

VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE
FAKULTA INFORMATIKY A STATISTIKY



BAKALÁRSKA PRÁCA

VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE
FAKULTA INFORMATIKY A STATISTIKY



**Empirická analýza faktorov volebného úspechu
krajnej pravice na Slovensku**

Autor: **Peter Osuský**
Katedra: **Katedra ekonometrie**
Odbor: **Matematické metódy v ekonómii**
Vedúci práce: **Ing. Jan Zouhar, Ph.D.**

Prehlásenie:

Prehlasujem, že som bakalársku prácu na tému „Empirická analýza faktorov volebného úspechu krajnej pravice na Slovensku“ spracoval samostatne. Všetku použitú literatúru a ďalšie podkladové materiály uvádzam v zozname použitej literatúry.

V Prahe dňa 24. mája 2017

.....

Peter Osuský

Pod'akovanie:

Rád by som na tomto mieste poďakoval Ing. Janovi Zouharovi, Ph.D. za vedenie mojej bakalárskej práce.

Abstrakt

Názov práce: Empirická analýza faktorov volebného úspechu krajnej pravice na Slovensku

Autor: Peter Osuský

Katedra: Katedra ekonometrie

Vedúci práce: Ing. Jan Zouhar, Ph.D.

Cieľom práce je nájsť faktory, ktoré boli katalyzátormi resp. inhibítormi nečakaného volebného úspechu slovenskej, krajne pravicovej strany Ľudová strana Naše Slovensko („ĽSNS“) a vytvoriť model, ktorý popisuje jej volebné zisky v okresoch Slovenska v ostatných parlamentných voľbách v roku 2016. Pre jeho dosiahnutie je v práci aplikovaná metóda regresnej a priestorovo-regresnej analýzy. Práca prichádza k záveru, že pomocou demografických, socioekonomických a v prípade priestorovej regresie aj geografických charakteristík jednotlivých okresov ako podiel vysokoškolsky vzdelanej populácie, podiel mužov, prítomnosť rómskej, či maďarskej menšiny atď. možno relatívne presne odhadnúť volebný zisk ĽSNS v nich.

Kľúčové slová: lineárna regresia, priestorová regresia, krajná pravica, Marian Kotleba, politický extrémizmus

Abstract

Title: The Electoral Success of Far-right in Slovakia: An Empirical Analysis

Author: Peter Osuský

Department: Department of Econometrics

Supervisor: Ing. Jan Zouhar, Ph.D.

The aim of the thesis is to find factors that were catalysts or inhibitors of the unexpected electoral success of the far-right Slovak party People's Party – Our Slovakia and to create a model that describes its electoral gains elections across districts in the latest 2016 Slovak parliamentary election. For this purpose, methods of linear regression and spatial regression analysis are applied. The thesis concludes that using specific demographic, socioeconomic and in the case of spatial regression as well geographic characteristics of individual electoral districts such as share of population with higher education, share of male population, substantial presence of Roma or Hungarian minority etc., the electoral gains of the party across them can be relatively accurately estimated.

Keywords: linear regression, spatial regression, far-right, Marian Kotleba, political extremism

Obsah

ÚVOD.....	1
1 HISTÓRIA EXTRÉMISTICKÝCH STRÁN A HNUTÍ NA SLOVENSKU.....	2
1.1 DO ROKU 2010.....	2
1.2 OD ROKU 2010 (VZNIK ĽUDOVEJ STRANY NAŠE SLOVENSKO)	3
2 DÁTOVÝ SÚBOR.....	5
2.1 PÔVOD DÁT.....	5
2.2 POPISNÁ ŠTATISTIKA	6
3 POUŽITÉ ŠTATISTICKÉ METÓDY.....	7
3.1 KORELÁCIA.....	7
3.2 LINEÁRNY REGRESNÝ MODEL	8
3.3 TESTY VÝZNAMNOSTI MODELU	10
3.4 RAMSEYHO RESET TEST.....	12
3.5 TESTY HETEROSKEDASTICITY	13
3.6 TESTY NORMALITY	15
3.7 PRIESTOROVÝ REGRESNÝ MODEL.....	17
4 TEORETICKÝ MODEL.....	20
5 APLIKÁCIA REGRESNEJ ANALÝZY A ZHODNOTENIE VÝSLEDKOV	23
5.1 VIACNÁSOBNÝ REGRESNÝ MODEL	23
5.2 DIAGNOSTIKA MODELU	25
5.3 PRIESTOROVÝ REGRESNÝ MODEL.....	26
5.4 ZHODNOTENIE VÝSLEDKOV.....	28
ZÁVER.....	30
POUŽITÉ ZDROJE A LITERATÚRA	31

Zoznam tabuliek a obrázkov

Tabuľka 1: Popisná štatistika dát	6
Tabuľka 2: Odhad viacnásobného regresného modelu MNŠ	24
Tabuľka 3: Korelačná matica regresorov	24

Tabuľka 4: Diagnostické testy	25
Tabuľka 5: Porovnanie odhadov modelov	27
Obrázok 1: Vizualizácia Moranovho I.....	19
Obrázok 2: Histogram rezíduí.....	25
Obrázok 3: Q-Q graf rezíduí	26
Obrázok 4: Priestorové mapy rezíduí.....	27

Úvod

V uplynulých rokoch zasiahlo Európu a Severnú Ameriku politické zemetrasenie v podobe vzostupu neštandardných, radikálnych či dokonca extrémistických strán a politikov. Medzi krajinami poznačenými týmto fenoménom sa ocitlo aj Slovensko, keď sa vo voľbách do Národnej Rady Slovenskej republiky (ďalej „NR SR“) v roku 2016 dostala do parlamentu aj krajne pravicová strana Ľudová strana Naše Slovensko (ďalej aj ako „ĽSNS“) a ktorej predseda Marian Kotleba sa ešte v roku 2013, po voľbách do vyšších územných celkov (VÚC), stal predsedom Banskobystrického samosprávneho kraja.

Cieľom tejto práce je identifikovať faktory, ktoré boli katalyzátormi resp. inhibítormi nečakaného volebného úspechu strany Mariana Kotlebu a vytvoriť ekonometrický model, ktorý dostatočne presne popisuje volebné zisky ĽSNS v okresoch Slovenska v ostatných voľbách do NR SR.

Prvá kapitola práce je venovaná stručnému prehľadu historického vývoja krajne pravicového extrémizmu na Slovensku.

Druhá kapitola detailne popisuje dátový súbor a odvodzuje metódy ekonometrickej analýzy naň aplikované.

Témou tretej kapitoly je formulácia teoretického modelu. Zaoberá sa výberom vysvetľujúcich premenných do modelu na základe sociologickej, ekonomickej a politologickej teórie a jednotlivými očakávaniami o numerických vzťahoch medzi závislými premennými a nezávislou premennou t. j. volebným ziskom ĽSNS v danom okrese.

Obsahom poslednej, štvrtej kapitoly je už samotná aplikácia regresnej analýzy na dátový súbor, zhodnotenie kvality modelu a výsledkov.

Záver prácu bilancuje z hľadiska naplnenia vytýčeného cieľu a zhrňuje výsledky a najdôležitejšie poznatky, ku ktorým autor dospel.

Pri zostavovaní dátového súboru a ekonometrického modelu a som použil programy *MS Excel*, *R* a *Stata 13*. Prvotné spracovanie dát bolo vykonané v prvej dvojici uvedených programov, ekonometrická analýza v treťom.

1 História extrémistických strán a hnutí na Slovensku

1.1 Do roku 2010

Slovenská krajne pravicová scéna začala naberať svoju dnešnú podobu ešte v osemdesiatych rokoch 20. storočia. Do krajiny vtedy začali prenikať extrémistické trendy predovšetkým z Nemecka a cez Českú republiku. (Mikušovič, 2009) Po Nežnej revolúcii dochádza ku spolupráni hnutia skinheads na Slovensku českými hudobnými skupinami ako Tři sestry či Orlík so svojim frontmanom Danielom Landom. Následne už začínajú vznikať aj domáce, hudobné skupiny (napr. Krátky Proces, Juden Mord, či Círhóza 88), ktoré scénu radikalizujú ešte viac. Ich texty sú ďaleko ostrejšie s otvorenými antisemitickými a rasistickými pasážami. (Milo, 2005)

V ďalšom vývoji vznikajú prvé politické organizácie. Časť orientovaná na nacizmus a ideológiu rasovej čistoty a iné vychádzajúce z klérofašistického vojnového Slovenského štátu s „aktualizačným“ panslavistickým podtónom, pre ktoré je typický skôr extrémny nacionalizmus, než násilná, neonacistická doktrína. (Biháiová, 2012)

Národno-socialistické hnutie Európy, označované ako prvé významnejšie zoskupenie, vzniklo v roku 1992. V jeho dvadsaťštyribodových stanovách prezentovalo myšlienku „oficiálneho zákazu degenerovania árijskej rasy“. Reálny význam hnutia však tkvel len v tom, že bolo prvou, registrovanou pravicovo-extrémistickou organizáciou na území Slovenska. Zaniklo už po troch rokoch existencie, v roku 1995. Jeho nasledovníci, ako napr. Klan Slovenských Rytierov (snažiaci sa nadviazať na odkaz amerického Ku Klux Klanu), Slovenský národný front, Biela slovenská jednota, Slovenská národná liga sa tiež vyznačovali relatívnu bezvýznamnosťou a ich aktivity sa obmedzovali takmer výlučne na maloobchod s oblečením, doplnkami (odznaky, nálepky, nášivky), cédečkami a kazetami krajne pravicových interpretov, či občasným vydávaním tlačovín. (Milo, 2005)

Za najvýznamnejšiu krajne pravicovú organizáciu, ktorá vznikla v 90. rokoch, možno označiť Slovenskú pospolitosť (ďalej „SP“), založenú bývalými prívržencami Slovenského národného frontu. Zaregistrovaná je 9. mája 1995 a od toho dňa funguje ako občianske

združenie so sídlom v Banskej Bystrici. Mikušovič rozdeľuje historický vývoj združenia do dvoch fáz (Mikušovič, 2007):

