 2013). A klasszifikációs fák a lehető legnagyobb homogenitás elérésére törekednek a kialakított csoportokon belül, előnyük a dolgozat szempontjából kiemelő, ugyanis előnyösebb a kvalitatív jellegű problémák megoldása során, szemben azokkal az esetekkel, amikor kvantitatív eredményekre vagyunk kíváncsiak (James et al., 2013).

Mivel egy egyedi döntési fa eredményére nagy befolyással bír a kiindulási adatok akár kismértékű megváltoztatása is, ezért ennek hatását az együttes tanulás módszerével igyekeztek csökkenteni (Nyitrai, 2021). Az eljárás lényege, hogy az egyes döntési fák elkészítése során csupán a rendelkezésre álló adatok egy véletlenszerűen választott részhalmazát vizsgáljuk, és az egyedi fák által adott eredmények összegzésével nyerjük a modell tényleges eredményét (Xu, 2013). Kumar és Thenmozhi (2006), illetve Palmer et al. (2007) munkájukban összehasonlított különböző gépi tanulási folyamatokat, és arra jutottak, hogy az RF módszer predikciós ereje közel olyan jó volt, sőt általában meghaladta más gépi tanulási modellekét, így dolgozatomban a következőkben én is ezt a modellt alkalmazom. Ezzel a módszerrel azonban csak akkor pontosíthatjuk a becslésünket, ha az egyes véletlen mintákon nyert eredmények eltérők (Nyitrai, 2021), azaz teljesül az eredmények diverzitása (Xia et al., 2020). A diverzitás tovább növelhető, ha független változókból való mintavételt is alkalmazunk a megfigyelésekből származó minták mellett (Tattar, 2018).

A véletlenerdő eljárásban az m elemű mintákból, visszatevés mellett m elemből álló véletlen mintákat nyerünk, ezzel összesen n darab döntési fát képzünk (Nyitrai, 2021). Ezek alkotják a modellfejlesztés adatállományát. Látható továbbá az is, hogy ez az n fa az őket alkotó m megfigyelés egy részhalmazán eltérhet egymástól. A fentebb említett módon a módszer része a független változókból való mintavétel is, amelyre minden döntési pontban sor kerül, az adott pontban visszatevés nélkül (Nyitrai, 2021).

A véletlenerdő eljárással először azt a kérdést vizsgáltam Ishwaran (2007) módszerét követve, hogy a korlátozási indexek alkotó 9 mutató között vannak-e olyanok, amelyek nagyobb hangsúllyal szerepelnek a korlátozási indexben. A változó fontossági (VIMP) mutató a bináris regressziós fákat tartalmazó véletlen erdőben a VIMP a változó véletlen permutációval zajosabbá tett és eredeti becsült értéke különbségeként határozható meg. A módszer változószelekcióra is alkalmas (Ishwaran, 2007), olyan módon, hogy az alacsony fontossági (importance) értékkel rendelkező változókat eltávolítjuk (ez ugyanis túllillesztéshez vezetne), majd információs kritériumok segítségével összehasonlítjuk az eredményeinket. Az információs output az az érték, amennyivel csökkenne a modell magyarázó ereje, ha azt a változót elhagynánk, amihez a kapott eredmény tartozik.

VI.3. A véletlenerdő-eljárás eredménye

A korlátozási indexet alkotó mutatók, az előző éves munkanélküliség alakulása, illetve a turizmusban alkalmazottak arányát tartalmazó modell esetén a fontossági sorrend a következők szerint alakult:

8. táblázat: A korlátozási indexet alkotó mutatók véletlenerdő eljárással nyert fontossági sorrendje, csökkenő sorrendbe rendezve

Sorszám	Mutató jele	Mutató neve
1.	H1	Nyilvános tájékoztató kampányok szintje
2.	C8	Nemzetközi utazások korlátozása
3.	$u_{i,t-1}$	Előző éves munkanélküliségi ráta változás logaritmusa
4.	C2	Munkahelyek bezárása
5.	C3	Nyilvános események lemondása
6.	C1	Iskola bezárása
7.	C7	Belföldi utazások korlátozása
8.	$jlh19_i$	Turizmus-vendéglátásban alkalmazottak aránya 2019-ben
9.	C4	Gyülekezési korlátozás
10.	C6	Lakóhely elhagyásának korlátozása
11.	C5	Tömegközlekedés leállítása

Az eredmény szerint kiemelkedően jelentős a H1 mutató, a lakosság tájékoztató kampányok szervezettsége, amely egy új eredmény a vizsgált szakirodalmakhoz képest, hiszen egy eddig kevéssé vizsgált mutatóról beszélünk. A propaganda szintjének erősödése önmagában nyilvánvalóan nem okozhatja a munkanélküliségi ráta változását, azonban határozottan a várakozásnak megfelelő országspecifikus mutatóról beszélhetünk. A jól szervezett médiakampányok és a lakosság számára határozottan kommunikált üzenetek esetén ugyanis

nagyfokú tudatos járványkezelésről beszélhetünk, ahol jelentős mértékű kormányzati figyelem arányul a lakosságra.

