

# Senátní volby 2010 a voličské přesuny mezi prvním a druhým kolem

*Kamil Gregor, Michal Pink<sup>1</sup>*

Tento text byl zpracován v rámci Výzkumného záměru *Politické strany a reprezentace zájmů v soudobých evropských demokraciích* (kód MSM0021622407).

Abstract: *The 2010 Senate Elections and Voter Shifts between the First and Second Rounds.*

*The article concerns elections to the Senate of the Czech Parliament. Primarily, it refers to potential shifts of voter preference between the first and second rounds of the Senate elections. The fundamental tool for determining this trend is the Bayesian ecological inference model, through the use of which the authors attempt to answer the question of how the voters from the first round voted in the second round one week later when the options were reduced to only two candidates. The text analyzes these processes in detail in five electoral districts.*

Key words: *Bayesian ecological inference, Czech politics, Senatorial election.*

## 1. Úvod

Ustavení Senátu českého parlamentu představovalo jednu z dlouho nenaplněných částí české ústavy z prosince 1992. Po čtyřech letech od rozpadu federace se nakonec uskutečnily první volby do horní komory zákonodárního tělesa, které se postupně staly nedílnou součástí českých volebních soutěží. Během patnácti let se uskutečnilo několik desítek voleb, jejichž výsledkem byl zisk senátorského mandátu, který svou délkou šest let představuje v českých podmínkách nejdelší volenou funkci. Spolu s postupným začleňováním zmíněného typu volební soutěže se začala i odborná veřejnost věnovat této tematice. Zatím nejobsáhlejší je výzkum Tomáše Lebedy, Kláry Vlachové a Blanky Řehákové (2009), který pracuje s agregovanými daty i výsledky dotazníkového šetření. Zabývá se ale jen prvními senátními volbami. Další studie sledují zejména účinky použitého volebního systému na chování voličů (Matějů, Vlachová 1997; Chytilík 2005a; Chytilík 2005b; Malcová 2010; Navrátil 2010).

Ambicí předloženého textu není jen představení nového typu statistických metod, ale i podnícení dalšího zájmu o výzkum senátních voleb, které jsou svým způsobem v českém prostředí unikátní. Jedná se o jediný případ, kdy se volí na základě pravidel většinového

---

<sup>1</sup> Kontakt: Katedra politologie, Fakulta sociálních studií Masarykovy univerzity, Joštova 10, 602 00 Brno; e-mail: [kamil.gregor@mail.muni.cz](mailto:kamil.gregor@mail.muni.cz) (Kamil Gregor), [pink@fss.muni.cz](mailto:pink@fss.muni.cz) (Michal Pink).

volebního systému, který se v jiných evropských zemích používá pro parlamentní volby velice zřídka.<sup>2</sup> Právě tato výjimečnost, kdy volič volí dvakrát po sobě v relativně krátkém časovém intervalu navozuje otázky: Jak se volič zachoval v případě, že do druhého kola postoupil jeho preferovaný kandidát z prvního kola? Jak se naopak rozhodl v případě, že jeho kandidát nezískal potřebný počet hlasů a z voleb „vypadl“? Vzhledem k uvedeným otázkám bude cílem následujícího textu pomocí statistického modelování představit přesuny voličské podpory jednotlivých kandidátů mezi prvním a druhým kolem voleb do Senátu Parlamentu České republiky v roce 2010. Základní nástroj bude představovat bayesovský model ekologické inference (*bayesian ecological inference model*).<sup>3</sup> Ekologická inference představuje pokročilou kvantitativní metodu, která nachází uplatnění v širokém okruhu oborů jako je sociologie, politologie, ekonomie, geografie, historie, politický marketing, geografie, ekologie, pedagogika či epidemiologie – prakticky pokaždé, kdy máme k dispozici agregovaná data o lidském chování, ale sběr dat na individuální úrovni by byl buď neúměrně nákladný, technicky náročný, nedostatečně spolehlivý nebo i zcela nemožný.

Předložený text je zaměřen v první řadě na metodologický výklad zmíněného nástroje. Jednotlivé výstupy ekologické inference budou použity k zodpovězení otázky, jak hlasovali voliči z prvního kola senátních voleb v kole druhém? Základní hypotézy, které budeme prověřovat, jsou:

1. Druhého kola se účastní především voliči, kteří se účastnili i kola prvního.
2. Voliči udělují ve druhém kole nejčastěji hlas stejnému kandidátovi jako v prvním kole, pokud tento kandidát do druhého kola postoupil.

---

<sup>2</sup> Tuto výjimku představují volby do Národního shromáždění ve Francii a částečně Maďarsko.

<sup>3</sup> Počet česky psaných článků na toto téma je minimální, proto neexistuje ustálený překlad anglických výrazů *ecological inference* a *ecological fallacy*. Pat Lyons (2008) překládá *ecological inference* jako „ekologické usuzování“, zatímco Daniel Čermák (2010) používá pojem „ekologická inference“. Domníváme se, že by bylo vhodné využít existence dvou českých výrazů k rozlišení „ekologické inference“ jakožto specifické statistické metody a „ekologického usuzování“ jako obecnějšího pojmu vztahujícího se k jakémukoli problému agregace dat, tedy například i k problému modifikovatelných územních jednotek (*modifiable areal unit problem*). Při překladu pojmu *ecological fallacy* si můžeme vybrat ze slov „zkreslení“, „omyl“ a „klam“. První varianta je vyloženě nevhodná, protože navozuje dojem, že data o chování jednotlivců máme k dispozici, ale jsou nějakým způsobem zkreslená. Ze zbývajících dvou slov odpovídá podle našeho názoru anglickému originálu přesněji spíše „klam“ než „omyl“, protože pojem *ecological fallacy* odkazuje na vlastnosti dat a nikoli výzkumníka. Existuje výraz *fallacious conclusion* („klamný závěr“), ale nikoli *fallacious researcher* („mýlící se výzkumník“).