1. fáza nízkeho zapojenia verejnosti – „tichá“ (do r. 2003)
2. fáza aktívnej spoločenskej organizácie – „zvučná“ (od roku 2003, v roku 2005 založená Slovenská pospolitosť – Národná strana (ďalej „SPNS“) vyvrcholila policajnými zásahmi proti manifestáciám SP v štyroch slovenských mestách na jeseň roku 2005, rozpustením SPNS a fiaskom v parlamentných voľbách v roku 2006)

Najvýznamnejšou osobnosťou SP a jej politického krídla SPNS bol bezpochyby jej dlhoročný „vodca“ Marian Kotleba. Po tom čo bola najprv Najvyšším súdom Slovenskej Republiky rozpustená dňa 1. marca 2006 SPNS a následne Ministerstvom vnútra Slovenskej republiky rozpustené i jej občianske združenie SP, nastalo dočasné obdobie utlmenia aktivít krajne pravicového spektra, napriek tomu, že 1. júla 2009 Najvyšší súd rozhodnutie ministerstva vnútra zrušil a SP mohla ako občianske združenie pôsobiť naďalej. Okrem neúspešnej kandidatúry Mariana Kotlebu vo voľbách do VÚC v roku 2009 na post predsedu Banskobystrického samosprávneho kraja, keď ako nezávislý kandidát obsadil v prvom kole štvrté miesto s 10,03 % hlasov (Štatistický úrad Slovenskej republiky, 2009), nepostupujúc tak do kola druhého, prišiel významnejší návrat krajnej pravice na slovenskú politickú scénu až v roku 2010, vznikom Ľudovej strany Naše Slovensko.

1.2 Od roku 2010 (vznik Ľudovej strany Naše Slovensko)

Ľudová strana Naše Slovensko, dnes bezpochyby najvýznamnejšia strana na krajne pravicovej časti slovenského politického spektra, vznikla transformáciou pôvodne recesistickej Strany priateľov vína (ďalej „SPV“), zaregistrovanej ešte v roku 2000. SPV sa na krajne pravicovú zmenila v roku 2010, keď sa ju za podivných okolností podarilo ovládnuť ľuďom z okolia Mariana Kotlebu. Prvé verejné vystúpenie pod hlavičkou strany sa udialo v januári 2010 na proteste proti rómskej kriminalite v Leviciach.

ĽSNS sa pôvodne vyznačovala úzkou spoluprácou so SP, z ktorej vzišla väčšina jej členskej základne. Vzájomné vzťahy oboch vedení sa však v rokoch 2013 – 2014 začali pozvoľna zhoršovať. Hoci sa žiadna zo strán k rozkolu nechce vyjadrovať, rozdielne názorové smerovanie je viditeľné napríklad v odlišných pohľadoch na zahraničnú politiku, konkrétne na tzv. „Euromajdan“ a následný konflikt na východnej Ukrajine. Zatiaľ čo Marian Kotleba, ako predseda ĽSNS, podporoval zvrhnutého prezidenta Viktora Janukovyča, vedenie SP

držalo stranu Pravému sektoru, krajne pravicovej ukrajinskej politickej strane, ktorej militantné krídlo sa zúčastnilo ako Euromajdanu tak aj následných bojov na Donbase.

V rokoch 2010 a 2012 sa Kotlebova ĽSNS zúčastnila parlamentných volieb, v ktorých získala 1,33 % (33 724) hlasov (Štatistický úrad Slovenskej republiky, 2010), resp. 1,58 % (40 460) hlasov (Štatistický úrad Slovenskej republiky, 2012). Prvý a dovtedy najväčší úspech dosiahla vo voľbách do VÚC v roku 2013. Jej predseda, Marian Kotleba, sa v nich stal, šokujúc ako politológov, politikov tak aj verejnosť, predsedom Banskobystrického samosprávneho kraja a demonštroval tým politický potenciál strany na regionálnej úrovni. V prvom kole volieb získal 21,30 % (26 251) hlasov (Štatistický úrad Slovenskej republiky, 2013) a vytlačil na tretie, nepostupové miesto kandidáta pravice Ľudovíta Kaníka. Vtedajší premiér Robert Fico sa po prvom kole vyjadril v zmysle, že netreba šíriť zbytočnú paniku, pretože pokiaľ ide o druhé kolo „Aj vrece zemiakov porazí pána Kotlebu.“ V druhom kole však Marian Kotleba ziskom 55,53 % (71 397) hlasov (Štatistický úrad Slovenskej republiky, 2013) porazil Ficovho straníckeho kolegu zo SMER – sociálna demokracia Vladimíra Maňku a stal sa banskobystrickým županom.

V parlamentných voľbách v roku 2016 prišiel druhý a do dneška najvýznamnejší úspech strany. ĽSNS sa v nich umiestnila piatom mieste ziskom 8,04 % (209 779) hlasov (Štatistický úrad Slovenskej republiky, 2016), čím prvý krát prekročila päť percentné volebné kvórum pre vstup do parlamentu, v ktorom obsadila 14 poslaneckých kresiel. Samotný Marian Kotleba získal 156 355 prednostných hlasov.

Strana sa dnes profiluje predovšetkým na rómskej problematike, ochrane tradičnej rodiny, odpore voči skorumpovaným politikom a medzinárodným organizáciám ako EÚ a NATO. Snaží sa ideologicky, názorovo aj činmi hrať viac na populistickú než radikálnu strunu a prilákať tak aj umiernenějších potenciálnych voličov.

2 Dátový súbor

2.1 Pôvod dát

Dátový súbor určený pre empirickú analýzu tvorí 79 pozorovaní demografických, socioekonomických, geografických a volebných charakteristík 79 okresov Slovenskej republiky.

Demografické dáta pochádzajú zo *Sčítania obyvateľov, domov a bytov 2011*, až na výnimku podielov občanov rómskej národnosti v jednotlivých okresoch. Ako zdroj pre ne sme zvolili tzv. kvalifikovaný odhad z publikácie *Čierne-biele svety. Rómovia v majoritnej spoločnosti na Slovensku*, keďže sa pri periodických sčítaniach obyvateľov značná časť Rómov ku svojej národnosti neprihlasuje. Dôležitú rolu zohráva to, že Rómovia sú na Slovensku stále výrazne stigmatizovanou skupinou a mnohí z nich sa radšej oficiálne hlásia k inej národnosti, najčastejšie slovenskej a maďarskej. Ďalšou z príčin prečo sa Rómovia nehlásia k svojej identite má svoj pôvod ešte v období socializmu, kedy neboli uznaní ako oficiálna menšina. Až dodnes si mnohí z nich preto tradične uvádzajú slovenskú národnosť.

Socioekonomické dáta boli získané z ročenky Štatistického úradu Slovenskej republiky (ŠÚ SR) *Zamestnanci a priemerné mesačné mzdy 2015*.

Zdrojom pre volebné zisky ĽSNS je ŠÚ SR a Štátna komisia pre voľby a kontrolu financovania politických strán.

Zemepisné súradnice okresov a mapové podklady pre priestorovú regresiu pochádzajú z Geodetického a kartografického ústavu Bratislava.

2.2 Popisná štatistika

Popisné štatistiky dátového súboru sú uvedené v prehľadnej tabuľke nižšie. Pre každú premennú uvádzame priemer, medián, minimum, maximum a smerodajnú odchýlku:

Tabuľka 1: Popisná štatistika dát

názov premennej	popis premennej	priemer	medián	minimum	maximum	sm. odchýlka
bbkraj	binárna premenná pre BB kraj	0,165	0	0	1	0,373
lsnspct	volebný zisk ĽSNS v okrese	8,632	8,78	1,29	15,35	2,703
muzipct	podiel mužov v % v okrese	48,64	48,7	45,8	50,4	0,794
rkpct	podiel rímskych katolíkov v % v okrese	61,34	65,35	7,484	92,46	16,89
vspct	podiel VŠ vzdelaných v % v okrese	13,17	11,22	7,09	35,4	5,308
nezamestnanipct	miera nezamestnanosti v % okrese	10,49	8,89	4,27	26,91	5,082
romoviapct	podiel Rómov v % v okrese	7,577	4,924	0	31,7	7,753
madaripct	podiel Maďarov v % v okrese	6,355	0,17	0,02	73,21	13,93

3 Použité štatistické metódy

3.1 Korelácia

Korelačný koeficient

Základnou mierou podobnosti dvoch náhodných veličín, nazvime ich X a Y , vyjadrených v kardinálnej škále je Pearsonov párový korelačný koeficient:

$$\rho_{YX} = \frac{\text{cov}(Y, X)}{\sigma_Y \sigma_X}$$

V praxi (v prípade náhodného výberu premenných o rozsahu n) sa však používa jeho výberový náprotivok:

$$r_{yx} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Ten nadobúda hodnotu od -1 do 1, pričom hodnota 1 vyjadruje maximálnu podobnosť resp. pozitívnu lineárnu závislosť a hodnota -1 naopak maximálny nesúhlas, resp. negatívnu lineárnu závislosť medzi dvoma premennými.

Korelačná matica

Majme vektory pozorovaní $\mathbf{x}_1, \dots, \mathbf{x}_n$ náhodných veličín $X_1, \dots, X_n$. Výberovou **korelačnou maticou** veličín $X_1, \dots, X_n$ potom nazveme maticu:

$$R = \begin{pmatrix} 1 & r_{\mathbf{x}_1\mathbf{x}_2} & \dots & r_{\mathbf{x}_1\mathbf{x}_k} \\ r_{\mathbf{x}_2\mathbf{x}_1} & \ddots & & r_{\mathbf{x}_2\mathbf{x}_k} \\ \vdots & & \ddots & \vdots \\ r_{\mathbf{x}_k\mathbf{x}_1} & \dots & \dots & 1 \end{pmatrix},$$

kde $r_{\mathbf{x}_i\mathbf{x}_j}$ sú párové korelačné koeficienty príslušných vektorov náhodných veličín X_i a X_j .

