

Priestorová analýza kriminality: jednorozmerný a viacrozmerný prístup

Michaela Chocholatá¹

Abstrakt

Príspevok sa zaoberá priestorovou analýzou kriminality v NUTS 3 (Nomenclature of Units for Territorial Statistics) regiónoch krajín Vyšehradskej štvorky. Využíva vybrané nástroje exploratanej analýzy priestorových dát ESDA, ktoré umožňujú skúmať štruktúru analyzovaných údajov, zistiť prítomnosť priestorovej závislosti a existenciu priestorových zhlukov. Lokálna priestorová závislosť bola preskúmaná pomocou jednorozmernej Gearyho štatistiky a pri uvažovaní ďalších relevantných premenných (HDP na obyvateľa, miera zamestnanosti a hustota obyvateľstva) tiež pomocou viacrozmernej Gearyho štatistiky. Kým lokálna jednorozmerná Gearyho štatistika umožňuje identifikovať podobnosť alebo rozdielnosť hodnôt v susedných regiónoch pri uvažovaní jednej premennej, jej viacrozmerná verzia umožňuje porovnať geograficky susedné regióny z viacrozmerného hľadiska.

Kľúčové slová

kriminalita, priestorová analýza, lokálna jednorozmerná a viacrozmerná Gearyho štatistika

Abstract

This paper deals with the spatial analysis of crime in the NUTS 3 (Nomenclature of Units for Territorial Statistics) regions of the Visegrad Four countries. It uses selected tools of exploratory spatial data analysis ESDA, which allow to examine the structure of the analysed data, to analyse the presence of spatial dependence and the existence of spatial clusters. Local spatial dependence was examined using the univariate Geary statistics and under consideration of the other relevant variables (GDP per capita, employment rate and population density) the multivariate Geary statistics was used, as well. Since the local univariate Geary statistic makes it possible to identify the similarity or difference of values in neighbouring regions when considering one variable, its multivariate counterpart allows to compare geographically neighbouring regions from a multidimensional point of view.

Key words

Crime, Spatial analysis, Local univariate and multivariate Geary statistics

JEL classification

K42, R15

1 Úvod

Analýza kriminality predstavuje zaujímavú oblasť nielen z pohľadu analýzy jej príčin a motivácie páchatel'ov k spáchaniu trestného skutku, ale aj z priestorového hľadiska. Je totiž zrejmé, že z časového i priestorového hľadiska nie je výskyt trestných skutkov náhodný. S rozmachom priestorovej analýzy dát ako aj popularizáciou nástrojov priestorovej ekonometrie, sa popri rôznych kriminologických smeroch (napr. sociologický, psychologický, biologický) do popredia opäť dostáva faktor miesta, resp. lokalita (viac pozri Ratcliffe, 2010). Za priekopnícku prácu v oblasti analýzy kriminality s ohľadom na faktor miesta možno považovať prácu Guerryho z roku 1833 (viac pozri Friendly, 2007), ktorý zhromažďoval údaje

¹ Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky, Katedra operačného výskumu a ekonometrie, Dolnozemska cesta 1, 852 35 Bratislava, michaela.chocholata@euba.sk.

o zločinoch, samovraždách, gramotnosti a ďalších „morálnych štatistikách“ pre francúzske kraje. Ako uvádzajú Dray a Jombart (2011), išlo o prvú analýzu na báze skutočných sociálnych údajov pomocou grafov a máp, ktoré zosumarizovali mnohorozmerný súbor georeferenčných údajov. Friendly (2007) popisuje Guerryho život, prácu, použité metódy a analýzy, ale tiež aplikuje na rovnaké dáta moderné prístupy zohľadňujúce tak viacrozmerne ako aj geografické aspekty dát. Dray a Jombart (2011) prezentujú rôzne spôsoby začlenenia priestorových obmedzení do viacrozmernej analýzy a ilustrujú tieto rôzne prístupy pomocou známeho Guerryho súboru údajov. Anselin a kol. (2000) a Ratcliffe (2010) poskytujú prehľad štúdií zaoberajúcich sa identifikáciou priestorových vzorcov v trestnej činnosti na základe použitia nástrojov exploratačnej analýzy priestorových dát (ESDA – Exploratory Spatial Data Analysis). Messner a kol. (1999) aplikáciou ESDA na preskúmanie rozdelenia vrážd potvrdili priestorové vzorce trestných skutkov. Chocholatá (2018a) realizovala priestorovú analýzu kriminality pre 113 NUTS 3 (Nomenclature of Units for Territorial Statistics) regiónov Českej republiky, Maďarska, Poľska a Slovenska a na základe nástrojov ESDA jasne potvrdila významnú pozitívnu priestorovú autokoreláciu a odhalila štatisticky významné zhluky regiónov s vysokou, resp. nízkou kriminalitou. Odhadnuté priestorové ekonometrické modely naznačujú, že miera kriminality nie je determinovaná iba ekonomickými a demografickými ukazovateľmi (HDP na obyvateľa, mierou zamestnanosti a hustotou obyvateľstva) analyzovaného regiónu, ale tiež počtom trestných skutkov v susedných regiónoch. Chocholatá (2018b), zohľadňujúc značnú priestorovú heterogenitu, prezentuje pre rovnakú údajovú základňu odhady lokálnych modelov s využitím geograficky váženej regresie (GWR – Geographically Weighted Regression).