3. Voliči udělují ve druhém kole častěji hlas kandidátovi ideologicky podobné strany, pokud jejich kandidát z prvního kola nepostoupil.
4. Pokud do druhého kola postoupili dva kandidáti ideologicky podobných stran, voliči kandidátů z opačného pólu ideologického spektra se druhého kola voleb nejčastěji neúčastní.

Z hlediska teorií volebního chování jsou tyto hypotézy poměrně triviální. Jejich platnost byla v prostředí českých senátních voleb před rokem 2010 již z velké části prokázána dotazníkovými šetřeními (Lebeda a kol. 2009). Velice vhodně ovšem ilustrují fenomén zvaný ekologický klam (*ecological fallacy*), tedy nebezpečí vyvozování chybných závěrů o nepozorovaném chování jednotlivců v populaci na základě pozorovaného chování populace jako celku.

Představme si, že v hypotetickém volebním obvodu postoupí do druhého kola dva pravicoví kandidáti a zároveň výrazně poklesne volební účast v okrscích s množstvím příznivců levice. Prohlásit na základě této skutečnosti čtvrtou hypotézu za potvrzenou by znamenalo podstupovat riziko, že podléháme ekologickému klamu. Můžeme nanejvýš odhadovat, že se mezi nevoliči ve druhém kole opravdu nachází především levicoví voliči. Jedná se totiž jen o jeden z mnoha možných vzorů individuálního chování produkujících totožná agregovaná data. Ekologická inference je metodou, s jejíž pomocí lze v tomto případě modelovat podíl voličů levicových kandidátů z prvního kola, kteří se ve druhém kole k volebním urnám nedostavili.

První část textu stručně charakterizuje volební systém a popisuje okolnosti voleb v roce 2010. Druhá část definuje takzvaný problém ekologické inference (*ecological inference problem*) a podává výklad o historickém vývoji a principu fungování metod snažících se tento problém vyřešit statistickým modelováním. Třetí část za pomoci výše uvedených hypotéz modeluje voličské přesuny mezi prvním a druhým kolem senátních voleb v roce 2010 a následně popisuje vzory těchto přesunů ve vybraných obvodech.

## 2. Volby do Senátu v roce 2010

Senát je horní komorou Parlamentu České republiky. Každé dva roky je volena třetina z celkového počtu 81 senátorů. Pro účely senátních voleb je území státu rozděleno do zhruba stejně „lidnatých“ volebních obvodů, přičemž obvody, v nichž volby probíhají současně, jsou rovnoměrně rozmístěny po celém území České republiky.

Zvolený volební systém můžeme charakterizovat jako dvoukolový většinový systém s uzavřeným druhým kolem v jednomandátovém obvodu (Chytilík a kol. 2009: 141-144). Každý volič uděluje jednomu z kandidátů hlas, a pokud žádný kandidát nezíská nadpoloviční většinu

hlasů, koná se následující týden druhé kolo voleb, do něhož postupují dva kandidáti s nejvyšším počtem hlasů. Díky tomu zvolený senátor obdrží vždy nadpoloviční počet hlasů. Základním charakteristickým rysem senátních voleb v českém prostředí je také nízká volební účast, která nenarůstá i přes fakt, že mimo rok 1996 se senátní volby konaly vždy současně s krajskými nebo komunálními volbami.<sup>4</sup> (Lebeda 2009: 25-30)

Sledované volby se konaly v říjnu roku 2010 společně s komunálními volbami, necelých pět měsíců po volbách do Poslanecké sněmovny, kdy nejvíce hlasů sice získala ČSSD, ale post ministerského předsedy obsadila ODS a vítěz voleb se ocitl v opozici. Senátních voleb se účastnilo 226 kandidátů z 33 stran (Volby.cz 2011)<sup>5</sup> a jedna nestranická kandidátka (Miluše Horská v obvodu Pardubice). Žádný senátor nebyl zvolen již v prvním kole. Tabulka č. 1 podává informace o relativní úspěšnosti politických stran a o složení Senátu před a po volbách. V prvním kole dosáhla volební účast 44,59 % registrovaných voličů a v druhém kole 24,26 % (Volby.cz 2011).

Tabulka č. 1. Úspěšnost politických stran ve druhém kole voleb<sup>6</sup>

Politická strana	Původní složení komory	Obhajovaných mandátů	Postupů do 2. kola	Nově získaných mandátů	Obhájených mandátů	Ztracených mandátů	Nové složení komory
ČSSD	29	0	22	12	0	0	41
KDU-ČSL	3	1	3	1	1	0	4
KSČM	3	1	0	0	0	1	2
NEZ	0	0	1	0	0	0	0
ODS	36	19	19	2	6	12	25
S.cz	0	0	2	2	0	0	2
SNK ED	1	1	0	0	0	1	0
TOP 09	7	4	5	0	2	2	5
VV	0	0	1	0	0	0	0
Nestraník	1	0	1	1	0	0	2
<b>Celkem</b>	80 <sup>7</sup>	26	54	18	9	16	81

Zdroj: Senát ČR: Seznam senátorů 2011.

<sup>4</sup> Ve všech případech kromě jedné výjimky v roce 2004 bylo synchronizováno první kolo senátních voleb.

<sup>5</sup> Tomáš Finger, společný kandidát SPOZ a strany Suverenita kandidující v obvodu Praha 6, je v této práci považován za kandidáta SPOZ. Rovněž kandidáti navržení koalicí strany TOP 09 a sdružení Starostové a nezávislí jsou považováni za kandidáty TOP 09.