3.2 Lineárny regresný model

Definícia

Uvažujme náhodný vektor premennej $\mathbf{y} = (y_1, \dots, y_n)^T$, predpokladajme vzťah:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

Kde $\mathbf{X}$ je matica pozorovaní premenných $x_1, \dots, x_k$ tvaru $n \times (k+1)$, pričom platí $n > k+1$. Prvý stĺpec matice $\mathbf{X}$ tvorí jednotkový vektor $\mathbf{x}_0$. $\boldsymbol{\beta}$ je vektor tzv. regresných koeficientov $\beta_0, \dots, \beta_k$ (kde β_0 je úrovňová konštanta) tvaru $(k+1) \times 1$ a $\mathbf{u}$ je vektor tvaru $n \times 1$ reprezentujúci idiosynkratickú chybu resp. náhodnú zložku. Takto zadefinovaný vzťah nazvime **lineárnym regresným modelom**. (Wooldridge, 2013)

Metóda najmenších štvorcov

Metóda najmenších štvorcov (skrátene „MNŠ“, ang. Ordinary least squares resp. skrátene „OLS“) odhaduje vektor regresných koeficientov $\boldsymbol{\beta}$ pre daný lineárny regresný model. Odhadnutý funkčný vzťah a ním vyrovnané hodnoty majú podobu:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$$

Zadefinujme $\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$. $\hat{\mathbf{u}}$ je **vektorom rezíduí** odhadnutého modelu $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$.

Skalár

$$\hat{\mathbf{u}}^T \hat{\mathbf{u}} = \hat{u}_1^2 + \hat{u}_2^2 + \dots + \hat{u}_n^2$$

sa potom nazýva **reziduálnym súčtom štvorcov** (RSS, v ang. tiež SSR z Sum of Squared Residuals).

Metóda najmenších štvorcov poskytuje práve také odhadnuté hodnoty $\hat{\beta}_0, \dots, \hat{\beta}_k$, ktoré reziduálny súčet štvorcov minimalizujú. Dosadením $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ do vzťahu pre SSR dostaneme:

$$SSR = \hat{\mathbf{u}}^T \hat{\mathbf{u}} = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

Upravujme ďalej:

$$\begin{aligned}
 SSR &= (\mathbf{y}^T - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T)(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\
 SSR &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \\
 SSR &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \\
 SSR &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} \\
 SSR &= \mathbf{y}^T \mathbf{y} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y}
 \end{aligned}$$

Podmienky prvého rádu pre minimalizáciu SSR môžeme zapísať ako:

$$\frac{\partial SSR}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} = 0$$

Po úprave získavame vzťah pre $\hat{\boldsymbol{\beta}}$:

$$\begin{aligned}
 -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} &= 0 \\
 \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}^T \mathbf{y} \\
 \hat{\boldsymbol{\beta}} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}
 \end{aligned}$$

Gauss–Markovove predpoklady

Uvažujme nasledovné predpoklady:

1. Správna špecifikácia modelu

Premenné v populácii popisuje vzťah lineárneho regresného modelu $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ pre nejaký konštantný vektor $\boldsymbol{\beta} \in \mathbb{R}^{k+1}$

2. Náhodný výber

Výber pozorovaní je z populácie nadobudnutý náhodným výberom. Hodnoty premenných rôznych pozorovaní sú teda na sebe nezávislé.

3. Exogenita

Stredná hodnota náhodnej zložky je nezávislá na hodnotách vysvetľujúcich premenných. Možno zapísať ako: $E(\mathbf{u} | \mathbf{X}) = \mathbf{0}$.

4. Absencia perfektnej multikolinearity

Neexistencia lineárnej závislosti medzi premennými (stĺpcami matice $\mathbf{X}$). $\mathbf{X}$ má teda plnú hodnotu.

5. Homoskedasticita

Rozptyl náhodnej zložky je konečný a konštantný. Nemení sa teda s hodnotami vysvetľujúcich premenných. Možno zapísať ako: $\text{var}(\mathbf{u} | \mathbf{X}) = \sigma^2 \mathbf{I}$, kde $\mathbf{I}$ je jednotkový vektor a $\sigma^2 \in (0; \infty)$

6. Normalita náhodnej zložky

Náhodná zložka má pri daných hodnotách premennej $\mathbf{x}$ normálne rozdelenie s nulovou strednou hodnotou a konečným rozptylom. Možno zapísať ako $\mathbf{u} | \mathbf{X} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$

Pokiaľ sú predpoklady 1. – 5. splnené, majú odhady $\hat{\beta}$ a poľažne aj $\hat{y}$, získané pomocou MNS, niektoré priaznivé vlastnosti. Sú nestranné, výdatné, konzistentné a majú najmenší rozptyl v množine daných lineárnych odhadových funkcií (ang. Best Linear Unbiased Estimator resp. skrátené „BLUE“). Predpoklad 6. pridávame pre zachovanie normality odhadov $\hat{\beta}$ v prípade malých výberových vzoriek. (Wooldridge, 2013)

3.3 Testy významnosti modelu

Studentov t-test

Uvažujme Gauss–Markovove predpoklady 1 – 6. za splnené. Potom majú odhady regresných parametrov $\hat{\beta}_j$ (kde $j = 0, \dots, k$) normálne rozdelenie. Platí teda, $\hat{\beta}_j \sim N(\beta_j, \text{var}(\hat{\beta}_j))$, kde sú rozptyly dané:

$$\text{var}(\hat{\beta}_0) = \sigma^2 v_{11}, \text{var}(\hat{\beta}_1) = \sigma^2 v_{22}, \dots, \text{var}(\hat{\beta}_k) = \sigma^2 v_{(k+1)(k+1)},$$

pričom $v_{11}, v_{22}, \dots, v_{(k+1)(k+1)}$ sú prvky na hlavnej diagonále matice $(\mathbf{X}^T \mathbf{X})^{-1}$. Nestranný odhad

rozptylu náhodnej zložky σ^2 získame ako $\hat{\sigma}^2 = \frac{SSR}{n - k - 1}$. Odhady $\text{var}(\hat{\beta}_j)$ sú potom dané:

$$\text{var}(\hat{\beta}_0) = \hat{\sigma}^2 v_{11}, \text{var}(\hat{\beta}_1) = \hat{\sigma}^2 v_{22}, \dots, \text{var}(\hat{\beta}_k) = \hat{\sigma}^2 v_{(k+1)(k+1)}$$

Druhé odmocniny týchto odhadov

$$s(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 v_{(j+1)(j+1)}}$$

sa nazývajú **smerodajné chyby odhadov regresných parametrov**.

Testy významnosti parametrov $\hat{\beta}_j$ sú založené na štatistikách

$$t = \frac{\hat{\beta}_j - \beta_j}{s(\hat{\beta}_j)},$$

ktoré majú Studentovo rozdelenie s $(n - k - 1)$ stupňami voľnosti. Testujeme nulovú hypotézu nulovosti regresného parametru $H_0: \beta_j = 0$, proti alternatívnej hypotéze $H_1: \beta_j \neq 0$.

V prípade, že platí nulová hypotéza má štatistika

$$t = \frac{\hat{\beta}_j - \beta_j}{s(\hat{\beta}_j)} = \frac{\hat{\beta}_j - 0}{s(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}$$

Studentovo rozdelenie $(n - k - 1)$ stupňami voľnosti (kde k je počet regresorov v modeli). Kritickou hodnotou zodpovedajúcou danej hladine významnosti α je potom kvantil $t_{1-\frac{\alpha}{2}(n-k-1)}$.

(Davidson, a iní, 2003)

Koeficient determinácie

Platí $y_i - \bar{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})$, z čoho možno odvodiť platnosť:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \rightarrow SST = SSE + SSR,$$

Kde je:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \mathbf{y}^T \mathbf{y} - n\bar{y}^2 = SST \text{ tzv. celkový súčet štvorcov (z ang. Sum of Squared Total)}$$

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \mathbf{y}^T \mathbf{y} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} = SSR \text{ reziduálny súčet štvorcov}$$

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} - n\bar{y}^2 = SSE \text{ vysvetlený súčet štvorcov (z ang. Sum of Squared}$$

Explained)

Podiel $\frac{SSE}{SST}$ resp. jemu ekvivalentný výraz $1 - \frac{SSR}{SST}$ nazývame **koeficientom determinácie**

a označujeme ho **R²** (tiež ang. R-squared). Koeficient determinácie nadobúda hodnoty z intervalu $\langle 0; 1 \rangle$ a udáva, aký podiel rozptylu závislej premennej sa podarilo regresnému

modelu vysvetliť. Teda čím bližšie je hodnota R^2 k 1, tým presnejšie model popisuje skutočnosť.

R^2 je však o. i. závislý na počte parametrov regresného modelu. T. j. pridávanie akýkoľvek, hoc' i nezmyselných regresorov, bude mať za následok približovanie sa R^2 k 1. Preto sa používa jeho korekcia $R^2_{adj} = 1 - (1 - R^2) \frac{n-1}{n-k-1}$, ktorú nazývame **upraveným koeficientom determinácie** (ang. Adjusted R-squared). (Wooldridge, 2013)

F-test významnosti modelu

Pri overovaní významnosti regresného modelu (s konštantou β_0) sa testuje nulová hypotéza:

$$H_0 : (\beta_1, \dots, \beta_k) = (0, \dots, 0), \text{ proti alternatívnej}$$

$$H_1 : (\beta_1, \dots, \beta_k) \neq (0, \dots, 0).$$

Platnosť H_0 udáva, že model je nevhodný a postačuje vyjadrenie s konštantou β_0 $y = x_0\beta_0 + u$, kde x_0 je jednotkový stĺpcový vektor.

Testové kritérium je štatistika:

$$F = \frac{R^2 / (k-1)}{(1-R^2) / (n-k)} \sim F(k-1, n-k),$$

H_0 zamietame, keď na hladine významnosti α platí $F \geq F_{1-\alpha}(k-1, n-k)$. (Wooldridge, 2013)

3.4 Ramseyho RESET test

Ramseyho RESET (z ang. Ramsey Regression Equation Specification Error Test) test (Ramsey, 1969) overuje správnosť funkčnej špecifikácie modelu.