Predmetom tohto príspevku je priestorová analýza kriminality pre rovnaký súbor údajov ako v Chocholatá (2018a) a Chocholatá (2018b) s využitím lokálnej jednorozmernej a viacrozmernej Gearyho štatistiky.

Štruktúra príspevku je nasledovná: po úvode v prvej časti, nasleduje časť venovaná metodológii popisujúca jednorozmernú a viacrozmernú Gearyho štatistiku, v tretej časti sú popísané dáta a empirické výsledky analýzy a príspevok uzatvára záver.

2 Jednorozmerná a viacrozmerná Gearyho štatistika - metodológia

Vo všeobecnosti existuje viacero nástrojov na exploratačnú analýzu priestorových dát ESDA, ktoré umožňujú skúmať štruktúru analyzovaných údajov, zisťovať prítomnosť priestorovej závislosti a existenciu priestorových zhlukov. Prvým krokom ESDA je obvykle vizualizácia dát pomocou rôznych grafov a máp. Kým s pomocou niektorých máp (box mapy, percentilové mapy, kvantilové mapy) možno síce identifikovať určité priestorové vzory, zhluky hodnôt, extrémne hodnoty či odľahlé hodnoty, tieto mapy však neposkytujú informáciu o štatistickej významnosti zhlukovania. Ďalším krokom ESDA je preto spravidla analýza priestorovej autokorelácie na globálnej i lokálnej úrovni. Priestorovú autokoreláciu možno všeobecne charakterizovať ako koreláciu premennej s jej priestorovo posunutou hodnotou, čo indikuje, že geograficky blízke regióny sú vzájomne previazané. Existuje niekoľko možností testovania priestorovej autokorelácie, pričom na základe rozsahu analýzy rozlišujeme globálne a lokálne štatistiky. Kým globálne štatistiky slúžia na meranie globálnej priestorovej autokorelácie a vyjadrené sú jednou hodnotou vzťahujúcou sa na celý súbor údajov, lokálne štatistiky zvyčajne hodnotia priestorovú autokoreláciu pre jednu konkrétnu priestorovú jednotku. Medzi najčastejšie používané štatistiky patria Moranovo I , Gearyho C a Getisova-Ordova G štatistika.

Pri analýze priestorových dát je v poslednom období dôraz kladený na lokálne hľadisko, čo sa odrazilo aj vo vzniku viacerých novších prístupov umožňujúcich analyzovať lokálnu priestorovú autokoreláciu, resp. lokálnu priestorovú heterogenitu (pre viac informácií pozri napr. Anselin, 2019a). Vzhľadom na zameranie príspevku sa budeme ďalej venovať analýze

priestorovej autokorelácie pomocou lokálnej jednorozmernej a viacrozmernej Gearyho štatistiky.

Lokálna jednorozmerná Gearyho štatistika bola prvýkrát prezentovaná v Anselin (1995) a ďalej rozpracovaná v Anselin (2019a). Možno ju zaradiť medzi lokálne indikátory priestorovej asociácie (LISA – Local Indicators of Spatial Association), a podobne ako jej globálna verzia (Geary, 1954), vychádza zo štvorcových diferencií, resp. nepodobnosti (rozdielnosti). Nízke hodnoty tejto štatistiky teda indikujú pozitívnu priestorovú autokoreláciu (podobnosť), kým jej vysoké hodnoty svedčia v prospech negatívnej autokorelácie (rozdielnosti). Lokálnu Gearyho štatistiku možno zapísať v tvare (Anselin, 2019a; Furková, 2019, Fischer a Wang, 2011):

$$c_i = \sum_j w_{ij} (x_i - x_j)^2 \quad (1)$$

kde x_i označuje hodnotu analyzovanej premennej v regióne i ($i = 1, 2, \dots, N$) a symbol w_{ij} označuje prvky priestorovej matice váh \mathbf{W} , ktorej špecifikácia predstavuje dôležitý bod analýzy (ohľadom rôznych prístupov ku konštrukcii matice \mathbf{W} pozri napr. Fischer a Wang, 2011). Anselin (2019a) poznamenáva, že vzhľadom na prítomnosť štvorcových odchýlok v matematickej formulácii (1) tejto štatistiky je irelevantné, či je analyzovaná premenná vyjadrená v pôvodnej mierke alebo v štandardizovanej forme, hoci pri výpočte mnohorozmernej štatistiky je preferovaná štandardizácia uvažovaných premenných (stredná hodnota štandardizovanej premennej je potom nulová a rozptyl nadobúda hodnotu jeden).