<sup>6</sup> Několik senátorů v období 2004–2010 změnilo stranickou příslušnost, zejména po vzniku TOP 09. Proto sloupec „Původní složení komory“ obsahuje počty členů senátních klubů a sloupec „Nové složení komory“ počty senátorů zvolených za jednotlivé strany.

<sup>7</sup> V obvodu Praha 6 bylo senátní křeslo uvolněno, když tamní senátor Karel Schwarzenberg získal v květnu roku 2010 poslanecký mandát. Doplňovací volby se zde nekonaly, protože se křeslo uvolnilo dříve než jeden rok před řádným termínem senátních voleb.

### 3. Historický vývoj a princip fungování metod ekologické inference

#### 3.1 Definice problému ekologické inference

Ekologická inference je označením používání agregovaných dat k odvozování závěrů o individuálním chování v situaci, kdy nejsou k dispozici žádná data na individuální úrovni (Schuessler 1999: 10578). Pro lepší uvedení ilustrujeme princip ekologické inference na jednoduchém příkladu. Představme si senátní obvod s  $n$  volebními okrsky. Každý zde registrovaný volič se při senátních volbách zachová jedním ze čtyř způsobů: 1. zúčastní se jen prvního kola voleb, 2. zúčastní se jen druhého kola, 3. zúčastní se obou kol voleb, 4. nezúčastní se ani jednoho kola. Za předpokladu, že elektorát obvodu zůstane v obou kolech voleb totožný a všechny odevzdané hlasy budou platné, můžeme pro okrsek  $i$  zapsat počty lidí, kteří se zachovali každým z těchto čtyř způsobů, do následující tabulky č. 2.

**Tabulka č. 2.** Možné kombinace účasti v prvním a druhém kole senátních voleb

$P_i$	$1 - P_i$	1
$\beta_i^v$	$D_i - \beta_i^v$	$D_i$
$\beta_i^n$	$(1 - D_i) - \beta_i^n$	$1 - D_i$

$P_i$  je podíl registrovaných voličů, kteří se zúčastnili prvního kola voleb,  $(1 - P_i)$  je podíl voličů, kteří se nezúčastnili prvního kola,  $D_i$  je podíl voličů, kteří se zúčastnili druhého kola, a  $(1 - D_i)$  je podíl voličů, kteří se nezúčastnili druhého kola.  $\beta_i^v$  udává podíl voličů účastnících se obou kol voleb,  $\beta_i^n$  udává podíl voličů, kteří se zúčastnili pouze prvního kola,  $(D_i - \beta_i^v)$  udává podíl voličů účastnících se jen druhého kola a  $[(1 - D_i) - \beta_i^n]$  udává podíl registrovaných voličů, kteří se nezúčastnili ani jednoho kola. Problém ekologické inference, známý i jako problém agregačního zkreslení (*aggregation bias problem*) nebo problém inference napříč úrovněmi (*cross-level inference problem*) (Lyons 2008: 51), označuje situaci, kdy pozorujeme pouze hodnoty  $P_i$  a  $D_i$  a hodnoty  $\beta_i^v$  a  $\beta_i^n$  zůstávají neznámé. Ekologická inference je definována jako metoda modelování hodnot těchto nepozorovaných proměnných na základě hodnot pozorovaných proměnných (Silva de Maltos, Veiga 2002: 4).

### 3.2 Počátky ekologického usuzování

Rané pokusy o ekologickou inferenci v dílech týkajících se politických jevů lze vystopovat až do 17. století (Graunt 1662; Petty 1690, 1691), přelomovou prací v této oblasti však byla až studie Wiliama Ogburna a Ineze Goltry z roku 1919, jedna z prvních politologických prací používajících vícerozměrnou statistickou metodu. Autoři zkoumali rozdíly v politických postojích mužů a žen, které krátce předtím získaly volební právo, a to pomocí statistické analýzy výsledků několika referend na různá témata ve městě Portland v Oregonu. Autoři uvádí: „Ačkoli volební pravidla neumožňují určit počet hlasů žen, nelze se nezamýšlet, zda existuje nějaká nepřímá metoda řešení tohoto problému.“ (Ogburn, Goltra 1919: 414)

Jejich metodou byla korelační analýza podílu žen a hlasů pro návrh v referendu ve velkém množství okrsků (*precints*). Na základě silných korelací mezi oběma proměnnými autoři odhadovali politické postoje žen. Zároveň ale předpokládali, že hlasování žen nebylo samo o sobě korelováno s jejich podílem na populaci okrsku. Kdyby to byli právě muži v okrscích s výraznou převahou žen, kdo ve velké většině hlasoval pro návrh, výsledné korelace by mohly být úplně stejné. Právě W. Ogburn a I. Goltra poprvé upozornili na podstatu tohoto problému, který je dnes známý pod pojmem ekologický nebo také skupinový klam (*group fallacy*).<sup>8</sup>

Záludnost ekologického klamu následně rozpoznala řada dalších autorů (například Allport 1924; Gehlke, Biehl 1934; Thorndike 1939; Deming, Stephan 1940; Yule, Kendall 1950), ale až William Robinson (1950) ho poprvé matematicky popsal a prokázal, že informaci o individuálním chování ztracenou při procesu agregace dat nelze žádným způsobem rekonstruovat. Je sice možné, aby korelace na úrovni agregovaných dat odrážely skryté vztahy na individuální úrovni, ale „podmínky, za kterých k tomu může dojít, jsou dosti vzdálené podmínkám, které se obvykle v datech vyskytují“ (Robinson 1950: 357). V extrémním případě může nastat takzvaný Simpsonův paradox, kdy jsou skutečné vztahy na individuální úrovni právě opačné než souvislosti pozorované na agregované úrovni (Simpson 1951).