Uvažujme lineárny regresný model v tvare

$$y = X\beta + u$$

a jeho odhadnutý náprotivok

$$\hat{y} = X\hat{\beta} \quad (*)$$

Ramseyho RESET test testuje, či má zavedenie ďalších vysvetľujúcich premenných, v podobe mocnín vyrovnaných hodnôt z pôvodne odhadnutej regresie (*)

$$\hat{\mathbf{y}}^{\circ 2} = (\mathbf{X}\hat{\boldsymbol{\beta}})^{\circ 2}, \dots, \hat{\mathbf{y}}^{\circ q} = (\mathbf{X}\hat{\boldsymbol{\beta}})^{\circ q} \quad (\text{kde operátor „}^{\circ q}\text{“ predstavuje Hadamardovu mocninu})$$

do modelu, štatisticky významný dopad na jeho kvalitu meraní koeficientom determinácie. Novovytvorený model a jeho odhad majú teda podobu:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{y}}^{\circ 2}\theta_1 + \dots + \hat{\mathbf{y}}^{\circ q+1}\theta_q + \mathbf{u}$$

$$\hat{\mathbf{y}}_n = \mathbf{X}\hat{\boldsymbol{\beta}}_n + \hat{\mathbf{y}}^{\circ 2}\hat{\theta}_1 + \dots + \hat{\mathbf{y}}^{\circ q+1}\hat{\theta}_q$$

Testové kritérium je potom štatistika:

$$F = \frac{(R_n^2 - R_o^2)/q}{(1 - R_n^2)/(n - k - q - 1)} \sim F(q, n - k - q - 1)$$

Kde R_n^2 a R_o^2 predstavujú koeficienty determinácie novovytvoreného a pôvodného odhadnutého modelu, k počet regresorov v pôvodnom modeli a q počet pridaných Hadamardových mocnín $\hat{\mathbf{y}}$, ako regresorov v novovytvorenom modeli.

Nulovú hypotézu H_0 , predpokladajúcu správnu špecifikáciu modelu, na hladine významnosti α zamietame, keď platí $F \geq F_{1-\alpha}(q, n - k - q - 1)$.

3.5 Testy heteroskedasticity

Breusch–Paganov test

Do diagnostických testov zameraných na testovanie prítomnosti heteroskedasticity v lineárnom modeli patrí Breusch–Paganov test (Breusch, a iní, 1979). Uvažuje regresný model v štandardnom tvare $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ a predpokladá, že prípadná heteroskedasticita náhodnej zložky má tvar:

$$\text{var}(\mathbf{u} | \mathbf{Z}) = \mathbf{Z}\boldsymbol{\delta}$$

Kde $\mathbf{Z} = (\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k)$ je matica zložená zo stĺpcových vektorov pozorovaní nejakých veličín $z_1, \dots, z_k$ a jednotkového vektoru $\mathbf{z}_0$. $\boldsymbol{\delta} = (\delta_0, \dots, \delta_k)^T$ je vektor im prislúchajúcim regresným koeficientom. Za $\mathbf{Z}$ sa často volia vysvetľujúce premenné z prvotného regresného modelu. Vzťah potom nadobúda formu:

$$\text{var}(\mathbf{u} | \mathbf{X}) = \mathbf{X}\boldsymbol{\delta}$$

Keďže je populačný rozptyl náhodnej zložky $\text{var}(\mathbf{u} | \mathbf{X})$ v praxi neznámy, nahradzuje sa jeho odhadom $\text{var}(\hat{\mathbf{u}}) = \hat{\mathbf{u}}^{\circ 2}$, kde $\hat{\mathbf{u}}$ je vektorom rezíduí z regresie $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}}$. Výsledný pomocný model a jeho OLS odhad vyzerajú nasledovne:

$$\hat{\mathbf{u}}^{\circ 2} = \mathbf{X}\boldsymbol{\delta}$$

$$\hat{\mathbf{u}}^{\circ 2} = \mathbf{X}\hat{\boldsymbol{\delta}} + \hat{\mathbf{v}}$$

Testuje sa nulová hypotéza H_0 konštantného rozptylu náhodnej zložky (homoskedasticity), ekvivalentná situácii, keď sú jednotlivé regresné koeficienty $\boldsymbol{\delta}$ nulové, proti alternatívnej hypotéze H_1 , predpokladajúcej prítomnosť heteroskedasticity:

$$H_0 : (\delta_1, \dots, \delta_k) = (0, \dots, 0)$$

$$H_1 : (\delta_1, \dots, \delta_k) \neq (0, \dots, 0)$$

Na hladine významnosti α je príslušný kritický obor nulovej hypotézy daný testovacou štatistikou nR^2 s χ^2 rozdelením:

$$nR^2 \geq \chi^2_{1-\alpha}(k-1)$$

Kde R^2 značí koeficient determinácie z odhadnutej regresie $\hat{\mathbf{u}}^{\circ 2} = \mathbf{X}\hat{\boldsymbol{\delta}} + \hat{\mathbf{v}}$.

Whiteov test

Ďalším z testov na prítomnosť heteroskedasticity v modeli je Whiteov test (White, 1980). Na rozdiel od verzie Breuschovho–Paganovho testu z predchádzajúcej kapitoly, pridáva jednoduchá forma Whiteovho testu do pomocného modelu vysvetľujúceho rozptyl náhodnej zložky k regresorom $\mathbf{x}_1, \dots, \mathbf{x}_k$ tvoriacim maticu $\mathbf{X}$ aj ich druhé mocniny a krížové súčiny:

$$\hat{\mathbf{u}}^{\circ 2} = \begin{pmatrix} 1 & \mathbf{x}_1 & \dots & \mathbf{x}_k & \mathbf{x}_1^{\circ 2} & \dots & \mathbf{x}_k^{\circ 2} & \mathbf{x}_1\mathbf{x}_2 & \dots & \mathbf{x}_{k-1}\mathbf{x}_k \end{pmatrix} \boldsymbol{\delta}$$

Nevýhodou Whiteovho testu je vzhľadom k veľkému počtu regresorov to, že pri nízkom počte pozorovaní resp. stupňov voľnosti, jeho sila slabne. Preto sa často namiesto druhých mocnín a krížových súčinov všetkých vysvetľujúcich premenných používa ich odhad v podobe druhej mocniny vyrovnaných hodnôt $\hat{\mathbf{y}}$:

$$\hat{\mathbf{u}}^{\circ 2} = \begin{pmatrix} \mathbf{1} & \mathbf{x}_1 & \dots & \mathbf{x}_k & \hat{\mathbf{y}}^{\circ 2} \end{pmatrix} \boldsymbol{\delta}$$

Nulová hypotéza opäť predpokladá situáciu, keď sú regresné koeficienty $\boldsymbol{\delta}$ nulové:

$$H_0 : (\delta_1, \dots, \delta_p) = (0, \dots, 0)$$

$$H_1 : (\delta_1, \dots, \delta_p) \neq (0, \dots, 0)$$

Testovacia štatistika a kritický obor sú obdobné ako v prípade Breuschovho–Paganovho testu:

$$nR^2 \geq \chi^2_{1-\alpha}(p)$$

3.6 Testy normality

Kvantil–kvantilový graf (Q–Q graf)

Prostredníctvom tzv. kvantil–kvantilového grafu resp. Q–Q (z ang. Quantile–quantile plot) grafu možno posúdiť, či výberové dáta pochádzajú z nejakého známeho rozdelenia. V prípade testovania normality volíme ako referenčné nejaké normálne rozdelenie s konečným rozptylom a nulovou strednou hodnotou. Graf porovnáva výberové kvantily s teoretickými kvantilmi zvoleného rozdelenia.

Majme usporiadané hodnoty výberu $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. Potom je i -tá usporiadaná hodnota považovaná za $p_i \cdot 100\%$ kvantil, kde $p_i = \frac{i - i_{adj}}{n + n_{adj}}$, pričom i_{adj} a $n_{adj} = 1 - 2i_{adj}$ sú tzv. korigujúce faktory a i_{adj} sa najčastejšie volí v intervale $\langle 0; 1/2 \rangle$. Ďalej označme $p \cdot 100\%$ kvantil zvoleného rozdelenia ako $Q(p)$, pričom platí $Q(p) = \Phi^{-1}(x)$, kde $\Phi(x)$ je distribučná funkcia zvoleného rozdelenia. Q–Q graf je potom bodový graf $(Q(p_i), x_{(i)})$. Týmto bodmi sa potom pomocou metódy najmenších štvorcov preloží priamka. Čím menej sa body od tejto priamky opticky odchyľujú, tým je vyšší súlad medzi rozdelením výberu a zvoleným teoretickým rozdelením. (Thode, 2002)

Shapiro–Wilkov test

Shapiro–Wilkov test (Shapiro, a iní, 1965) sa používa predovšetkým pre výbery o malých rozsahoch $n \leq 50$. Test je založený na analýze Q–Q grafu, teda či sa jeho body výraznejšie odchyľujú od regresnej priamky nimi preloženej. Na hladine významnosti α

testuje nulovú hypotézu, že výber $\mathbf{x} = (x_1, \dots, x_n)$ pochádza z normálneho rozdelenia, teda že $\mathbf{x} \sim N(\mu, \sigma^2)$. Testovacia štatistika má tvar:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})}$$

Kde $x_{(i)}$ sú tzv. poradové štatistiky a a_i koeficienty odvodené zo stredných hodnôt a kovariančnej matice poradových štatistík náhodného výberu o rozsahu n z normovaného normálneho rozdelenia $N(0,1)$. Štatistika W bude nadobúda hodnoty od 0 do 1, pričom čím je jej hodnota vyššia, tým vyššiu zhodu s normálnym rozdelením výberové dáta vykazujú. Presnejšie, v prípade, že bude testovacia štatistika nižšia, než príslušná kritická hodnota $W_{n,\alpha}$, nulová hypotéza o normalite výberu sa na hladine významnosti α zamietajú.