Az eredmény továbbá megerősíti a Del Boca et al. (2020), illetve Fabrizio et al. (2021) eredményeit, miszerint az otthonról történő munkavégzés szintje fontos mutató a munkanélküliség alakulására nézve. A C8 mutató Nelson (2021) által bemutatott hatásokat tükrözheti: a nemzetközi mobilitás változása felerősítheti a munkanélküliségi ráták változásának tovaryűrűző hatásait. Az eredmények szerint a munkanélküliségre kevésbé hatnak lokális jellegű mutatók, sőt elhagyásukkal a modell magyarázó ereje javítható. Nelson (2021) eredményei szerint a tömegközlekedési mutató a fejlődő országokban van szignifikáns jelen a foglalkoztatottsági folyamatokban, azonban az általam vizsgált országok nem tartoznak ebbe a kategóriába.

Összevetve Nelson (2021) 24 országra kiterjedő kérdőíves felméréssel, illetve logit becsléssel végzett kutatásával, a nemzetközi utazások korlátozása mellett a nyilvános események lemondása, illetve a tömegközlekedésre vonatkozó korlátozások bizonyultak a foglalkoztatottságra leginkább befolyással bíró változóknak, bár ez a becslés a H1 változót nem tartalmazta.

A turizmusban foglalkoztatottak inszignifikanciáját a lineáris becslés után a véletlenerdő-eljárás is megerősítette. Doti (2021) eredményei szerint több munkahely megőrzésére volt lehetőség azokban az amerikai államokban, ahol az előző évben is hasonló trend volt megfigyelhető. Az előző éves trendek fontosságát tükrözi az általam kapott eredmény is.

VII. Összegzés

Dolgozatomban arra a kérdésre kerestem a választ, hogy hatással van-e a munkanélküliségi ráta 2020-ban bekövetkezett változására az, hogy egy adott ország milyen szigorítási intézkedéseket alkalmazott az év során a koronavírus-járvánnyal szembeni védekezés céljából.

A szakirodalomban elterjedt éves lineáris modell teszteléséhez először előállítottam a 2019-es évre még nem elérhető turizmusban foglalkoztatottak arányát, illetve az éves korlátozási indexet, majd elvégeztem a becslést. A modell eredményei szerint az Oxfordi Egyetem által kidolgozott korlátozási index szignifikáns hatással volt a munkanélküliségi ráta változására, azonban az intuícióval ellentétes előjelű eredményre vezetett. Ahol magasabb volt a korlátozások éves átlagos mértéke ott a munkanélküliség is hozzávetőlegesen kevésbé emelkedett, míg ahol a vizsgálatot megelőző évben is nagyobb emelkedést tapasztalhattunk, ott 2020-ban is jobban emelkedett a munkanélküliség.

Az elterjedt OLS módszer tesztelése után az éves adatokat tovább finomítottam havi gyakoriságú megfigyelésekre. A korábban leginkább a fertőzöttek és halálesetek számának előrejelzésére használt panelregressziós modell már a várakozásoknak megfelelő eredményre vezetett, vagyis magasabb korlátozási szint mellett nagyobb növekedést jelzett a munkanélküliségi ráta esetén is.

Azonban a paneladatokon erős fixhatások tapasztalhatók, így valószínűsíthető, hogy a korlátozási index erősen pozitívan korrelál valamilyen országspecifikus mutatóval. A szakirodalomban elterjedt nézetek szerint a fiskális politika magyarázatot adhat a különböző országok esetén tapasztalt eltérésekre, hiszen az európai országok meglehetősen eltérő fiskális intézkedéseket alkalmaztak a járvány 2020-as szakaszában.

A modell továbbfejlesztése céljából lefuttatott gépi tanulásra épülő véletlenerdő-eljárás rávilágított arra, hogy az index H1 jelű „információs tájékoztató kampányok szintje” komponense bír a legnagyobb előrejelző erővel a munkanélküliségi ráta alakulására vonatkozóan. Ennek a mutatónak az elemzése a szakirodalomban a munkaerőpiac kapcsán még nem elterjedt, azonban tükrözi az országok egyedi hatásainak paneladatok által is megerősített kiemelt szerepét.