Robinsonova pesimistická zjištění vedla k mohutnému odklonu od využívání agregovaných dat ve prospěch sběru dat na individuální úrovni (King 1997: 4-7). Zároveň však podnítila řadu pokusů, jak problém ekologické inference řešit pomocí statistického usuzování. Jednou z raných metod byla takzvaná ekologická regrese Lea Goodmana (1953, 1959). Vrátime-li se k hypotetickému senátnímu obvodu z tabulky č. 2, můžeme podíl voličů hlasujících v prvním kole vyjádřit jako

---

<sup>8</sup> Samotný pojem *ecological fallacy* zavedl Hanan C. Selvin (1958).

$$P_i = \beta_i^y D_i + \beta_i^n (1 - D_i). \quad (1)$$

Dva členy této rovnice neznáme, takže je principiálně neřešitelná. Lze ji ale převést na regresní model a přepokládat, že hodnoty  $\beta_i^y$  a  $\beta_i^n$  jsou konstantní ve všech okrscích. Právě to je podstatou Goodmanovy ekologické regrese (King 1997: 37-40). V důsledku tohoto silného a velice nerealistického předpokladu produkuje ekologická regrese často nesmyslné výsledky.<sup>9</sup> Goodmanova regrese nebere v potaz ani skutečnost, že počet voličů v obvodu nemůže být neomezeně veliký. I přes tyto problémy byla dlouho jedinou metodou ekologické inference a v zásadě každá další metoda na ni nějakým způsobem navazovala (King 1997: 35-41).

Odlíšný přístup zvolili ve stejné době Otis D. Duncan a Beverly Davis (1953). Jejich metoda hranic (*method of bounds*) umožňuje vypočítat nejvyšší a nejnižší možnou hodnotu obou nepozorovaných proměnných. Formálně lze výpočty metody hranic zapsat jako:

$$\beta_i^y \in \left[ \max\left(0, \frac{P_i - (1 - D_i)}{D_i}\right); \min\left(1, \frac{P_i}{D_i}\right) \right], \quad (2)$$

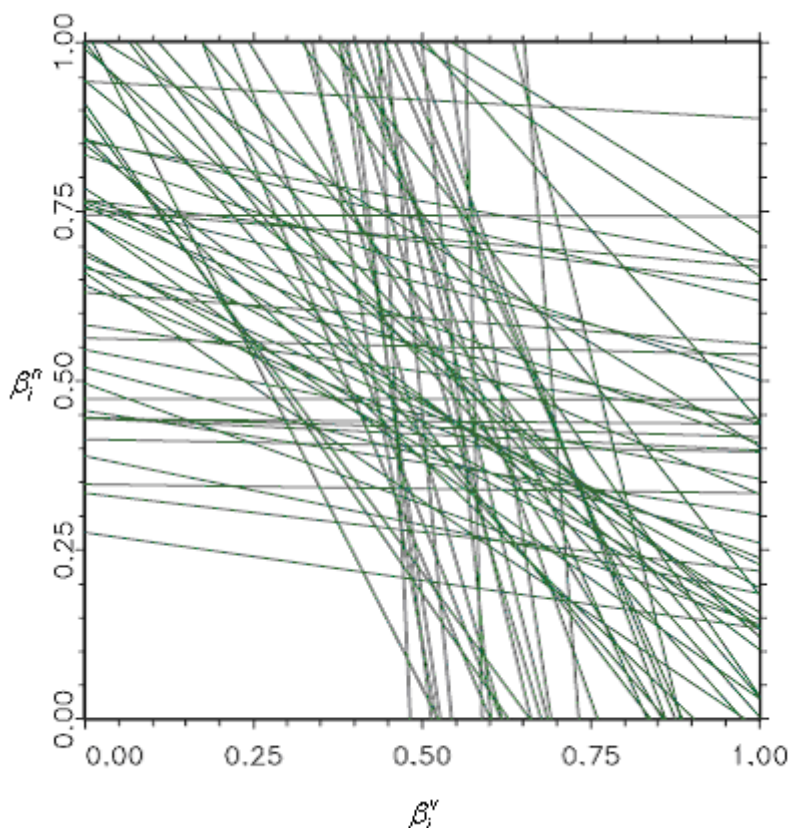
$$\beta_i^n \in \left[ \max\left(0, \frac{P_i - D_i}{1 - D_i}\right); \min\left(1, \frac{P_i}{1 - D_i}\right) \right]. \quad (3)$$

Hodnoty  $\beta_i^y$  a  $\beta_i^n$  musí samozřejmě vždy ležet na intervalu [0,1]. Metoda hranic ovšem umožňuje tento interval zúžit vzhledem k hodnotám  $P_i$  a  $D_i$ . Obrázek č. 1 je grafickým výstupem metody hranic. Pro každý okrsek je na horní a dolní osu grafu vynesena minimální a maximální možná hodnota  $\beta_i^y$  a na levou a pravou osu minimální a maximální možná hodnota  $\beta_i^n$ . Spojením těchto bodů vzniká linie, na níž se nachází všechny přípustné kombinace hodnot obou proměnných v daném okrsku.

---

<sup>9</sup> Při soudním přezkoumávání vytyčení hranic volebních obvodů ve státě Ohio v roce 1990 byl soudní znalec nucen prohlásit, že 109,63 % voličů černé pleti odevzdalo při posledních volbách v jednom z okrsků svůj hlas kandidátovi Demokratické strany (King 1997: 15). Různí autoři navrhli řadu způsobů, jak se s takovými výsledky vypořádat, například dovolili hodnotám  $\beta_i^y$  a  $\beta_i^n$  variovat, ale pouze v lineární závislosti na hodnotách  $P_i$ . Žádné z navrhovaných řešení nicméně neřeší podstatu problému (King 1997: 41, 57).