Jarqueov–Berov test

Jarqueov–Berov test (Jarque, a iní, 1980) je test založený na porovnávaní rozdielov koeficientov šikmosti a špicatosti výberových dát s koeficientami šikmosti a špicatosti normálneho rozdelenia. Na zvolenej hladine významnosti α testuje nulovú hypotézu:

$$H_0 : F(x) = \Phi(x), \text{ proti alternatívnej}$$

$$H_1 : F(x) \neq \Phi(x),$$

Kde F je empirická distribučná funkcia výberových dát a $\Phi(x)$ je teoretická distribučná funkcia uvažovaného normálneho rozdelenia $N(\mu, \sigma^2)$ s neznámymi parametrami.

Testové kritérium je štatistika:

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4} (K - 3)^2 \right) \sim \chi^2_{1-\alpha}(2),$$

kde S a K sú postupne koeficienty šikmosti (z ang. skewness) a špicatosti (z ang. kurtosis) výberového súboru. Hypotézu H_0 zamietame, na hladine významnosti α , keď $JB \geq \chi^2_{1-\alpha}(2)$.

Problémom tohto testu je, že platí len asymptoticky, predpokladá teda veľký rozsah výberového súboru. Pri nižších rozsahoch $n < 100$, je potrebná ostražitosť, keďže je vtedy test príliš citlivý, často zamietajúci nulovú hypotézu napriek jej pravdivosti (chyba prvého druhu).

3.7 Priestorový regresný model

Cieľom priestorovej regresie je podobne ako pri klasickej regresii, čo najpresnejšie kvantifikovanie vzťahu medzi závislou premennou y a jej prediktormi $\mathbf{X}$, avšak navyše berúc do úvahy priestorovú závislosť (autokoreláciu) medzi pozorovaniami závislej premennej. Prítomnosť priestorovej autokorelácie možno všeobecne zadať ako podobnosť medzi hodnotami pozorovaní premennej a miestami, kde boli dané hodnoty namerané resp. získané. Formálnejší zápis vyzerá nasledovne:

$$\text{Cov}(y_i, y_j) = E(y_i y_j) - E(y_i)E(y_j) \neq 0 \text{ pre } i \neq j,$$

kde y_i a y_j sú pozorovania náhodnej premennej v lokalitách i a j , pričom i a j môžu byť reprezentované ako bodmi (napr. zemepisné súradnice), tak aj oblasťami (napr. okresy či kraje). (Anselin, a iní, 2003)

Dvomi najčastejšie používanými typmi modelov sú **priestorový intervalový model** (ang. Spatial Lag Model) a **priestorový chybový model** (ang. Spatial Error Model). Oba modely sú odhadované metódou maximálnej vierohodnosti (ang. Maximum Likelihood Estimation), keďže priamo z ich definícií plynie porušenie Gauss–Markovových predpokladov. V práci bude použitý prvý menovaný.

Matica priestorových váh

To, ako veľmi sú jednotlivé lokality pozorovaní vzájomne priestorovo príbuzné a či má zmysel overovať závislosť medzi k nim prislúchajúcimi hodnotami premennej je formulované pomocou tzv. **matice priestorových váh**. Označuje sa $\mathbf{W}$, je tvaru $n \times n$ (n je počet pozorovaní) a jej prvky w_{ij} sú nenulové v prípade, že sú lokality pozorovaní i a j považované „susedné“. „Susednosť“ dvoch lokalít môže byť zadefinovaná rôznymi spôsobmi. Pri lokalitách reprezentovaných bodmi najčastejšie prostredníctvom ich vzdialenosti, pri oblastiach zdieľaním spoločnej hranice, vzdialenosťou ich geometrických ťažísk (tzv. centroidy), alebo kombináciou oboch. Zaužívanou konvenciou je to, že diagonálne prvky matice $\mathbf{W}$ sú nulové (teda, že jednotlivé lokality „nesusedia sami so sebou“) a matica býva často štandardizovaná po riadkoch. (Anselin, a iní, 2003)

Priestorový intervalový model

Priestorový intervalový model inkorporuje priestorovú závislosť do klasického regresného vzťahu priamo, pridaním ďalšej nezávislej premennej $\mathbf{W}y$:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}y + \mathbf{u},$$

kde je ρ priestorový regresný parameter a $\mathbf{W}$ matica priestorových váh. Model teda predpokladá, že hodnota závislej premennej v nejakej lokalite je ovplyvnená jej hodnotami v susedných lokalitách. (Anselin, a iní, 2003)

Moranovo I

Jednou z najpoužívanějších mier výskytu priestorovej závislosti v priestorovo usporiadateľných premenných je tzv. Moranovo I. Je definované nasledovne (Moran, 1950):

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

kde w_{ij} je prvok príslušnej matice priestorových váh $\mathbf{W}$ a $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ je potom celkový súčet váh. Moranovo I nadobúda hodnôt od -1 do 1 , pričom hodnota 1 vyjadruje perfektnú pozitívnu (t. j. zhukovanie podobných hodnôt pozorovaní premennej v priestore) a hodnota -1 perfektnú negatívnu (t. j. vysoké hodnoty premennej sú obklopené prevažne tými nízkymi, alebo naopak) priestorovú závislosť (viď Obrázok 1). Nulová hodnota naznačuje, že sú hodnoty premennej v priestore usporiadané náhodne, na lokalite teda nezávisia.

Testovacou štatistikou Moranovho I je z-skóre:

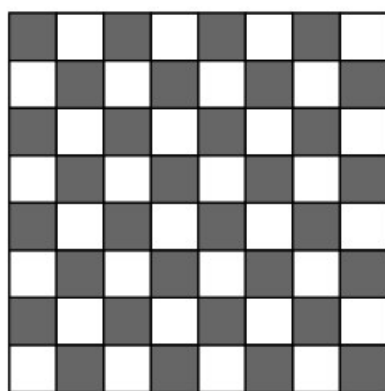
$$z_I = \frac{I - E(I)}{\sqrt{\text{var}(I)}},$$

kde $E(I) = \frac{-1}{(n-1)}$ a $\text{var}(I) = E(I^2) - E(I)^2$. Testuje nulovú hypotézu:

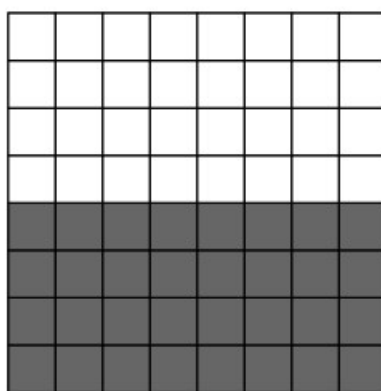
$$H_0 : I = 0, \text{ proti alternatívnej}$$

$$H_1 : I \neq 0$$

H_0 o neprítomnosti priestorovej autokorelácie na zvolenej hladine α zamietame, keď platí $|z_I| > z_{1-\alpha}$.



$I = -1$



$I = 1$

Obrázok 1: Vizualizácia Moranovho I

4 Teoretický model

Identifikovanie charakteristík skupín obyvateľstva, ktoré stoja za, alebo naopak, brzdia vzostup krajnej pravice, môže byť ťažké a to napriek rozsiahlej akademickej debate o umiestnení jednotlivých strán a hnutí na politickom spektre a analýze ich ideologického profilu. Jednou z príčin je to, že voliči voliaci extrémistické strany ako LSNS, tak často robia z úplne iných, ako ideologických dôvodov. Táto kapitola je preto venovaná hľadaniu socioekonomických, demografických a iných faktorov, ktoré môžu zdôvodniť voličskú preferenciu Ľudovej Strany Naše Slovensko a tiež antipatiu k nej. Východiskom budú sociologické a politologické práce a niektoré špecifiká Slovenska, ako volebného teritória.

Pohlavie ako faktor

Jedným z najzreteľnejších znakov, ktorým možno charakterizovať elektorát krajnej pravice je ten, že sú v ňom disproporčne viac zastúpení muži. Výskumy venujúce sa tejto problematike prítomnosť tzv. „gender gap“ potvrdzujú, a to aj po zohľadnení socioekonomických činiteľov. (Givens, 2004)

S vysvetlením tohto fenoménu, prichádza Hartevelde a kolektív. Podľa nich sú muži náchylnejší prisudzovať témam, ktoré krajne pravicové strany prioritizujú (právo a poriadok, odpor voči imigrácií atď.) vyššiu dôležitosť. (Hartevelde, a iní, 2015)

Vplyv vzdelania

Extrémistické strany mávajú vyššiu mieru podpory u menej vzdelaných skupín obyvateľstva a naopak, vysokoškolsky vzdelaní ľudia majú v radoch ich voličov nižšie zastúpenie. Možným zdôvodnením je to, že voliči s vyššími stupňami vzdelania častejšie vyznávajú liberálne hodnoty (Weakliem, 2002), ktoré sú v priakom rozpore s krajne pravicovými postojmi. Taktiež sa možno domnievať, že rétorika krajnej pravice namierená proti „elitám“ rezonuje ďaleko viacej u ľudí s nižším vzdelaním, ktorí sa ako súčasť „elity“ cítia v ďaleko menšej miere, ako vzdelanejšie vrstvy.

Nezamestnanosť

Ďalšími z faktorov, ktoré pravdepodobne prispievajú k rastu volebnej podpory krajnej pravice je nepochybne vysoká miera nezamestnanosti. Tú dávajú extrémistickí politici za vinu

často imigrantom a cudzincom, ktorí ich slovami „kradnú“ pracovné miesta domácomu obyvateľstvu. (Jackman, a iní, 1996) A podobne, ako v prípade menej vzdelaných skupín obyvateľstva, aj nezamestnaní ľudia na „antielitársku“ rétoriku krajnej pravice reflektujú, vzhľadom k svojmu nižšiemu sociálnemu postaveniu, častejšie.

Katolicizmus, Tiso a ĽSNS

Nezanedbateľná časť slovenských, katolíckych veriacich vníma vojnový Slovenský štát (resp. prvú Slovenskú republiku, obdobie 1939 – 1945) a jej vrcholného predstaviteľa Jozefa Tisu za výnimočne pozitívne obdobie a výnimočnú postavu slovenských dejín. A keďže je kresťanský prezentujúca sa ĽSNS jedinou politicky relevantnou silou, ktorá vyzýva pamiatku popraveného prezidenta Slovenského štátu, existuje už z tohto dôvodu prirodzená afinita časti rímskokatolíckych veriacich k tejto politickej strane.