Na rozdiel od lokálnej Getisovej-Ordovej G štatistiky a lokálnej Moranovej I štatistiky nie je interpretácia štatistiky významnej lokálnej Gearyho štatistiky taká priamočiara. Kladná a štatisticky významná hodnota Getisovej-Ordovej G štatistiky indikuje zhuk vysokých hodnôt (hot spot), zatiaľ čo jej záporná a štatisticky významná hodnota naznačuje zhuk nízkych hodnôt (cold spot). Hodnoty lokálnej Moranovej I štatistiky sú interpretované v kontexte kvadrantov tzv. Moranovho rozptylového diagramu, pričom rozlišujeme priestorové zhuky typu vysoká-vysoká (high-high) a nízka-nízka (low-low) indikujúce podobnosť hodnôt a tiež priestorové odľahlé hodnoty (spatial outliers) typu vysoká-nízka (high-low) a nízka-vysoká (low-high) indikujúce rozdielnosť hodnôt (Anselin, 1995; Anselin, 2019a). V prípade lokálnej Gearyho štatistiky rozlišujeme tri prípady pozitívnej lokálnej priestorovej autokorelácie, a to okrem typov vysoká-vysoká (high-high) a nízka-nízka (low-low) tiež tzv. typ iná (other) v prípade, že hodnoty v analyzovanej dvojici regiónov sú síce z hľadiska ich veľkosti podobné, jedna z nich je však nižšia ako stredná hodnota, druhá naopak vyššia ako stredná hodnota (štvorcová diferenciacia v (1) teda prechádza „cez“ strednú hodnotu). Pri negatívnej lokálnej priestorovej autokoreláci je táto indikovaná vysokou hodnotou lokálnej Gearyho štatistiky (1), vzhľadom na použitie štvorcových diferencií vo výpočte nie je však možné určiť, či ide o typ vysoká-nízka (high-low) alebo nízka-vysoká (low-high).

Rozšírenie jednorozmernej lokálnej Gearyho štatistiky do viacrozmernej podoby bolo prezentované Anselinom (2019a), ktorý uvádza tiež stručný prehľad štúdií zaoberajúcich sa priestorovou autokoreláciou z viacrozmerného hľadiska. Viacrozmerná lokálna Gearyho štatistika je zvyčajne počítaná pre štandardizované premenné, ktorých stredná hodnota je nulová a rozptyl nadobúda hodnotu jeden. Viacrozmerná lokálna Gearyho štatistika je sumou individuálnych lokálnych Gearyho štatistik získaných pre každú uvažovanú premennú a zodpovedá váženému priemeru štvorcových vzdialeností (Furková, 2019) vo viacrozmernej priestore určenom uvažovanými premennými (atribútmi) medzi hodnotami v geografickej lokalite (regióne) i a jeho geografickými susedmi j definovanými priestorovou maticou váh. Ak uvažujeme k premenných indexovaných pomocou v , viacrozmernej lokálna Gearyho štatistika má potom tvar (Anselin, 2019c):

$$c_{k,i} = \sum_{v=1}^k \sum_j w_{ij} (x_{vi} - x_{vj})^2 \quad (2)$$

Z pohľadu interpretácie je dôležité zdôrazniť, že hoci je viacrozmerná lokálna Gearyho štatistika súčtom jednorozmerných štatistík, neznamená to, že lokality (regióny) identifikované ako štatisticky významné pre jednorozmernú štatistiku budú automaticky štatisticky významné aj vo viacrozmernej kontexte (Anselin, 2019a a Anselin, 2019c). Testovanie štatistickej významnosti je založené na podmienenom permutačnom prístupe (podobne ako pri jednorozmernej štatistike), odporúča sa však namiesto klasickej hladiny významnosti α využiť tzv. pseudo p -hodnotu (hladinu významnosti α vydelené počtom uvažovaných premenných k) alebo False Discovery Rate (FDR) – pre viac informácií pozri Anselin (2019b) a Anselin (2019c).

3 Dáta a empirické výsledky analýzy

Analýza v tomto príspevku vychádza z regionálnych údajov za 114 NUTS 3 regiónov krajín Vyšehradskej štvorky (14 českých regiónov, 20 maďarských regiónov, 72 poľských regiónov a 8 slovenských regiónov) a hoci je zrejmé, že z hľadiska hlbšej analýzy priestorových súvislostí by väčšiu výpovednú schopnosť malo využitie dát pre nižšie regionálne úrovne, vzhľadom na nedostupnosť, resp. neporovnateľnosť dát nebolo možné takúto analýzu zrealizovať. Získanie údajov o kriminalite bolo pomerne problematické, nakoľko regionálna databáza Eurostatu obsahuje iba historické údaje (2008 - 2010) o trestných činoch zaznamenaných políciou. Údaje na analýzu - počet registrovaných trestných činov v roku 2016 - boli získané z rôznych zdrojov: Polície Českej republiky (2018), Štatistický úrad Maďarska (Hungarian Central Statistical Office, 2018), Štatistická ročenka regiónov - Poľsko 2017 (Statistics Poland, 2018) a DATAcube. databáza Štatistického úradu Slovenskej republiky (Štatistický úrad Slovenskej republiky, 2018). Ďalším limitujúcim faktorom príspevku je použitie údajov za celkovú kriminalitu, čo bolo determinované dostupnosťou údajov za jednotlivé krajiny Vyšehradskej štvorky. Kým Štatistický úrad Slovenskej republiky vykazuje trestné činy v členení na všeobecné, ekonomické a zostávajúce, poľská Štatistická ročenka regiónov uvádza členenie na kriminálne a ekonomické trestné činy, Štatistický úrad Maďarska vykazuje trestné činy proti osobe, trestné činy proti zákonu a poriadku a trestné činy proti majetku a Polícia Českej republiky udáva členenie na vraždy, násilnú kriminalitu, mravnostnú kriminalitu, krádeže vlámaním, krádeže prosté, ostatnú majetkovú kriminalitu, majetkovú kriminalitu, ostatnú kriminalitu, obecnú kriminalitu, zostávajúcu kriminalitu, hospodársku kriminalitu, vojenské a protiústavné činy.