**Obrázek č. 1.** Tomograf intervalů možných hodnot nepozorovaných proměnných ve volebních okrscích



*Zdroj: Schnessler 1999: 10580.*

Metoda hranic nepracuje s žádnými apriorními předpoklady o hodnotách nepozorovaných proměnných, takže její výsledky jsou z empirického hlediska bezproblémové. Získané intervaly mají nicméně ve většině případů příliš široký rozsah a nejsou při vysvětlování volebního chování zajímavé (Lyons 2008: 53).

I přesto může být metoda hranic užitečným nástrojem při výzkumu okrsků, kde hodnota jedné z proměnných dosahuje extrémních hodnot. Dejme tomu, že chceme odpovědět na otázku, zda je volební účast v rasově heterogenní zemi vyšší u bělošského nebo u černošského obyvatelstva. Určit podíl černošských voličů v okrsku se 100 % černošských obyvatel je snadné – černé pleti je zde každý volič. Na výše položenou otázku proto lze odpovědět srovnáním volební účasti v okrscích s výhradně bělošským nebo černošským obyvatelstvem.

Dá se však předpokládat, že v rasově heterogenní zemi bude počet takových okrsků poměrně malý. Můžeme ho ale zvýšit, označíme-li za rasově homogenní každý okrsek, kde podíl obyvatel jedné rasy přesahuje, řekněme, 95 %. Pak lze porovnávat intervaly hodnot získané pomocí metody hranic (Herron, Sekhon 2005; Lau a kol. 2007: 45).

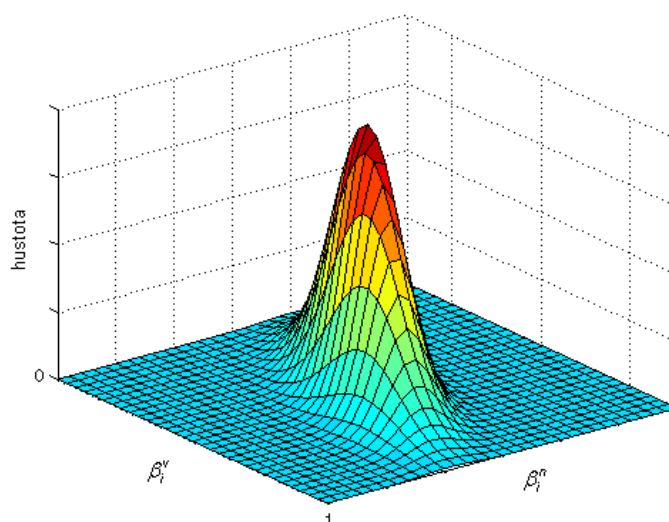


V tomto případě již vzniká riziko ekologickému klamu. Šance, že se dopustíme chybných závěrů, je tím vyšší, čím nižší podíl obyvatel jedné rasy stanovíme jako kritérium etnický homogenního okrsku. Nemůžeme navíc ověřit, zda naše zjištění platí i v rasově heterogenních okrscích. Neodhalíme například Simpsonův paradox, tedy situaci, kdy jsou skutečné vztahy na individuální úrovni právě opačné než souvislosti pozorované na agregované úrovni (Simpson 1951).

### 3.3 Nebayesovský model ekologické inference Garyho Kinga

Na další krok ve vývoji ekologické inference si výzkumníci museli počkat téměř 50 let. Byl jím inferenční model Garyho Kinga (1997), který jako první propojil metodu hranic s Goodmanovou ekologickou regresí. Model se skládá ze dvou kroků. Nejprve je metodou hranic určen interval, v němž leží skutečné hodnoty nepozorovaných proměnných. Poté je všem hodnotám na tomto intervalu přiřazena určitá relativní pravděpodobnost, že budou pravdivými hodnotami. Použité vzorce přitom vychází z výpočtů ekologické regrese. Na rozdíl od ní však Kingův model nevyžaduje extrémně nerealistický předpoklad konstantní hodnoty  $\beta_i^v$  a  $\beta_i^n$  ve všech okrscích. Místo toho předpokládá, že rozdělení hodnot těchto dvou proměnných odpovídá dvourozměrnému normálnímu rozdělení.<sup>10</sup> Jeden z možných tvarů tohoto rozdělení ukazuje obrázek č. 2.