A téma további kutatásának lehetséges területe egy olyan fiskális index kidolgozása a munkanélküliségi ráta előrejelzése céljából, amely széles körben használható az európai

országok esetében. A korlátozási index munkanélküliségi rátára gyakorolt hatásának elemzése is tovább folytatható, a kérdést a munkavégzés rendje-, a nem-, illetve életkor szerinti bontásban vizsgálva. A dolgozat eredményeinek felhasználásával egy lehetséges további vizsgálati irány a korlátozások európai országokra is kiterjedő költség-haszon elemzésének elkészítése.

Hivatkozásjegyzék

Airy, Sir G. B. (1861): On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations. Macmillan, Cambridge–London.

Anderson, J. – Bergamini, E. – Brekelmans, S. – Cameron, A. – Darvas, Zs. – Jiménez, M. D. – Lenaerts, K. – Midoes, C. (2020): The fiscal response to the economic fallout from the coronavirus. <https://www.bruegel.org/publications/datasets/covid-national-dataset/> , letöltés dátuma: 2021. 12. 23.

Apeti, A. E. – Combes, J.-L. – Debrun, X. – Minea, A. (2021): Did fiscal space influence Covid-19's fiscal response? Covid Economics, 74. sz. 71.-93. o.

Baek C. W. – McCrory, P. – Mui, P. (2020): Unemployment Effects of Stay-at-Home Orders: Evidence from High Frequency Claims Data. The Review of Economics and Statistics, 1–72. o. https://doi.org/10.1162/rest_a_00996, letöltés dátuma: 2021. 02. 18.

Balázs L. – Divényi J. K. – Kézdi G. – Mátyás L. (2014): A közgazdasági adatforradalom és a panelökonometria. Közgazdasági Szemle, 61. évf. 11. sz. 1319–1340. o.

Breiman, L. – Friedman, J. H. – Olshen, R. – Stone, C. J. (1984). Classification and Regression Trees. Wadsworth International Group, Belmont

Breiman, L. (2001): Random Forest. Machine Learning, 45. sz. 5-32. o.

Chernozhukov, V. – Kasahara, H. – Schrimpf, P. (2021): The association of opening K-12 schools and colleges with the spread of Covid-19 in the United States: County-level panel data analysis. Covid Economics, 70. sz. 70-108. o.

Choubin, B. – Darabi, H. – Rahmati, O. – Sajedi-Hosseini, F. – Klove, B. (2018): River suspended sediment modelling using the CART model: A comparative study of machine learning techniques. Science of The Total Environment, 615 sz. 272-281. o.

Croissant, Y. – Millo, G. (2008): Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, 27. évf. 2. sz. 1-43. o.

Del Roca, D. – Oggero, N. – Profeta, P. – Rossi, M. C. (2020): Women's and Men's Work, Housework and Childcare, before and during COVID-19. Review of Economics of the Household, 2020. 13. sz. 1001-1017. o.

- Del Roca, D. – Oggero, N. – Profeta, P. – Rossi, M. C. (2021): Household division of labor during two waves of COVID-19 in Italy. *Covid Economics*, 75. sz. 60.- 80. o.
- Doti J. L. (2021): Benefit-cost analysis of COVID-19 policy intervention at the state and national level. *Covid Economics*, 67. sz. 94-127.o.
- Eisenhart, C. (1947): The Assumptions Underlying the Analysis of Variance. *Biometrics*, 3. évf. 1. sz. 1–21. o.
- Fabrizio, S. – Gomes, D. B. P. – Tavares, M. M. (2021): The COVID-19 she-cession: The employment penalty of taking care of young children. *Covid Economics*, 72. sz. 136.-166. o.
- Fairlie, R. W. – Couch, K. – Xu, H. (2020): The Impacts of COVID-19 on Minority Unemployment: First Evidence from April 2020 CPS Microdata. DOI: 10.3386/w27246, <https://www.nber.org/papers/w27246>, letöltés dátuma: 2021. 03. 09.
- Fisher, R. A. (1925): *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh–London
- Guaitoli, G. – Tochev, T. (2021): Do localised lockdowns cause labour market externalities? *Covid Economics*, 69. sz. 46-86. o.
- Hale, T. – Webster, S. – Petherick, A. – Philips, T. – Kira, B. (2020): Oxford COVID-19 Government Response Tracker, Blavatnik School of Government. *Natural Human Behaviour*, <https://doi.org/10.1038/s41562-021-01079-8>, letöltés dátuma: 2021. 02. 04.
- Hanck C. – Arnold M. – Gerber A. – Schmelzer M. (2020): *Introduction to Econometrics with R*. University of Duisburg-Essen, Németország, Essen
- Hanck, C. _ Arnold, M. – Gerber, A. – Schmelzer, M. (2020): *Introduction to Econometrics with R*. <https://www.econometrics-with-r.org/ITER.pdf>, letöltés dátuma: 2021. 13. 20.
- Hayashi Y. – Oishi T. (2018): High accuracy-priority rule extraction for reconciling accuracy and interpretability in credit scoring. *New Generation Computing*. 36. évf. 8. sz. 393–418. o.
- Hershbein, B. J. – Holzer, H. J. (2021): The COVID-19 pandemic’s evolving impacts on the labour market: Who’s been hurt and what we should do. *Covid Economics*, 72. sz. 68.-120. o.
- Hildreth, C. (1950): *Combining Cross Section Data and Time Series*. Cowles Commission. Discussion Paper, 347. sz. 1-13. o.