Kriminalita je vzhľadom na charakter analýzy v tomto príspevku vyjadrená ako prevrátená hodnota celkového počtu trestných činov v roku 2016² pripadajúca na obyvateľa konkrétneho regiónu (t.j. ako prevrátená hodnota miery kriminality). Takto vyjadrená premenná teda udáva priemerný počet obyvateľov regiónu pripadajúci na 1 trestný čin v danom regióne³. Z pohľadu viacrozmernej analýzy boli ďalej uvažované nasledujúce premenné: HDP na obyvateľa definovaný v PPS na rok 2015, miera zamestnanosti (vypočítaná ako pomer zamestnaných osôb k priemernému ročnému obyvateľstvu na výpočet údajov o regionálnom

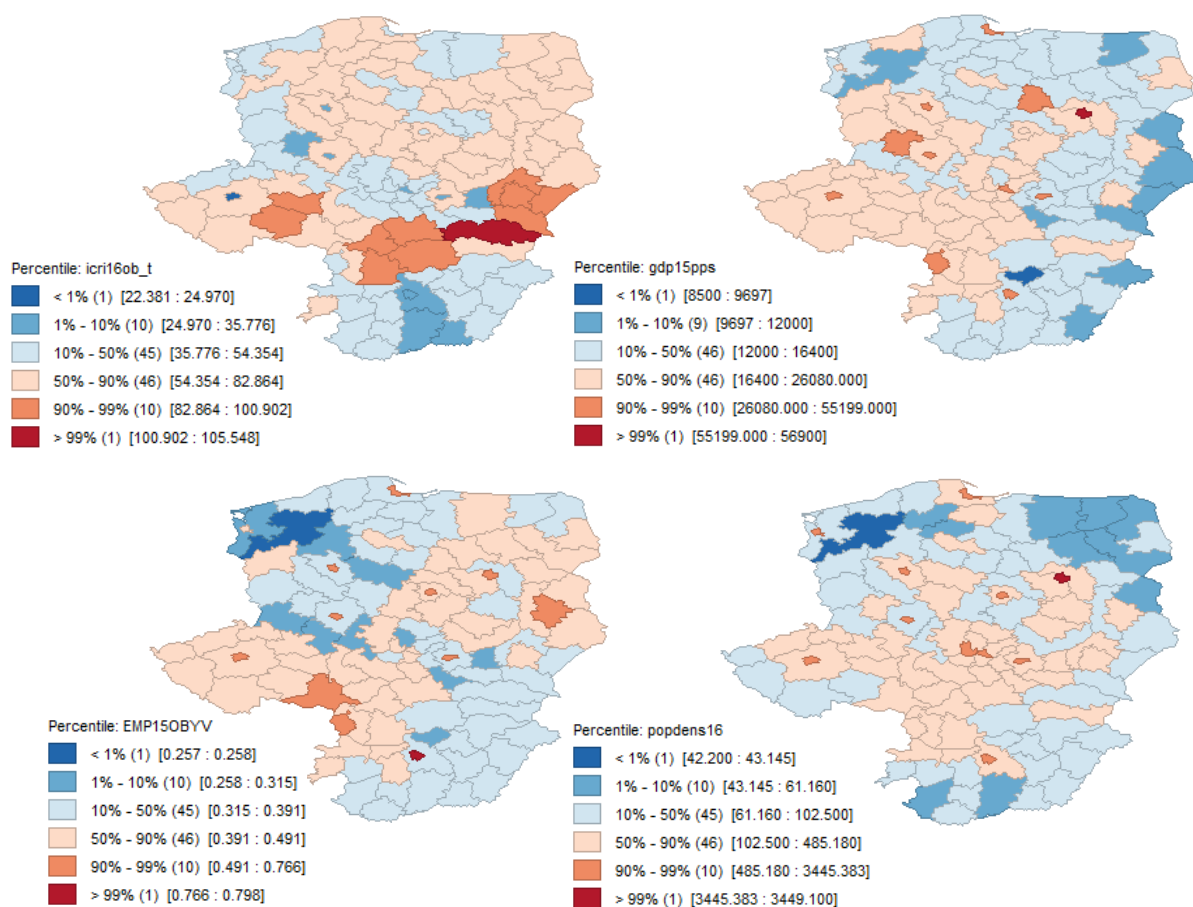
² Vzhľadom na údaje o počte registrovaných trestných činov z maďarského regiónu Zala, ktorý vykazoval v roku 2016 až 26 593 registrovaných trestných činov v porovnaní so 6 222 registrovanými trestnými činmi v roku 2015, sme sa rozhodli tento región vylúčiť z analýzy a zrealizovať analýzu len pre zostávajúcich 113 regiónov NUTS 3.