**Obrázek č. 2.** Dvourozměrné normální rozdělení

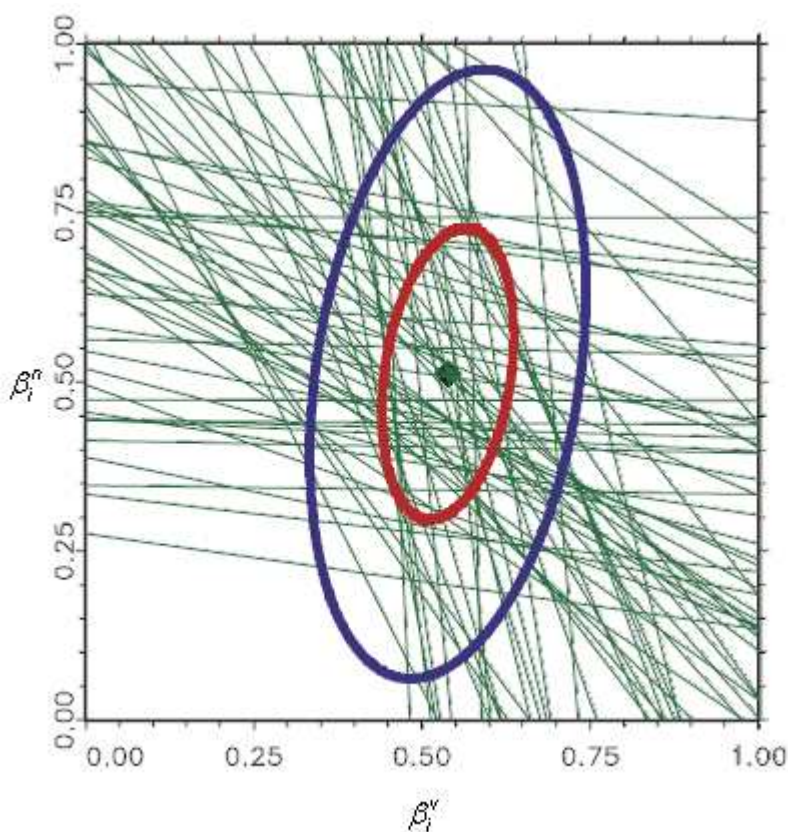


Zdroj: Autoři na základě *Multivariate Normal Distribution* 2011.

<sup>10</sup> Dalším předpokladem je absence prostorové autokorelace hodnot nepozorovaných proměnných (Schuessler 1999: 10580).

Obrázek č. 3 graficky znázorňuje podstatu využití tohoto rozdělení. Černý bod označuje místo s relativně nejvyšší hustotou linií. Graf je pomyslně překryt dvourozměrným normálním rozdělením podobným tomu na obrázku č. 2. Jeho tvar odpovídá výstupům Goodmanovy ekologické regrese, přičemž vrchol rozdělení se nachází právě v bodě s nejvyšší hustotou čar. Modrá a červená křivka vymezuje oblast grafu s 95% a 50% pravděpodobností výskytu pravdivých hodnot.

**Obrázek č. 3.** Tomograf intervalů možných hodnot nepozorovaných proměnných překrytý dvourozměrným normálním rozdělením



Zdroj: *Schuessler 1999: 10580.*

Validita Kingova modelu závisí na splnění předpokladu o rozdělení hodnot  $\beta_i^v$  a  $\beta_i^n$ . Tento poměrně silný předpoklad proto kritizuje řada autorů. Pomocí simulací nebo porovnávání výsledků modelu a dotazníkových šetření upozorňují na chybné závěry v situacích, kdy předpoklad není splněn (například Freedman a kol. 1998; Cho 1998; Gelman a kol. 2001; Cho, Gaines 2004; Park 2004, 2008; Wakefield 2008).

Výběr nesprávného rozdělení není zapříčiněn jeho chybnou diagnostikou. Základní nedostatek je mnohem závažnější. Skutečné rozdělení nepozorovaných hodnot v populaci totiž z principu zjistit nelze. Narážíme zde na problém zásadní statistické neurčitelnosti (*basic indeterminacy problem*). Validita statistického usuzování je vždy založena na platnosti předpokladů, jenže v tomto případě nemáme pro prověření jejich platnosti dost informací. Kdyby byly takové údaje k dispozici, nemuseli bychom modelování vůbec provádět (King 1997: 39-46).

### 3.4 Bayesovská ekologická inference

G. King, Ori Rosen a Martin Tanner (1999) se tento problém snažili vyřešit pomocí bayesovského hierarchického modelování. Navrhují zařadit před původní dva kroky inferenčního modelu třetí krok, jehož cílem je určit za pomoci Bayesova teorému vhodné rozdělení hodnot proměnných  $\beta_i^v$  a  $\beta_i^n$  na základě rozdělení hodnot pozorovaných proměnných. Autoři nahrazují dvourozměrné normální rozdělení beta rozdělením, jehož výhodou je skutečnost, že může mít více než jeden vrchol. Konkrétní tvar rozdělení je stanoven metodou matematické simulace za použití některého z výpočetních algoritmů označovaných jako Markovovy řetězce (*Markov chains*) (viz formální popis modelu v King a kol. 1999; King a kol. 2004; Silva de Maltos, Viega 2004).

Zatímco v původním Kingově modelu upozorňoval na pravděpodobnou neplatnost předpokladů o rozdělení nepozorovaných proměnných neobvyklý tvar linií v grafu metody hranic (viz obrázek č. 1), v případě hierarchických bayesovských modelů jsou veškeré diagnostické nástroje přímo součástí modelu. Nedostatek dodatečných informací o zkoumaném jevu znemožňuje provádět „obvyklé“ testy statistické signifikance. Nutno podotknout, že tato komplikace nebrání řadě sociálních vědců aplikovat modely ekologické inference i v případech, kdy jejich validitu nelze nijak ověřit, a následně zpochybňovat dosavadní znalosti o daném jevu shromážděné jinými metodami (například O'Loughlin 2002; Kopstein, Wittenberg 2003, 2004, 2009, 2010a, 2010b, 2011a, 2011b; Herron, Sekhon 2005; King a kol. 2008).

Vzhledem k problému základní statistické neurčitelnosti se klíčovou vlastností modelu ekologické inference stává jeho robustnost, tedy míra chyby modelu při odchýlení rozdělení hodnot nepozorovaných proměnných od předpokládaného rozdělení. Bayesovské modely ekologické inference jsou poměrně nové, proto je literatura na toto téma zatím relativně chudá. Zcela nedávným příspěvkem v této oblasti je studie Martina Elffa, Thomase Gschwenda a Rona J. Johnstona (2008), kteří konstruují bayesovský model ekologické inference, jehož validitu lze měřit na základě principu maximální entropie (*principle of maximum entropy*). Softwarové nástroje nutné pro programování tohoto modelu ale zatím nejsou volně dostupné.