Hosny, A. (2021): The sooner (and the smarter), the better: COVID-19 containment measures and fiscal responses. Covid Economics, 73. sz. 137.-157. o.

Ishwaran, H. (2007): Variable importance in binary regression trees and forests. Electronic Journal of Statistics, 1. évf. 4.sz. 519–537. o.

James G. – Witten D. – Hastie T. – Tibshirani R. (2013): An Introduction to Statistical Learning with Applications in R. Springer, New York 2013

Kong, E. – Prinz, D. (2020): Disentangling policy effects using proxy data: Which shutdown policies affected unemployment during the COVID-19 pandemic? Journal of Public Economics, 189. évf. 104257. sz. 1-39. o.

Köllő J. (2020): Foglalkoztatás a koronavírus-járvány első hullámának idején. Közelkép, 220-234. o.

Kumar, M. – Thenmozhi, M. (2006): Forecasting stock index movement: A comparison of support vector machines and random forest. Indian Institute of Capital Markets 9th Capital Markets Conference, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=876544, letöltés dátuma: 2021. 04. 10.

Madaras Sz. (2009): A munkanélküliség elemzése a Központi Régió megyéiben. Közgazdász Fórum, 2009. 8. sz. 45-58. o.

Maruthappu, M. – Da Zhou, C. – Williams, C. – Zeltner, T. – Atun, R. (2015): Unemployment, Public-sector Health Care Expenditure and HIV Mortality: An Analysis of 74 Countries, 1981-2009. Journal of Global Health, 5. évf. 1. sz. 1-9. o.

Moreno, A. – Ongena, S. – Veghazy, A. V. – Wagner, A. F. (2021): Misfortunes never come alone: From the financial crisis to the Covid-19 pandemic. Covid Economics, 71. sz. 59.-82. o.

Nelson, M. A. (2021): COVID-19 closure and containment policies: A first look at the labour market effects in emerging nations. Covid Economics, 66. sz. 89.-114. o.

Nerlove, M. (2005): Essays in Panel Data Econometrics. Cambridge University Press, Cambridge

Nyitrai T. (2021): A gépi tanulás módszereinek alkalmazása R-ben. Statisztikai szemle, 99. évf. 2. sz. 173–198. o.

OECD Foglalkoztatottsági előrejelzés (2020): COVID-19: From a health to a jobs crisis. OECD Employment Outlook 2020, 21-115. o.

- Oxfordi COVID-19 Kormányzati reakció követő (2020): Calculation and presentation of the Stringency Index 4.0. https://github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/interpretation_guide.md,_letöltés dátuma: 2020. 03. 15.
- Palmer, D. – O’Boyle, N. – Glen, R. – Mitchell, J. (2007). Random forest models to predict aqueous solubility. *Journal of Chemical Information Modelling*, 47. évf. 1. sz. 150–158. o.
- Pesaran, M. H. (2004): General Diagnostic Tests for Cross Section Dependence in Panels. IZA Discussion Paper, 1240. sz. 1-39. o.
- Petersen, E.– Koopmans, M.–Go, U.–Hamer, D. H.–Petrosillo, N.–Castelli, F.–Storgaard, M.–Al Khalili, S.–Simonsen, L. (2020): Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases*, 20. évf. 9. sz. 238–244. o.
- Scheffé, H. (1956): Alternative Models for the Analysis of Variance. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, 27. évf. 2. sz. 251–271. o.
- Tattar, P. N. (2018): Hands-On Ensemble Learning with R. Packt Publishing, Birmingham
- Várad B. (2020): Magyarországi közpolitikai reakciók a koronavírus-járványra 2020 első félévében. *Közlelkép*, 213-219. o.
- Zádori I. – Nemeskéri Zs. – Szabó Sz. (2020): Deglobalizáció vagy reglobalizáció? Munkaerőpiac a vírus előtt, alatt és után. *Új Munkaügyi Szemle*, 1. évf. 3. sz. 2-13. o.
- Wooldridge, J.M. (2002): *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, Cambridge
- Xia, Y. – Zhao, J. – He, L. – Li, Y. – Niu, M. (2020): A novel tree-based heterogeneous ensemble method for credit scoring. *Expert Systems with Applications*, 159. évf. 113615. sz. 1-18. o.
- Xu, R. (2013): Improvements to random forest methodology. <https://core.ac.uk/download/pdf/38926405.pdf>, letöltés dátuma: 2021. 04. 09.