Rómsky problém

Zrejme najvýraznejšou témou, na ktorej sa Ľudová strana Naše Slovensko profiluje, je rómska problematika. Práve problémy súžitia väčšinového obyvateľstva s neprispôsobivými členmi rómskej menšiny a ich dlhodobé neriešenie všetkými doterajšími vládami bez ohľadu na ich politickú príslušnosť, je živnou pôdou pre akceptovanie zdanlivo úderných a jednoduchých riešení, ktoré ponúka strana Mariana Kotlebu, voličmi žijúcich v takto postihnutých okresoch.

Maďarská menšina

Takmer dokonalá voličská vernosť maďarského obyvateľstva na Slovensku k stranám reprezentujúcim ich etnickú menšinu (toho času Strana maďarskej komunity a Most – Híd), negatívne ovplyvňuje tendenciu voliť iné strany a celkom očakávateľne medzi nimi aj nacionalistickú a vo vzťahu k menšinám minimálne kontroverznú ak nie nepriateľskú stranu akou je ĽSNS.

Banskobystrický kraj

Banskobystrický kraj a Banskú Bystricu možno považovať za politickú „baštu“ Mariana Kotlebu a jeho strany. V Banskej Bystrici vznikol najvyšším súdom rozpustený, ideologický predchodca ĽSNS, Slovenská pospolitosť, v Banskobystrickom samosprávnom kraji prišiel prvý nečakaný úspech Mariana Kotlebu, jeho zvolenie za „župana“ vo voľbách do VÚC v roku 2013, ktoré ĽSNS vydláždilo ďalšiu cestu do parlamentu a z Banskej Bystrice pochádza aj samotný Marian Kotleba. Preto možno predpokladať, že voliči žijúci v okresoch

Banskobystrického kraja, budú mať tendenciu voliť ĽSNS vo vyššej miere, než voliči vo zvyšku Slovenska.

Priestorovo-regresný člen

Relatívne ťažko kvantifikovateľným javom je pravidelná sociálna interakcia obyvateľov susediacich okresov, ktorá môže mať za následok podobné volebné preferencie u voličov. Ide teda o akési „prelievanie“ volebných sympatií k politickej strane cez hranice okresu. Tú však možno do modelu spoločne s prípadnými ďalšími ťažko kvantifikovateľnými resp. nepozorovanými, priestorovo korelovanými faktormi, vniesť do regresného modelu prostredníctvom priestorových regresných parametrov a k nim prislúchajúcim maticiam priestorových váh.

5 Aplikácia regresnej analýzy a zhodnotenie výsledkov

5.1 Viacnásobný regresný model

Východiskom regresnej analýzy je počiatočný viacnásobný regresný model (1), do ktorého sú zahrnuté všetky potenciálne vysvetľujúce premenné diskutované v predchádzajúcej kapitole:

$$\text{lsnspct} = \beta_0 + \beta_1 \text{bbkraj} + \beta_2 \text{muzipct} + \beta_3 \text{rkpct} + \beta_4 \text{vspct} + \beta_5 \text{nezamestnanipct} + \beta_6 \text{romoviapct} + \beta_7 \text{madaripct} + u$$

Po odhade MNŠ pozorujeme, že premenná *nezamestnanipct* bola v modeli (1) vyhodnotená ako štatisticky nevýznamná s takmer nulovým koeficientom. Príčinou tohto výsledku nemusí byť nutne jej kauzálna nezávislosť s vysvetľovanou premennou, ale multikolinearita s premennými *vspct* a *romoviapct* (vysoký podiel vysokoškolsky vzdelaných ľudí implikuje prirodzene nižšiu a vysoký podiel často sociálne vylúčených Rómov naopak vyššiu nezamestnanosť) v kombinácii s relatívne obmedzeným počtom pozorovaní (79 volebných okresov v r. 2016). Premennú *nezamestnanipct* pre jej nulový prínos napriek tomu z finálneho modelu (2) vypúšťame. Dopad na hodnoty zvyšných regresných koeficientov je minimálny, možno však badať mierne zvýšenie štatistickej významnosti premennej *romoviapct* (z desaťpercentnej hladiny na päťpercentnú), čo je však do istej miery zrejme následkom vyššie diskutovanej, vysokej závislosti medzi oboma premennými. Upravený model (2) má teda tvar:

$$\text{lsnspct} = \beta_0 + \beta_1 \text{bbkraj} + \beta_2 \text{muzipct} + \beta_3 \text{rkpct} + \beta_4 \text{vspct} + \beta_6 \text{romoviapct} + \beta_7 \text{madaripct} + u$$

Finálny model dosiahol relatívne vysokú hodnotu $R^2 = 0,812$ ($R^2_{adj} = 0,796$) so všetkými premennými významnými už na päťpercentnej hladine významnosti. Konštanta modelu je štatisticky nevýznamná, čo však vďaka jej nepraktickej interpretácii (volebný zisk ĽSNS hypotetickom okrese mimo Banskobystrický kraj, s nulovým podielom mužov, rímskych katolíkov, vysokoškolsky vzdelaných, Maďarov a Rómov), nie je až tak podstatné. Koeficienty ostatných regresorov, pokiaľ ide o ich kladnosť a zápornosť, nadobudli predpokladané hodnoty z teoretickej časti.

Tabuľka 2: Odhad viacnásobného regresného modelu MNŠ

PREMENNÉ	(1) lsnspct	(2) lsnspct
bbkraj	3.034*** (0.430)	3.028*** (0.403)
muzipct	0.617** (0.293)	0.615** (0.288)
rkpct	0.0430*** (0.00992)	0.0430*** (0.00968)
vspct	-0.141*** (0.0433)	-0.141*** (0.0421)
nezamestnanipct	-0.00223 (0.0555)	
romoviapct	0.0569* (0.0328)	0.0559** (0.0216)
madaripct	-0.104*** (0.0112)	-0.104*** (0.0111)
Konštanta	-22.39 (14.41)	-22.34 (14.24)
Pozorovania	79	79
R ²	0.812	0.812
Upravený R ²	0.794	0.796
F-test	43.85	51.88

Smerodajné chyby odhadov v zátvorkách

*** p<0.01, ** p<0.05, * p<0.1

Tabuľka 3: Korelačná matica regresorov

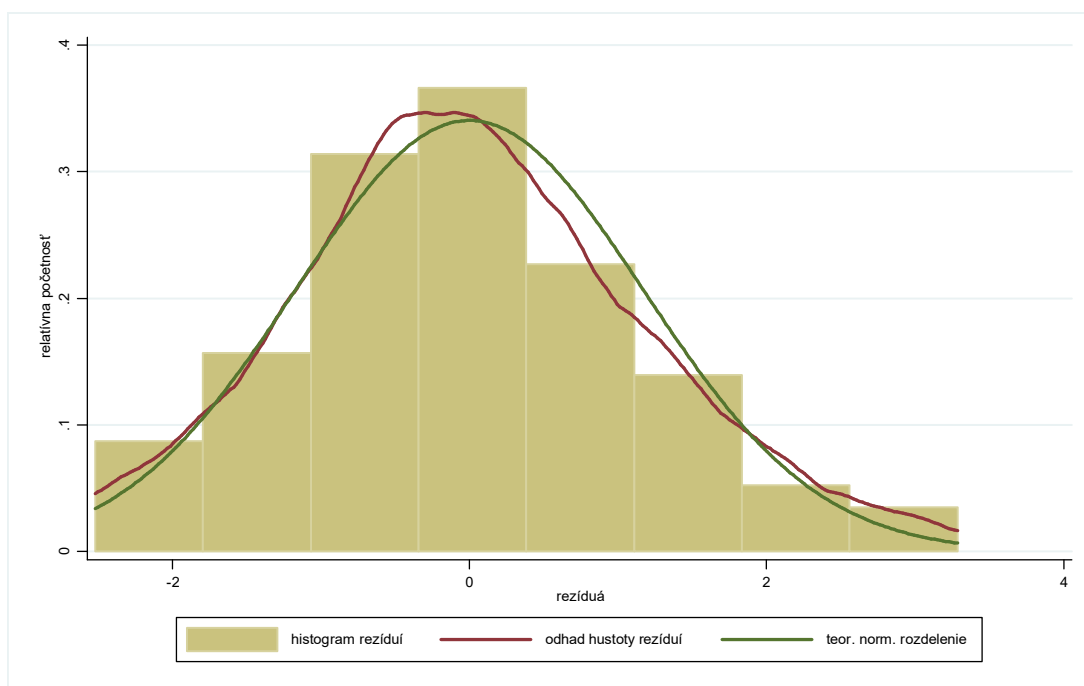
PREMENNÉ	bbkraj	muzipct	rkpct	vspct	nezamestnanipct	romoviapct	madaripct
bbkraj	1.00						
muzipct	-0.13	1.00					
rkpct	-0.03	0.42	1.00				
vspct	-0.13	-0.39	-0.28	1.00			
nezamestnanipct	0.36	0.30	-0.15	-0.50	1.00		
romoviapct	0.24	0.21	-0.20	-0.39	0.83	1.00	
madaripct	0.06	-0.14	-0.04	-0.18	0.18	0.21	1.00