Při hodnocení robustnosti modelu proto nezbývá než odkázat na odbornou literaturu zabývající se měřením přesnosti modelů v simulovaných extrémních případech nebo srovnáváním jejich výsledků s výsledky dotazníkových šetření (Cho 1998; Ferree 1999; King a kol. 1999; Rosen a kol. 2001; Silva de Maltos, Veiga 2002; Tomz a kol. 2002; Cho, Gaines 2004; King a kol. 2004; Park 2004; Elff a kol. 2008; Park 2008; Wakefield 2008; Glynn, Wakefield 2009). Zatím nejobsáhlejší takovou studii provedli Lucas Leemann a Philipp Leimgruber (2009), kteří měřili „přesnost“<sup>11</sup> šesti modelů ekologické inference náboženského vyznání voličů hlasujících ve 113 švýcarských referendech.

Vzhledem ke složitosti celé problematiky se výzkumníci až donedávna věnovali téměř výhradně problému ekologické inference v případech, kdy sledované proměnné nabývají pouze dvou hodnot. Možnosti praktického využití takových metod jsou ovšem velmi omezené, protože naprostá většina proměnných v sociálních vědách nabývá více než dvou hodnot. I při studiu dokonale dvoustranického systému pracujeme se třemi kategoriemi („volič strany A“, „volič strany B“, „nevolič“).

Zpočátku odborníci doporučovali prostě sloučit kategorie modelovaných proměnných do dvou (Ferree 1999: 1). Takový postup samozřejmě není ve většině případů žádoucí, pokud je vůbec smysluplný. Teprve bayesovské usuzování umožnilo vytvořit funkční modely ekologické inference proměnných s více než dvěma kategoriemi (například Ferree 1999; Rosen a kol. 2001).

Model používaný v této práci je rozšířením výše popsaného baysovského modelu (King, Rosen, Tanner 1999). Hlavní rozdíl spočívá v nahrazení beta rozdělení Dirichletovým rozdělením, které je jeho vícerozměrným zobecněním. Použitým Markovovým řetězcem je takzvaný *Gibbs sampler* (viz formální popis modelu v Rosen a kol. 2001). V současné době jde o jeden z mála baysovských modelů ekologické inference vícekategoriálních proměnných, pro který existuje software, byť zatím jen v základní, uživatelsky nepříliš přátelské podobě.<sup>12</sup> Tento model je rovněž zdaleka nejčastěji prakticky využíván.

---

<sup>11</sup> „Přesnost“ je v případě podobných studií vhodné uvádět v uvozovkách, protože i výsledky dotazníkových šetření se samozřejmě do určité míry odchylují od reálných nepozorovaných hodnot.

<sup>12</sup> Výpočty v této práci byly provedeny pomocí volně dostupného softwarového balíčku *Zelig* (Imai a kol. 2007), který využívá programovací jazyk R určený pro statistickou analýzu dat (The R Project for Statistical Computing 2011).

## 4. Voličské přesuny mezi prvním a druhým kolem senátních voleb v roce 2010

### 4.1 Testování vytyčených hypotéz

Tato část článku popisuje platnost jednotlivých hypotéz představených v úvodu práce, a to na základě výsledků modelů ekologické inference. Začneme první hypotézou, že se druhého kola senátních voleb účastní především voliči, kteří k urnám přišli i v kole prvním. V zásadě existují tři možné vztahy mezi účastí v obou kolech. Buď ve druhém kole hlasují zcela jiní voliči než v kole prvním, nebo jsou voliči druhého kola podmnožinou voličů z kola prvního, nebo se obě množiny částečně překrývají. Tabulka č. 3 ukazuje výsledek modelování volební účasti v prvním a druhém kole. V tomto případě byla při modelování použita data ze všech 27 volebních obvodů.<sup>13</sup> 90,67 % voličů druhého kola se účastnilo i kola prvního. Pouze 9,33 % voličů, kteří přišli k urnám ve druhém kole, nehlasovalo v kole prvním. Na základě tohoto zjištění můžeme první hypotézu prohlásit za prokázanou, tedy alespoň v případě senátních voleb v roce 2010.

Tabulka č. 3. Přesuny mezi voliči a nevoliči<sup>14</sup>

		První kolo		Celkem
		Voliči	Nevoliči	
Druhé kolo	Voliči	90,67 %	09,33 %	100 %
	Nevoliči	25,30 %	74,70 %	100 %

Zdroj: Autoři na základě výpočtu v programu *Zelig* (Imai a kol. 2007), použitá data lze nalézt na serveru *Volby.cz* 2011.

Podle druhé hypotézy voliči udělují ve druhém kole nejčastěji hlas kandidátovi z prvního kola, pokud postoupil. Tabulka č. 4 ukazuje, že tomu tak v roce 2010 skutečně bylo. 57,18 % voličů, jejichž kandidát z prvního kola postoupil, se zúčastnilo i kola druhého a dalo mu hlas znovu. Pouze 9,84 % voličů hlasovalo pro protikandidáta. Z tabulky je rovněž patrné, že voliči kandidátů

<sup>13</sup> Nutno podotknout, že mezi prvním a druhým kolem došlo k drobným změnám v počtu registrovaných voličů v jednotlivých okrscích, což pochopitelně zkresluje výsledky modelu. Tyto změny byly ovšem zcela marginální a můžeme je ignorovat. Opomenuty jsou rovněž neplatné hlasy.