MELLÉKLETEK

1. számú melléklet: A vizsgált európai országok listája

- | | |
|----------------------------|-------------------|
| 1. Ausztria | 17. Lettország |
| 2. Belgium | 18. Litvánia |
| 3. Bulgária | 19. Luxemburg |
| 4. Ciprus | 20. Magyarország |
| 5. Csehország | 21. Málta |
| 6. Dánia | 22. Németország |
| 7. Egyesült királyság (UK) | 23. Norvégia |
| 8. Észtország | 24. Olaszország |
| 9. Finnország | 25. Portugália |
| 10. Franciaország | 26. Románia |
| 11. Görögország | 27. Spanyolország |
| 12. Hollandia | 28. Svájc |
| 13. Horvátország | 29. Svédország |
| 14. Írország | 30. Szlovákia |
| 15. Izland | 31. Szlovénia |
| 16. Lengyelország | |

2. számú melléklet: A korlátozási indexet alkotó mutatók bemutatása

Mutató jele	Mutató megnevezése	Lehetséges értékek	Értékek magyarázata
C1	Iskolák bezárása	0	Nincsenek korlátozások.
		1	Az iskolák bezárása javasolt.
		2	Csak meghatározott szinten, részlegesen kötelező bezárások.
		3	Minden oktatási szinten kötelező az iskolák bezárása.
		NA/ üres	Nincs adat.
C2	Munkahelyek bezárása	0	Nincsenek korlátozások.
		1	A munkahelyek bezárása és/vagy otthoni munkavégzés javasolt.
		2	A munkahelyek bezárása vagy otthoni munkavégzés csak meghatározott szektorokban vagy más kijelölt munkavállalói csoportok esetén kötelező.
		3	Minden szinten kötelezően a munkahelyek bezárásra vagy az otthoni munkavégzésre átállás, kivéve a létfontosságú munkahelyek esetén (például élelmiszerüzletek, orvosi rendelők).
		NA/ üres	Nincs adat.
C3	Nyilvános események lemondása	0	Nincsenek korlátozások.

		1	A nyilvános események lemondása ajánlott.
		2	A nyilvános események lemondása kötelező.
		NA/ üres	Nincs adat.
C4	Gyülekezési korlátozás	0	Nincsenek korlátozások.
		1	Szigorítási intézkedések bevezetése 1000 fő feletti létszámkorlát mellett.
		2	Szigorítási intézkedések bevezetése a 100 és 1000 fő közé eső létszámú események esetén is.
		3	Szigorítási intézkedések bevezetése a 10 és 100 fő közé eső létszámú események esetén is.
		4	Szigorítási intézkedések bevezetése a 10 fő alatti események esetén is.
		NA/üres	Nincs adat.
C5	Tömegközlekedés leállítása	0	Nincsenek korlátozások.
		1	A tömegközlekedés leállítása ajánlott, vagy szignifikáns csökkenés figyelhető meg az igénybe vehető tömegközlekedési eszközök esetén: a járatok számában, az útvonal hosszában, igénybe vehető tömegközlekedési eszközök fajtáiban.
		2	A leállítás kötelező, vagy a legtöbb lakos számára tiltott az igénybevétele.
		NA/üres	Nincs adat.

C6	Lakóhely elhagyásának korlátozása	0	Nincsenek korlátozások.
		1	A lakóhely elhagyása nem ajánlott.
		2	A lakóhely elhagyása csak a napi testmozgás, élelmiszer vásárlás és elkerülhetetlenül szükséges utazások megtétele esetén lehetséges.
		3	A lakóhely elhagyása csak minimális számú kivételes esetben lehetséges (például a lakóhely elhagyása hetente egyszer engedélyezett, vagy egyszerre csak egy ember hagyhatja el a házat.).
		NA/üres	Nincs adat.
C7	Belföldi utazások korlátozása;	0	Nincsenek korlátozások.
		1	Országban belüli utazások korlátozása ajánlott régiók vagy városok között.
		2	Országban belüli utazások korlátozásának bevezetése.
		NA/üres	Nincs adat.
C8	Nemzetközi utazások korlátozása	0	Nincsenek korlátozások.
		1	A beutazók szűrése és/vagy tesztelése.
		2	A magas víruskockázatú országokból érkezőkre kötelező karantén bevezetése.