³ Hodnoty premennej boli získané vydelením údajov Eurostatu o priemernej ročnej populácii na výpočet údajov o regionálnom HDP podľa NUTS 3 regiónov v roku 2015 (Eurostat, 2018a) celkovým počtom trestných činov v jednotlivých regiónoch v roku 2016.

HDP - v roku 2015) a hustota obyvateľstva (počet obyvateľov na km²) v roku 2016. Priestorová analýza bola vykonaná v softvéri GeoDa, súbor .shp pre európske regióny v súvislosti s klasifikáciou NUTS 2013 bol stiahnutý z webovej stránky Eurostatu (2018b) a následne adekvátne upravený v softvéri GeoDa.

Obr. 1 obsahujúci percentilové mapy pre kriminalitu (*icri16ob_t*), HDP na obyvateľa (*gdp15pps*), mieru zamestnanosti (*EMP15OBYV*) a hustotu obyvateľstva (*popdens16*) v jednotlivých regiónoch krajín Vyšehradskej štvorky zachytáva nerovnomernosť rozdelenia analyzovaných premenných v priestore vrátane značných regionálnych rozdielov v rámci jednotlivých krajín.

Obr. 1: Percentilové mapy pre kriminalitu⁴ (*icri16ob_t*), HDP v PPS (*gdp15pps*), mieru zamestnanosti (*EMP15OBYV*) a hustotu obyvateľstva (*popdens16*)



Zdroj: vlastné spracovanie v GeoDa

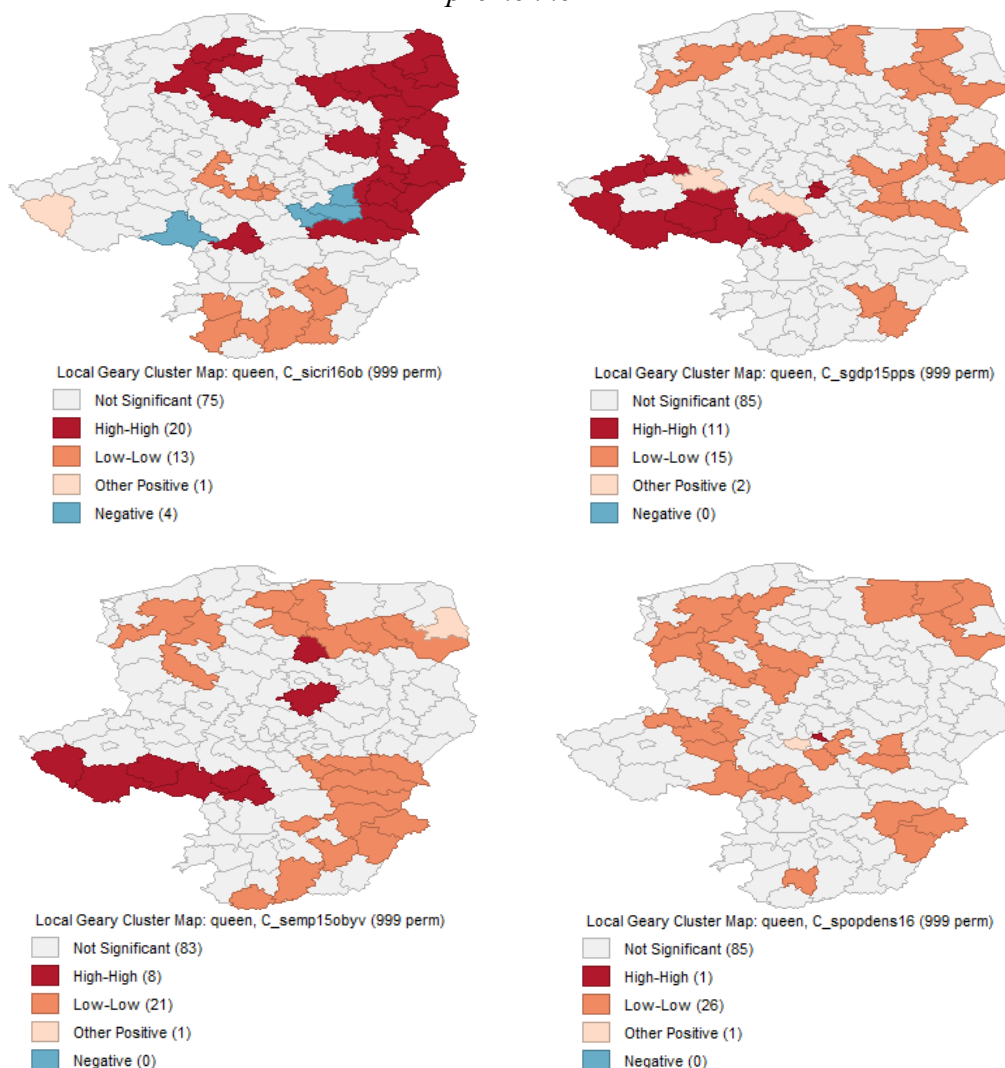
Pokiaľ ide o kriminalitu, najvyššiu mieru kriminality (t.j. najnižšiu hodnotu premennej *icri16ob_t*) sme zistili v Prahe, nasledovali 4 maďarské a 6 poľských regiónov koncentrovaných okolo veľkých miest. Najmenšia miera kriminality bola preukázaná v slovenskom regióne Prešovský kraj. Vysoké hodnoty zvyšných troch premenných - HDP na obyvateľa, miery zamestnanosti a hustoty obyvateľstva - boli zistené v regiónoch zodpovedajúcich hlavným mestám jednotlivých krajín ako aj v niektorých ďalších regiónoch koncentrovaných okolo väčších miest. Nízke hodnoty sme zaznamenali vo väčšine maďarských