5.2 Diagnostika modelu

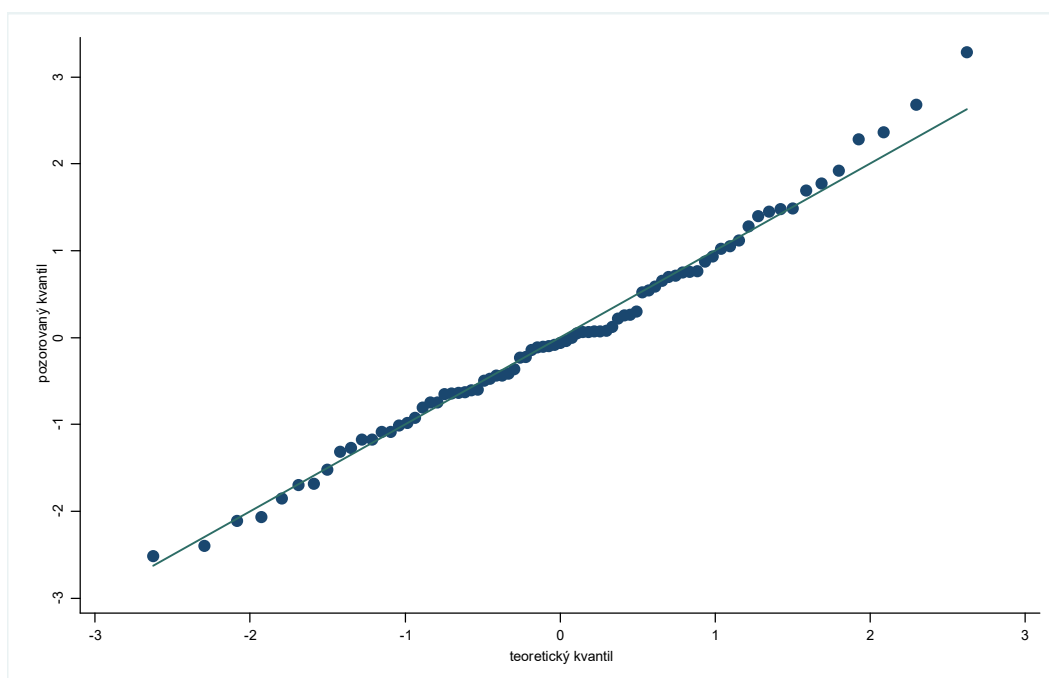
Model úspešne splnil všetky diagnostické testy. V žiadnom z nich neboli zamietnuté nulové hypotézy o správnej špecifikácii modelu, homoskedasticite, ani o normalite rezíduí. Kvantil-kvantilový graf a histogram rezíduí takisto nevykazujú žiadne extrémnejšie odchýlky od normálneho rozdelenia. Rezíduá však podľa z-skóre Moranovho I vykazujú priestorovú autokoreláciu. „Susednosť“ okresov v príslušnej matici priestorových váh **W** je zadefinovaná zdieľaním spoločnej hranice.

Tabuľka 4: Diagnostické testy

typ testu	testovacia štatistika	p-hodnota
RESET test	$F = 0.63$	0.5958
Breuschov–Paganov test	$nR^2 = 8.26$	0.2196
Whiteov test	$nR^2 = 21.25$	0.729
Jarqueov–Berov test	$JB = 1.127$	0.5692
Shapirov–Wilkov test	$W = 0.98975$	0.78572
Moranovo I	$z = 2.409$	0.008



Obrázok 2: Histogram rezíduí



Obrázok 3: Q-Q graf rezíduí

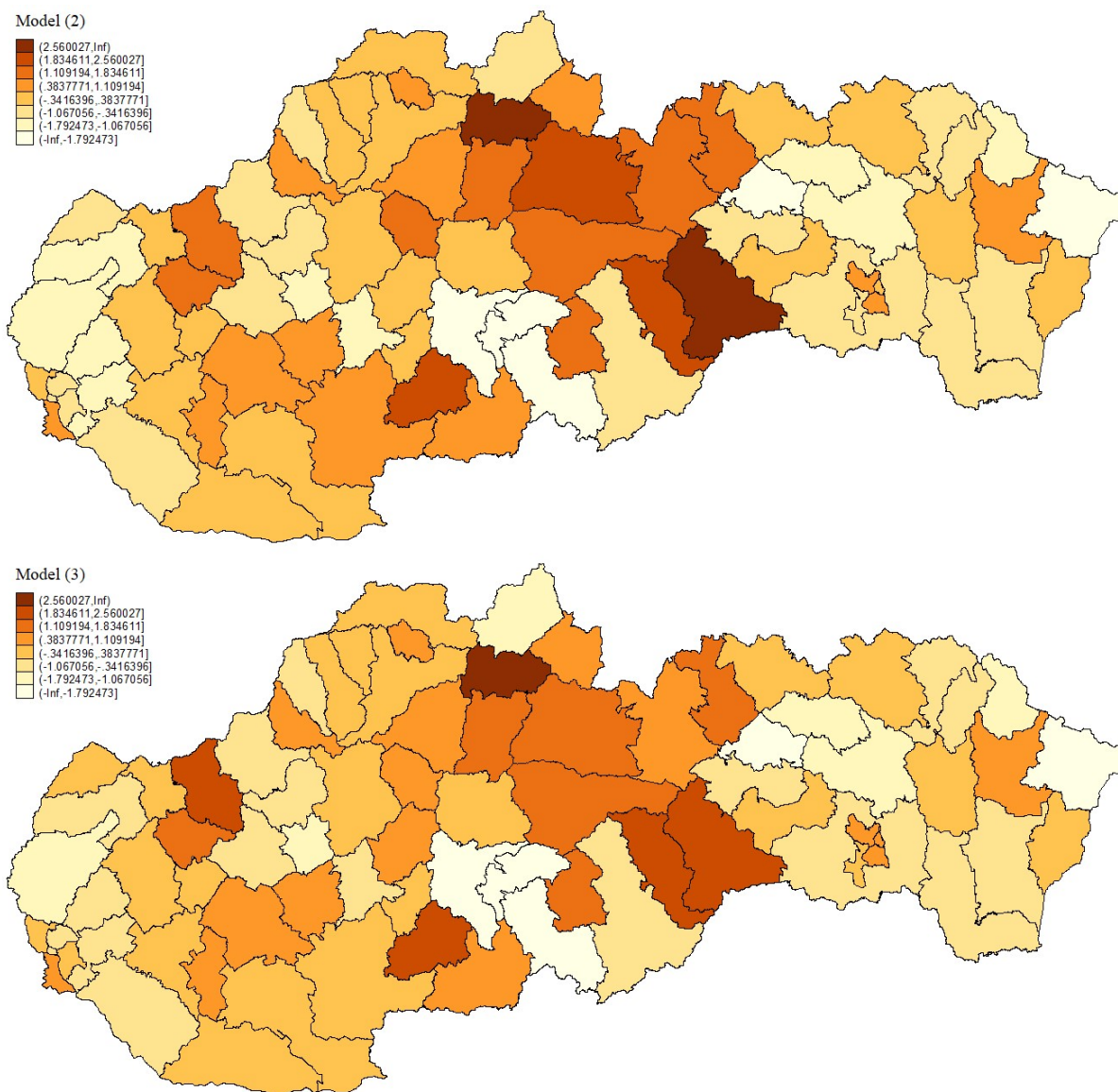
5.3 Priestorový regresný model

Keďže bola vo finálnom viacnásobnom regresnom modeli z predchádzajúcej podkapitoly pozorovaná štatisticky významná (p-hodnota = 0,008) priestorová autokorelácia rezíduí, pokúsime sa ju zohľadniť vnesením člena $\rho \mathbf{Wlsnspct}$ do regresnej rovnice, kde ρ je priestorový regresný parameter a $\mathbf{Wlsnspct}$ priestorovou maticou váh $\mathbf{W}$ vážené pozorovania vysvetľovanej premennej *lsnspct*. Výsledný priestorový intervalový model (3) má teda podobu:

$$\begin{aligned} \mathbf{lsnspct} = & \beta_0 + \beta_1 \mathbf{bbkraj} + \beta_2 \mathbf{muzipct} + \beta_3 \mathbf{rkpct} + \beta_4 \mathbf{vspct} + \beta_5 \mathbf{nezamestnanipct} \\ & + \beta_6 \mathbf{romoviapct} + \beta_7 \mathbf{madaripct} + \rho \mathbf{Wlsnspct} + \mathbf{u} \end{aligned}$$

V odhadnutom modeli je priestorový regresný parameter štatisticky významný na desaťpercentnej hladine významnosti (p-hodnota = 0,054). Všetky ostatné premenné si v porovnaní s predchádzajúcim modelom zachovali svoju štatistickú významnosť na príslušných hladinách, absolútne hodnoty ich koeficientov, až na výnimku premennej *rkpct*, mierne poklesli. Hodnota R^2 mierne vzrástla na 0,817 ($R^2_{adj} = 0,799$) a to napriek tomu, že bol odhad realizovaný metódou maximálnej vierohodnosti, ktorá R^2 v porovnaní s MNŠ nutne nemaximalizuje. Samotný efekt použitia priestorového regresného modelu v porovnaní s klasickou regresiou je badateľný, priestorová autokorelácia rezíduí je však napriek vzrastu

p-hodnoty z-skóre Moranovho I (z 0,008 na 0,038) stále štatisticky významná na päťpercentnej hladine. Mierne optické „vyčistenie“ zhľuku vyšších hodnôt rezíduí v oblasti stredného Slovenska možno pozorovať aj na priestorovej mape rezíduí.