		3	A magas kockázatú országokba történő kiutazás és az onnan történő fogadás megtiltása.
		4	Teljes határzár.
		NA/üres	Nincs adat.
H1	Nyilvános tájékoztató kampányok szintje	0	Nincsen Covid-19-hez köthető nyilvános tájékoztató kampány.
		1	Az állami tisztviselők felhívják a figyelmet a COVID-19 járványra.
		2	Koordinált nyilvános tájékoztató kampányok, például a hagyományos- vagy a közösségi médiában.
		NA/üres	Nincs adat.

3. számú melléklet: A lineáris modell futtatásához szükséges programkód

#Adatforrások beolvasása és tárolása:

#korlátozási index

```
covidmeasures <- data.frame(read.csv("C:/Users/Asus/Documents/corvinus/6/TDK_makro/
                                     data/total.csv"))
```

#éves munkanélküliségi ráta változás:

```
annual_un    <- read_excel("C:/Users/Asus/Documents/corvinus/6/TDK_makro/data/
                             tisztított/unemployment_data_annual.xlsx" )
```

#2019-es adatok, az összes foglalkoztatottan belül a turizmusban foglalkoztatottak aránya

```
turism_employment_rate <- read_excel("C:/Users/Asus/Documents/corvinus/6/TDK_makro/
                                       data/tisztított/tourism_employment_rate_eurostat_i_n.xlsx")
```

#elkészítem a vizsgált országok neveit tartalmazó listát

#(összesen 31, az 1. sz. melléklet szerint):

```
countries      <- list("Austria", "Belgium", "Bulgaria", "Croatia", "Cyprus",
                        "Czechia", "Denmark", "Estonia", "Finland", "France",
                        "Germany", "Greece", "Hungary", "Iceland", "Ireland",
                        "Italy", "Latvia", "Lithuania", "Luxembourg", "Malta",
                        "Netherlands", "Norway", "Poland", "Portugal", "Romania",
                        "Slovakia", "Slovenia", "Spain", "Sweden", "Switzerland",
                        "United Kingdom")
```

```
countries2     <- data.frame(countries)
```

#az éves korlátozási index átlagok meghatározása

```
library("dplyr")
```

```
for (i in 1:31) {
```

```
  covm_con      <- covidmeasures %>% group_by(CountryName) %>%
    filter(all(countries2[,i] %in% CountryName)) %>%
    filter(Date<=20201231)
```

```
  cov_m_annual <- filter(covm_con, Date<=20201231)
```

```
  strin_an_country[i] <- mean(cov_m_annual$stringencyIndex)}
```

#az adattábla elkészítése:

```
strin_n_an_country      <- as.numeric(unlist(strin_an_country))
countries3              <- as.character(unlist(countries2))
turism_emp_rate         <- as.numeric(unlist(turism_employment_rate$`employment
rate in tourism 2019`)))
```

#munkanélküliségi ráták logaritmizálása:

```
annual_un$`20192020`    <- log(annual_un$`20192020`, base=exp(1))
annual_un$`20202021`    <- log(annual_un$`20202021`, base=exp(1))
```

a modell tartalmazza a turizmusra vonatkozó adatokat is:

```
reg_an_with_turism      <- data.frame(cbind(countries3, annual_un$`20192020`,
                                             annual_un$`20202021`,      strin_n_an_country,
                                             turism_emp_rate))
```

#adattisztítás:

```
reg_an_with_turism2     <- reg_an_with_turism %>%
  filter(reg_an_with_turism$strin_n_an_country!="NaN"
  & reg_an_with_turism$V3!="NaN"
  & reg_an_with_turism$V2!="NaN"
  & reg_an_with_turism$V3!="-Inf")
reg_an_with_turism2     <- drop_na(reg_an_with_turism2)
```

#az eldobott megfigyelések száma:

```
length(reg_an_with_turism$countries3)-length(reg_an_with_turism2$countries3)
```

#megmaradt megfigyelések száma:

```
length(reg_an_with_turism2$countries3)
```

#az első lineáris modell illesztése:

```
emp_rate_tur           <- unfactor(reg_an_with_turism2$turism_emp_rate)
un_annual_2020_wt       <- unfactor(reg_an_with_turism2$V3)
s_annual_wt             <- unfactor(reg_an_with_turism2$strin_n_an_country)
un_annual_2019_wt       <- unfactor(reg_an_with_turism2$V2)
```

```
ols_annual_with_turism    <- lm(un_annual_2020_wt ~ s_annual_wt +  
                                un_annual_2019_wt + emp_rate_tur)
```

a turizmus változó elhagyásával végzett illesztés:

```
ols_annual_without_tourism <- lm(un_annual_2020_wt ~ s_annual_wt +  
                                un_annual_2019_wt )
```