⁴ Vzhľadom na to, že premenná kriminalita je v príspevku vyjadrená ako priemerný počet obyvateľov regiónu pripadajúci na 1 trestný čin v danom regióne, znamená to, že vysoké hodnoty tejto premennej označujú regióny s nízkou mierou kriminality, kým nízke hodnoty označujú regióny s vysokou mierou kriminality.

regiónov, v regiónoch na juhovýchode a východe Slovenska a tiež v poľských regiónoch lokalizovaných v rôznych častiach krajiny (predovšetkým na severozápade, severovýchode, ale aj v južnej časti).

V ďalšom kroku ESDA sme pristúpili k analýze priestorovej autokorelácie pomocou lokálnej Gearyho štatistiky. Obr. 2 znázorňuje klastrové mapy pre štandardizované premenné kriminalita (*sicri16ob*), HDP na obyvateľa (*sgdp15pps*), miera zamestnanosti (*semp15obyv*) a hustota obyvateľstva (*spopdens16*) zodpovedajúce hladine významnosti 0,05 s využitím 999 permutácií. Kým v prípade kriminality môžeme identifikovať tak regióny s pozitívnou ako aj negatívnou lokálnou priestorovou autokoreláciou, v prípade zvyšnej trojice uvažovaných premenných štatistická významnosť negatívnej lokálnej priestorovej autokorelácie nebola potvrdená pre žiadny región.

Obr. 2: Klastrové mapy na základe jednorozmerných Gearyho štatistik pre štandardizované premenné



Zdroj: vlastné spracovanie v GeoDa

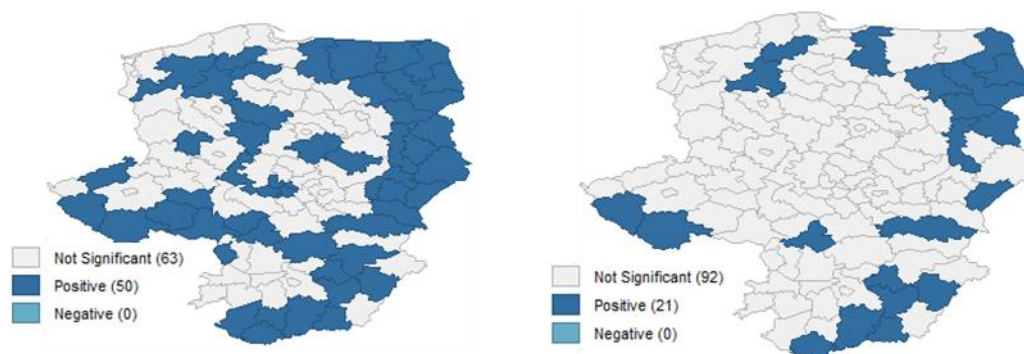
V prípade kriminality sú regióny so štatisticky významnou pozitívnou lokálnou priestorovou autokoreláciou typu vysoká-vysoká (indikujúce nízku mieru kriminality) lokalizované prevažne vo východnej časti Poľska, ale tiež v severozápadnej časti Poľska a do tejto kategórie patria tiež dva slovenské regióny. Regióny so štatisticky významnou pozitívnou

lokálnou priestorovou autokoreláciou typu nízka-nízka (indikujúce vysokú mieru kriminality) sú lokalizované v strednej a v južnej časti Maďarska a v južnej časti Poľska. Štatisticky významná pozitívna lokálna priestorová autokorelácia typu iná bola potvrdená pre Plzenský kraj v Českej republike. Štatisticky významná negatívna lokálna priestorová autokorelácia bola potvrdená pre trojicu poľských regiónov a jeden český región (Juhomoravský kraj).

Pre premenné HDP na obyvateľa a mieru zamestnanosti bol identifikovaný podobný počet regiónov so štatisticky významnou pozitívnou lokálnou priestorovou autokoreláciou typu vysoká-vysoká (11, resp. 8), pričom väčšina z nich bola lokalizovaná v južnej časti Českej republiky. V prípade poslednej uvažovanej premennej, hustota obyvateľstva, bola priestorová autokorelácia takéhoto typu potvrdená len pre jeden región v južnej časti Poľska. Regióny so štatisticky významnou pozitívnou lokálnou priestorovou autokoreláciou typu nízka-nízka boli v prípade tejto trojice premenných na mape značne viac rozptýlené. Taktiež bolo možné identifikovať pre jednotlivé premenné 1-2 regióny so štatisticky významnou pozitívnou lokálnou priestorovou autokoreláciou typu iná.