Obrázok 4: Priestorové mapy rezíduí

Tabuľka 5: Porovnanie odhadov modelov

PREMENNÉ	(2) lsnspct	(3) lsnspct
bbkraj	3.028*** (0.403)	2.684*** (0.416)
muzipct	0.615** (0.288)	0.536** (0.271)
rkpct	0.0430***	0.0446***

	(0.00968)	(0.00905)
vspct	-0.141***	-0.114***
	(0.0421)	(0.0417)
romoviapct	0.0559**	0.0549***
	(0.0216)	(0.0201)
madaripct	-0.104***	-0.0946***
	(0.0111)	(0.0114)
Konštanta	-22.34	-20.42
	(14.24)	(13.29)
ρ		0.172*
		(0.0893)
Pozorovania	79	79
R^2	0.812	0.817
Upravený R^2	0.796	0.799
z-skóre Moranovho I	2.409***	1.779**

Smerodajné chyby odhadov v zátvorkách

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.4 Zhodnotenie výsledkov

Klasickému viacnásobnému (2) aj priestorovému regresnému modelu (3) sa podarilo vysvetliť viac, než 81 % rozptylu vysvetľovanej premennej *lsnspct*. Aplikácia priestorovej regresie v modeli (3) bola však úspešná len sčasti, keďže štatisticky významnú priestorovú autokoreláciu rezíduí vykazuje aj ten. Regresné koeficienty štatisticky významných vysvetľujúcich premenných dosiahli v oboch modeloch relatívne podobných (a čo sa týka znamienka neprekvapujúcich) hodnôt. Ich interpretácia pre potreby práce nie je až tak dôležitá, keďže jej cieľom bolo v prvom rade identifikovať faktory, ktoré boli „motorom“ resp. „brzdou“ volebného úspechu ĽSNS. Napriek tomu ju však uvedieme:

- V okresoch Banskobystrického kraja dosiahla strana Mariana Kotlebu v priemere o 3,028 % (2,684 %) vyšší volebný zisk, než v okresoch mimo neho.
- Každé percento podielu mužov v okrese navyše, implikovalo v priemere o 0,615 % (0,536 %) vyšší volebný zisk pre ĽSNS. Tento konkrétny koeficient je vzhľadom k jeho zjavne privysokej výške v oboch modeloch nepresne odhadnutý. To naznačuje už vyššia p-hodnota 0,036 (0,049) a široký interval spoľahlivosti 0,041 až 1,19 (0,041 až 1,067). Dôvodom môže byť prirodzene nízka variabilita (sm. odchýlka 0,794) pozorovaní príslušného regresoru, ktorým je percentuálny podiel mužskej populácie v okresoch.

- Každé percento rímskych katolíkov podielu v okrese navyše zvýšilo volebnú podporu strany Mariana Kotlebu o 0,043 % (0,045 %).
- Jednopercentný nárast podielu vysokoškolsky vzdelaných ľudí v okrese malo za následok pokles volebného zisku ĽSNS v ňom o 0,141 % (0,114%).
- Jedno percento podielu obyvateľov rómskej národnosti v okrese navyše zvýšilo volebný zisk ĽSNS v ňom o 0,056% (0,055%).
- Každé percento obyvateľov maďarskej národnosti v okrese navyše znížilo volebnú podporu strany Mariana Kotlebu o 0,104 % (0,095%).
- Kumulovaný efekt priestorovo korelovaných faktorov medzi susediacimi okresmi je 0,172% na každú jednotku pomocnej premennej **Wlsnspect**.

Všetky efekty sú samozrejme uvádzané „ceteris paribus“, teda pri nemennosti ostatných premenných.

Záver

Cieľom práce bolo identifikovať faktory, ktoré naprieč Slovenskom, na základe analýzy volebných výsledkov v okresoch, zvyšovali resp. znižovali volebný zisk Ľudovej Strany Naše Slovensko vo voľbách do NR SR v roku 2016. Zároveň bolo jej zámerom vytvoriť ekonometrický model, ktorý dostatočne presne popisuje volebné zisky ĽSNS v okresoch Slovenska. Pre dosiahnutie stanoveného cieľa boli v práci aplikované metódy viacnásobnej a priestorovej regresie.

V praktickej časti práce bolo zvolených osem potenciálnych faktorov, ktoré mohli mať vplyv na volebné výsledky ĽSNS v okresoch SR. Boli nimi percentuálne podiely mužov, vysokoškolsky vzdelaných, nezamestnaných, rímskych katolíkov, príslušníkov maďarskej a rómskej menšiny v okrese, príslušnosť okresu k Banskobystrickému kraju a pre priestorovú regresiu priestorovo-regresný člen reprezentujúci koreláciu volebného zisku ĽSNS medzi susednými okresmi, nezachytenú zvyšnými premennými. Sedem z nich bolo v následnej regresnej analýze identifikovaných ako štatisticky významných. Ako štatisticky nevýznamný bol identifikovaný faktor nezamestnanosti, príčinou toho však môže byť fakt, že regresná analýza bola limitovaná len 79 pozorovaniami. Rovnaká príčina zrejme prispela k tomu, že rozšírenie modelu o priestorovo-regresný člen, nemalo na jeho kvalitu až tak výrazný dopad. Všetky štatisticky významné premenné dosiahli, čo sa týka kladnosti a zápornosti, očakávaných hodnôt. Na základe nich možno usudzovať, že muži, rímski katolíci, občania žijúci v okresoch s výraznejšou prítomnosťou rómskej menšiny a občania okresov Banskobystrického kraja boli tými skupinami obyvateľstva, ktoré sa podieľali na prekvapivo vysokom volebnom výsledku strany Mariana Kotlebu. Naopak, vysokoškolsky vzdelaní a príslušníci maďarskej menšiny boli tými skupinami obyvateľstva, ktoré tento výsledok znižovali.

Rozšírenie práce by mohlo spočívať v zahrnutí ďalších dát vyplývajúcich z budúcich parlamentných volieb a k ich dátumu aktualizovaných demografických a socioekonomických údajov. Na základe nich by bolo možné oba doterajšie modely verifikovať resp. spresniť vnesením ďalších relevantných vysvetľujúcich premenných.

Použité zdroje a literatúra

- Anselin, Luc a Baltagi, Badi. 2003.** *A Companion to Theoretical Econometrics*. Malden : Wiley-Blackwell, 2003. s. 311-330. ISBN 978-1405106764.
- Biháriová, Irena. 2012.** *Vývoj pravicového extrémizmu na Slovensku v rokoch 1989 – 2012*. Bratislava : Friedrich Ebert Stiftung, 2012.
- Breusch, Trevor a Pagan, Adrian. 1979.** A simple test for heteroscedasticity and random coefficient variation. *Econometrica*. roč. 47, 1979, č. 5, s. 1287-1294.
- Davidson, Russell a MacKinnon, James. 2003.** *Econometric Theory and Methods*. New York : Oxford University Press, 2003. ISBN 9780195123722.
- Geodetický a kartografický ústav Bratislava. 2016.** Základná úroveň/ZBGIS - Administratívne hranice, mapový podklad formátu Esri SHP. [Online] 2016. https://www.geoportal.sk/files/zbgis/na_stiahnutie/shp/hranice_okresy_0.zip.
- Givens, Terri. 2004.** The Radical Right Gender Gap. *Comparative Political Studies*. 2004, č. 37, s. 30-54.
- Harteveld, Eelco, a iní. 2015.** The gender gap in populist radical right voting: examining the demand side in Western and Eastern Europe. *Patterns of Prejudice*. 2015, č. 49, s. 103-104.
- Jackman, Robert a Volpert, Karin. 1996.** Conditions Favouring Parties of the Extreme Right in Western Europe. *British Journal of Political Science*. 26, 1996, 4, s. 501–521.
- Jarque, Carlos a Bera, Anil. 1980.** Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*. roč. 6, 1980, č. 3, s. 255-259.
- Mikušovič, Dušan. 2009.** *Militantní demokracie na Slovensku?* Fakulta sociálnych štúdií, Masarykova univerzita v Brně. Brno : s.n., 2009. Diplomová práca.
- . **2007.** *Mimoparlamentná krajná pravica na Slovensku*. Fakulta sociálnych štúdií, Masarykova univerzita v Brně. Brno : s.n., 2007. Bakalárska práca.
- Milo, Daniel. 2005.** *Rasistický extrémizmus v Slovenskej republike*. Bratislava : Ľudia proti rasizmu, 2005.

Moran, Patrick. 1950. Notes on Continuous Stochastic Phenomena. *Biometrika*. roč. 37, 1950, č. 1-2, s. 17-23.

Podolinská, Tatiana a Hruštič, Tomáš. 2016. Čierno-biele svety. Rómovia v majoritnej spoločnosti na Slovensku. Bratislava : VEDA, 2016.

Ramsey, James. 1969. Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society*. roč. 31, 1969, č. 2, s. 350-371.

Shapiro, Samuel a Wilk, Martin. 1965. An analysis of variance test for normality (complete samples). *Biometrika*. roč. 52, 1965, č. 3-4, s. 591-611.

Štatistický úrad Slovenskej republiky. 2011. SODB 2011. [Online] 2011. [Dátum: 15. apríl 2017.] <https://census2011.statistics.sk/>.

—. **2010.** Voľby do Národnej rady Slovenskej republiky 2010. [Online] 2010. [Dátum: 15. apríl 2017.] <http://volby.statistics.sk/nrsr/nrsr2010/sr/tab3.jsp@lang=sk.htm>.

—. **2012.** Voľby do Národnej rady Slovenskej republiky 2012. [Online] 2012. [Dátum: 15. apríl 2017.] <http://volby.statistics.sk/nrsr/nrsr2012/sr/tab3.jsp@lang=sk.htm>.

—. **2016.** Voľby do Národnej rady Slovenskej republiky 2016. [Online] 2016. [Dátum: 15. apríl 2017.] <http://volby.statistics.sk/nrsr/nrsr2016/sk/data02.html>.

—. **2009.** Voľby do orgánov samosprávnych krajov 2009. [Online] 2009. [Dátum: 15. apríl 2017.] <http://volby.statistics.sk/osk/osk2009/sr/tab5.jsp@lang=sk.htm>.

—. **2013.** Voľby do orgánov samosprávnych krajov 2013. [Online] 2013. [Dátum: 15. apríl 2017.] http://volby.statistics.sk/osk/osk2013/VUC/Tabulka7_sk.html.

—. **2013.** Voľby do orgánov samosprávnych krajov 2013. [Online] 2013. [Dátum: 15. apríl 2017.] http://volby.statistics.sk/osk/osk2013/VUC/Tabulka5_sk.html.

Thode, Henry. 2002. *Testing For Normality*. Boca Raton : CRC Press, 2002. 9780824796136.

Weakliem, David. 2002. The Effects of Education on Political Opinions: An International Study. *International Journal of Public Opinion Research*. 2002, č. 14, s. 141-157.

White, Halbert. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*. roč. 48, 1980, č. 4, s. 817-838.

Wooldridge, Jeffrey. 2013. *Introductory Econometrics: A Modern Approach*. 5th edition. Mason : South-Western Cengage Learning, 2013. ISBN 9781111531041.