4. számú melléklet: A panelmodell futtatásához szükséges programkód

#havi munkanélküliségi ráták változásának beolvasása Excel fájlból:

```
monthly_un <- read_excel("C:/Users/Asus/Documents/corvinus/6/TDK_makro/  
data/tisztított/unemployment_monthly_data_2019_2020_.xlsx")
```

#időintervallumok kialakítása listában:

```
d <- c (20191231, 20200131, 20200229, 20200331, 20200430, 20200531,  
20200630, 20200731, 20200831, 20200930, 20201031, 20201130, 20201231,  
20210131)
```

```
month_names <- c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct",  
"Nov", "Dec")
```

havi stringency indexek létrehozás:

#készítünk üres listákat amiben eltároljuk a kapott eredményeket

```
strin_mon_country <- list()
```

```
country_month <- list()
```

```
month_name_list <- list()
```

```
econ_index_monthly <- list()
```

#az i az országok futóindexe, a j a hónapokat meghatározó dátumoké a következő ciklusban:

```
for (i in 1:31) {
```

```
  for (j in 1:12){
```

```
    covm_con_mon <- covidmeasures %>% group_by(CountryName) %>%  
      filter(all(countries2[,i] %in% CountryName)) %>%  
      filter(Date<=20201231)
```

```
    cov_m_mon <- filter(covm_con_mon, Date>d[j] & Date<=d[j+1] )
```

```
    strin_mon_country[(i-1)*12+j] <- mean(cov_m_mon$StringencyIndex)
```

```
    econ_index_monthly[(i-1)*12+j] <- mean(cov_m_mon$EconomicSupportIndex)
```

```
    country_month[(i-1)*12+j] <- countries[i]
```

```
    month_name_list[(i-1)*12+j] <- month_names[j]
```

```
  }
```

```
}
```

#ezt követően wide-ról long formájú adattá transzponálom a havi munkanéliségi adatokat

```
un_month_long2019      <- list()
un_month_long2020      <- list()
un_month_longcountry    <- list()
for (i in 1:31){
  for (j in 1:12) {
    un_month_longcountry[(i-1)*12+j] <- monthly_un[i,1]
    un_month_long2019[(i-1)*12+j]   <- monthly_un[i,j+1]
    un_month_long2020[(i-1)*12+j]   <- monthly_un[i,j+13]
  }
}
```

#adattábla elkészítése a munkanélküliségi változókból, majd logaritmizálásuk:

```
un_month_long      <- data.frame(cbind(un_month_longcountry,      month_name_list,
un_month_long2019, un_month_long2020))

un_month_long$un_month_long2019      <-
  log(unlist(un_month_long$un_month_long2019),
  base=exp(1))

un_month_long$un_month_long2020      <-
  log(unlist(un_month_long$un_month_long2020),
  base=exp(1))
```

#adattáblát készítünk, amibe beletesszük az országok nevét és a kétféle havi munkanélküliségi ráta változásokat

```
data_monthly      <- data.frame(cbind(country_month,month_name_list,
un_month_long$un_month_long2019,
un_month_long$un_month_long2020, strin_mon_country))
```

#megtisztítjuk az adatot a hibás megfigyelésektől, az előbbihez hasonló hibáktól

```
data_monthly2      <- data_monthly %>% filter(data_monthly$V3!="NaN"      &
data_monthly$V3!="-Inf"      & data_monthly$V4!="NaN"      &
data_monthly$V4!="-Inf"      & data_monthly$V3!="NA"      &
data_monthly$V4!="NA"      &
data_monthly$strin_mon_country!="NaN")

data_monthly2 <- drop_na(data_monthly2)
```

#az eldobott megfigyelések száma:

```
length(data_monthly$country_month)- length(data_monthly2$country_month)
```

#megmaradt megfigyelések száma:

```
length(data_monthly2$country_month)
```

A plm csomag használatához szükséges formára kell rendezni az adattáblát:

```
un_monthly_2020 <- unlist(data_monthly2$V4)
```

```
s_monthly <- unlist(data_monthly2$strin_mon_country)
```

```
un_monthly_2019 <- unlist(data_monthly2$V3)
```

```
country_d <- as.factor(unlist(data_monthly2$country_month))
```

```
month_d <- as.factor(unlist(data_monthly2$month_name_list))
```

```
data_monthly2$country_d <- country_d
```

```
data_monthly2$month_d <- month_d
```