Po výpočte jednorozmerných lokálnych Gearyho štatistík sme pre premennú kriminalita a ostatné tri uvažované premenné zrealizovali viacrozmernú analýzu na báze viacrozmernej lokálnej Gearyho štatistiky (2). Viacrozmerná lokálna Gearyho štatistika je „kombináciou“ vzdialenosti vo viacrozmernej priestore určenom uvažovanými premennými (atribútmi) a geografickou vzdialenosťou. Je určitým indikátorom, ktorý udáva, či priemerná vzdialenosť medzi hodnotami uvažovaných premenných v regióne i a hodnotami týchto premenných v jeho susedných regiónoch je menšia alebo väčšia, ako by bola pri uvažovaní priestorovej náhodnosti. Kým prvá situácia zodpovedá pozitívnej priestorovej autokoreláci, druhá odráža existenciu negatívnej priestorovej autokorelácie (Anselin, 2019c). Obr. 3 poskytuje klastrové mapy pre štvorrozmernú Gearyho štatistiku pre hladinu významnosti 0,05 (vľavo) a pre FDR 0,00929 (vpravo) indikujúce existenciu 50, resp. 21 regiónov s pozitívnou priestorovou autokoreláciou. Znamená to, že tieto regióny majú podobný viacrozmerný profil ako ich susedné regióny. Je však dôležité si v tomto kontexte uvedomiť, že ak je niektorý región identifikovaný ako klastor so štatisticky významnou priestorovou autokoreláciou v prípade všetkých jednorozmerných Gearyho štatistík, nemusí to ešte automaticky znamenať, že je štatisticky významný aj vo viacrozmernej kontexte (Furková, 2019).

Obr. 3: Klastrové mapy na základe viacrozmernej Gearyho štatistiky (všetky štyri premenné) pre rôzne definované hladiny významnosti (0,05 a 0,00929)



Zdroj: vlastné spracovanie v GeoDa

4 Záver

Príspevok poskytuje priestorovú analýzu kriminality pre 113 NUTS 3 regiónov krajín Vyšehradskej štvorky. Okrem kriminality boli analyzované aj ďalšie relevantné premenné (HDP na obyvateľa, miera zamestnanosti a hustota obyvateľstva). Napriek niektorým čiastočne

limitujúcim faktorom (použitie dát na úrovni NUTS 3, a to vzhľadom na nedostupnosť, resp. neporovnateľnosť dát; použitie údajov za celkovú kriminalitu, čo bolo determinované nesúladosť vo vykazovaných druhoch kriminality v jednotlivých krajinách Vyšehradskej štvorky) bola zrealizovaná priestorová analýza s cieľom preskúmať previazanosť geograficky blízkych regiónov (priestorovú autokoreláciu) pre jednotlivé analyzované premenné na báze lokálnych jednorozmerných Gearyho štatistík, ako aj viacrozmerná analýza umožňujúca posúdiť profil, resp. podobnosť profilu jednotlivých regiónov na báze viacerých ukazovateľov s využitím v minulom roku publikovanej viacrozmernej Gearyho štatistiky.

Z pohľadu jednorozmernej Gearyho štatistiky pre kriminalitu bolo možné identifikovať celkovo 34 regiónov s podobnou úrovňou kriminality ako ich susedné regióny. V prípade vysokej miery kriminality ide o regióny lokalizované v strednej a južnej časti Maďarska (pri hraniciach s Chorvátskom, Srbskom a čiastočne aj Rumunskom) a v južnej časti Poľska; v prípade nízkej miery kriminality sú to takmer všetky regióny na východe Poľska (s výnimkou regiónu zahŕňajúceho mesto Lublin), štyri regióny na severozápade Poľska (mimo veľkých miest) a dva slovenské regióny (rozľahlý región Prešovský kraj a Trenčiansky kraj). Problematickejšia je interpretácia pre Plzenský kraj s tzv. iným typom pozitívnej lokálnej priestorovej autokorelácie, keďže podobnosť so susednými regiónmi nie je možné jednoznačne interpretovať. Odlišný profil od susedných regiónov bol potvrdený pre štvoricu regiónov (Juhomoravský kraj, Tarnowski, Nowosadecki a Nowotarski) vykazujúcich negatívnu lokálnu priestorovú autokoreláciu. Podobne boli identifikované skupiny regiónov s podobným profilom (pozitívnu autokoreláciu) a odlišným profilom (negatívnu autokoreláciu) aj pre zvyšné tri premenné (HDP na obyvateľa, mieru nezamestnanosti a hustotu obyvateľstva).

V ďalšom kroku bola analýza zameraná na posúdenie podobnosti, resp. rozdielnosti regiónov z viacrozmerného hľadiska berúc do úvahy všetky štyri uvažované premenné. Agregáciou individuálnych lokálnych Gearyho štatistík získaných pre každú uvažovanú premennú sme získali viacrozmernú Gearyho štatistiku, ktorá umožnila identifikáciu 50, resp. 21 regiónov s podobným viacrozmerným profilom (v dôsledku potvrdenia pozitívnej priestorovej autokorelácie) ako ich susedné regióny. Výsledky analýz na báze tak jednorozmerných ako aj viacrozmernej Gearyho štatistiky potvrdzujú teda nielen významný vplyv geografickej polohy regiónu, ale upriamujú tiež pozornosť na tzv. „zaujímavé“ regióny, ktoré majú podobný jednorozmerný, resp. viacrozmerný profil ako ich susedné regióny.

Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/0193/20 Vplyv priestorových spillover efektov na inovačné aktivity a rozvoj regiónov EÚ.

Literatúra

- [1] Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis*, 27(2), 93-115. doi:10.1111/j.1538-4632.1995.tb00338.x
- [2] Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G. (2000). Spatial Analyses of Crime. In *Criminal justice 2000: Vol. 4. Measurement and analysis of crime and justice*. (pp. 213-262). Washington, DC: National Institute of Justice.
- [3] Anselin, L. (2019a). A Local Indicator of Multivariate Spatial Association: Extending Geary's c. *Geographical Analysis*, 51(2), 133-150. doi: 10.1111/gean.12164
- [4] Anselin, L. (2019b). GeoDa An Introduction to Spatial Data Analysis, Local Spatial Autocorrelation (1). Univariate Local Statistics. Retrieved May 24, 2020, from https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html#local-geary.
- [5] Anselin, L. (2019c). GeoDa An Introduction to Spatial Data Analysis. Local Spatial Autocorrelation (2). Advanced Topics. Retrieved May 24, 2020, from https://geodacenter.github.io/workbook/6b_local_adv/lab6b.html#multivariate-local-geary.

- [6] Dray, S., & Jombart, T. (2011). Revisiting Guerry's data: Introducing spatial constraints in multivariate analysis. *The Annals of Applied Statistics*, 5(4), 2278-2299. doi: 10.1214/10-AOAS356
- [7] Eurostat. (2018a). Retrieved June 24, 2018, from <http://ec.europa.eu/eurostat/web/rural-development/data/database>.
- [8] Eurostat. (2018b). Retrieved June 24, 2018, from <http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units>.
- [9] Fischer, M.M., & Wang, J. (2011). *Spatial Data Analysis. Models, Methods and Techniques*. Heidelberg: Springer.
- [10] Friendly, M. (2007). A.-M. Guerry's Moral Statistics of France: Challenges for multivariable spatial analysis. *Statistical Science*, 22(3), 368-399. doi: 10.1214/07-STS241.
- [11] Furková, A. (2019). Spatial Cluster Analysis of the Regional Poverty in the European Union. In *Proceedings of the International Conference on Contemporary Issues in Economics, Management and Business*. Hanoi: National Economics University Publishing House, 873-883.
- [12] Geary, R. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5, 115-145. doi: 10.2307/2986645.
- [13] Hungarian Central Statistical Office. (2018). *Number of registered crimes by selected groups of crimes (2000–)*. Retrieved June 27, 2018, from https://www.ksh.hu/docs/eng/xstadat/xstadat_annual/i_zji001b.html.
- [14] Chocholatá, M. (2018a). Spatial econometric analysis of crime: Evidence from NUTS 3 regions of V4 countries. In *Proceedings of the 34th International Scientific Conference on Economic and Social Development – XVIII International Social Congress (ISC-2018)*. Moscow, Varazdin: Varazdin Development and Entrepreneurship Agency, 194-202.
- [15] Chocholatá, M. (2018b). Crime Determinants from the Spatial Perspective: a GWR Approach. In *Nové trendy v ekonometrii a operačným výzkumu, zborník príspevkov z medzinárodného vedeckého seminára*, Praha, Bratislava: Vydavateľstvo EKONÓM, 68-76.
- [16] Messner, S.F., Anselin L., Baller, R.D., Hawkins, D.F., Deane, G., & Tolnay, S.E. (1999). The spatial patterning of county homicide rates: an application of exploratory spatial data analysis. *Journal of Quantitative Criminology*, 15(4), 423-450. doi: 10.1023/A:1007544208712.
- [17] Policie České republiky. (2018). *Statistické přehledy kriminality za rok 2016*. Retrieved June 27, 2018, from <http://www.policie.cz/docDetail.aspx?docid=22346473&docType=ART>.
- [18] Ratcliffe, J. (2010). Crime Mapping: Spatial and Temporal Challenges. In: *Handbook of Quantitative Criminology* (pp. 5-24), New York: Springer-Verlag.
- [19] Statistics Poland. (2018). *Statistical Yearbook of the Regions - Poland 2017*. Retrieved June 27, 2018, from <https://stat.gov.pl/en/topics/statistical-yearbooks/statistical-yearbooks/statistical-yearbook-of-the-regions-poland-2017,4,12.html>.
- [20] Štatistický úrad Slovenskej republiky. (2018). DATAcube. Database. Retrieved June 27, 2018, from http://datacube.statistics.sk/#!/view/sk/VBD_SK_WIN/sk3003rr/Kriminalita%20pod%20C4%BEa%20z%20C3%A1kladn%C3%BDch%20skup%C3%ADn%20trestn%C3%BDch%20C4%8Dinov%205Bsk3003rr%5D.