#a faktorváltozóra alakított ország és hónap értékeknek az első oszlopokban kell szerepelniük:

```
data_monthly3 <- data_monthly2[, c("country_d", "month_d", "country_month",  
                                   "month_name_list", "un_month_long2019",  
                                   "un_month_long2020", "strin_mon_country" ) ]
```

```
library("plm")
```

1.1.: within modell (fixed-effects):

```
plm_monthly_un_within <- plm( un_monthly_2020 ~ s_monthly,  
                               data = data_monthly3,  
                               index=c("country_d", "month_d"),  
                               model = "within",  
                               effect = "twoways")
```

#csak ország dummy esetén:

```
plm_monthly_un_within_coun <- plm( un_monthly_2020 ~ s_monthly,  
                                   data = data_monthly3,  
                                   index=c("country_d"),
```

```
model = "within" )
```

1.2.: pooling modell (nincsenek fixhatások)

```
plm_monthly_un_pooling <- plm(un_monthly_2020 ~ s_monthly ,  
                               data = data_monthly3,  
                               index=c("country_d", "month_d"),  
                               model = "pooling",  
                               effect = "twoways")
```

1.3.: between modell (mintha éves adatokkal dolgoznánk a)

```
plm_monthly_un_between <- plm(un_monthly_2020 ~ s_monthly,  
                               data = data_monthly3,  
                               index=c("country_d", "month_d"),  
                               model = "between")
```

az előző három modellhez tartozó együtthatók:

```
summary(plm_monthly_un_within)$coeff  
summary(plm_monthly_un_pooling)$coeff  
summary(plm_monthly_un_between)$coeff
```

5. számú melléklet: A véletlenerdő-modell futtatásához szükséges programkód

#a korlátozási indexet alkotó mutatók éves átlagos értékének előállítása:

```
c1_an <- list()
c2_an <- list()
c3_an <- list()
c4_an <- list()
c5_an <- list()
c6_an <- list()
c7_an <- list()
c8_an <- list()
h1_an <- list()
```

#a következő ciklus az éves az első modellnél használt éves korlátozási index meghatározásával egyszerre is futtatható

```
library("dplyr")
for (i in 1:31) {
  covm_con      <- covidmeasures %>% group_by(CountryName) %>%
  filter(all(countries2[,i] %in% CountryName)) %>%
  filter(Date<=20201231)

  cov_m_annual  <- filter(covm_con, Date<=20201231)
  c1_an[i]      <- mean(cov_m_annual$C1_School.closing)
  c2_an[i]      <- mean(cov_m_annual$C2_Workplace.closing)
  c3_an[i]      <- mean(cov_m_annual$C3_Cancel.public.events)
  c4_an[i]      <- mean(cov_m_annual$C4_Restrictions.on.gatherings)
  c5_an[i]      <- mean(cov_m_annual$C5_Close.public.transport)
  c6_an[i]      <- mean(cov_m_annual$C6_Stay.at.home.requirements)
  c7_an[i]      <- mean(cov_m_annual$C7_Restrictions.on.internal.movement)
  c8_an[i]      <- mean(cov_m_annual$C8_International.travel.controls)
  h1_an[i]      <- mean(cov_m_annual$H1_Public.information.campaigns)
}
```

```

c1_an <- unlist(c1_an)
c2_an <- unlist(c2_an)
c3_an <- unlist(c3_an)
c4_an <- unlist(c4_an)
c5_an <- unlist(c5_an)
c6_an <- unlist(c6_an)
c7_an <- unlist(c7_an)
c8_an <- unlist(c8_an)
h1_an <- unlist(h1_an)

```

#adattábla kialakítása a korábban bevezetett változók felhasználásával:

```

rf_data <- data.frame(cbind(countries3, annual_un$`20192020`,
annual_un$`20202021`, turism_emp_rate, c1_an, c2_an, c3_an,
c4_an, c5_an, c6_an, c7_an, c8_an,
e1_an, e2_an, e3_an, e4_an, h1_an, h2_an, h3_an, h4_an, h5_an, h6_an, h7_an
))

```

#véletlerdő modell:

```

library("varhandle")
library("randomForestSRC")
rf_data2$unemployment <- unfactor(rf_data2$V3)
korl_index_fa <- rfsrc(unemployment~ unemp_t1+ turism_emp_rate + c1_an
+c2_an+c3_an+c4_an+c5_an+c6_an+c7_an+c8_an+h1_an,
data = rf_data2,
importance=TRUE)

```

#fontossági sorrend output:

```
korl_index_fa$importance
```