<sup>14</sup> Datový soubor je tvořen úplnými výsledky obou kol senátních voleb na úrovni 4.860 volebních okrsků. Podrobný popis dílčích úprav a kalibrace modelů prezentovaných v této práci by byl z velké části příliš technický a vyžadoval by alespoň stručný výklad o základech bayesovské pravděpodobnostní statistiky. Přesto považujeme za nutné uvést zde některé údaje, s jejichž pomocí lze výpočty přesně replikovat. Při výpočtu všech modelů bylo provedeno 100.000 simulací aposterioriho rozdělení. Počáteční inicializace generátoru pseudonáhodných čísel Markovova řetězce (*random seed*) je posloupností celých kladných čísel. Parametry Dirichletova a gamma rozdělení byly ponechány přednastavené (viz Imai a kol. 2007: 184).

vypadlých po prvním kole se většinou (60,69 %) již druhého kola neúčastnili. I druhou hypotézu proto můžeme v případě senátních voleb v roce 2010 považovat za potvrzenou.

**Tabulka č. 4.** Přesuny voličů mezi totožnými a jinými kandidáty

		První kolo		
		Voliči prvních dvou	Voliči ostatních	Nevoliči
Druhé kolo	Voliči stejného	57,18 %	00,00 %	03,91 %
	Voliči jiného	09,84 %	39,31 %	
	Nevoliči	32,98 %	60,69 %	96,09 %
<b>Celkem</b>		100,00 %	100,00 %	100,00 %

*Zdroj: Autoři na základě výpočtu v programu Zelig (Imai a kol. 2007), použitá data lze nalézt na serveru Volby.cz 2011.*

Při prověřování hypotézy, že voliči udělují ve druhém kole hlas kandidátovi ideologicky podobné strany, jestliže jejich kandidát z prvního kola nepostoupil, narážíme na dvě komplikace. Především je nutné kategorizovat politické strany z hlediska vzájemné ideologické podobnosti. Roman Chytilík a Otto Eibl (2011) identifikují v nejnovějším příspěvku k diskuzi o podobě politického prostoru v České republice dvě relevantní dimenze, pro jednoduchost ovšem můžeme předpokládat jednorozměrné rozdělení stran na pravolevé ose se třemi kategoriemi: levicové, pravicové a ostatní strany. Mezi levicové strany počítáme ČSSD, KSČM a SPOZ a mezi pravicové strany KDU-ČSL, ODS, SNK ED, SSO a TOP 09.

Nezařazeny zůstávají strany s velmi malým počtem voličů (například ČSNS, DSSS, KČ), regionální strany (M, NB, S.cz a podobně) a strany, u nichž je jednoznačné zařazení na pravolevé ideologické škále problematické (zejména ČPS, NK, SZ, VV). Při řazení stran do kategorií se lze ve většině případů celkem spolehlivě opřít o dosavadní odbornou literaturu (například Malíř, Marek 2005: 1413-1488, 1503-1530, 1653-1665), v případě nově vzniklých stran SPOZ a TOP 09 snad nebude přiřazení k levici, respektive k pravici, diskutabilní. Kandidáti nezařazených stran získali ve všech 27 obvodech v prvním kole celkem 17,22 % hlasů. Domníváme se, že z hlediska testování hypotéz je rigoróznější ponechat tyto strany v kategorii „ostatní“, než je za každou cenu přiřazovat k levici či pravici, v mnoha případech čistě arbitrárně.

Druhá komplikace spočívá ve skutečnosti, že se v jednotlivých obvodech voleb účastnili kandidáti různých politických stran. Proto je v tomto případě třeba provést 27 samostatných modelů, přičemž tabulka č. 5 představuje jednoduchou, ale zároveň smysluplnou agregaci těchto výpočtů. Hodnoty  $p_i$  jsou aritmetickými průměry podílů voličů, kteří se účastnili druhého kola v obvodech, kde jejich kandidát z prvního kola nepostoupil. Základem těchto podílů přitom není počet všech registrovaných voličů v obvodu, ale voličů, kteří ve druhém kole přišli k urnám. Tabulka dále udává směrodatnou odchylku hodnot od aritmetického průměru ( $\sigma$ ) a počet

případů, kdy se daná kombinace kandidátů vyskytla ( $n_t$ ). Vyloučeny byly případy, kde do druhého kola nepostoupil kandidát ideologicky příbuzné strany.<sup>15</sup>

**Tabulka č. 5.** Přesuny voličů mezi ideologicky spřízněnými kandidáty

		Druhé kolo											
		ČSSD			KDU-ČSL			ODS			TOP 09		
		$p_t$	$\sigma_t$	$n_t$	$p_t$	$\sigma_t$	$n_t$	$p_t$	$\sigma_t$	$n_t$	$p_t$	$\sigma_t$	$n_t$
První kolo	ČSSD				12,31 %	0,055	3	00,00 %	0	0	00,00 %	0	0
	KDU-ČSL	43,60 %	0,135	11				55,65 %	0,133	13	53,50 %	0,058	3
	KSČM	75,72 %	0,098	21	26,77 %	0,101	3	22,30 %	0,097	14	18,47 %	0	1
	ODS	28,37 %	0,072	6	73,13 %	0,081	3				73,03 %		1
	SNK ED	53,23 %	0	1	00,00 %	0	0	46,77 %	0	1	00,00 %	0	0
	SPOZ	56,06 %	0,148	14	41,18 %	0,131	3	44,43 %	0,171	10	39,76 %	0	1
	SSO	37,49 %	0,148	3	00,00 %	0	0	59,19 %	0,193	2	69,15 %	0	1
	SZ	61,25 %	0,139	2	00,00 %	0	0	44,88 %	0,145	3	42,86 %	0	1
	<b>TOP 09</b>	27,72 %	0,107	18	77,69 %	0,051	3	69,14 %	0,107	15			

Zdroj: Autoři na základě výpočtu v programu Zelig (Imai a kol. 2007), použitá data lze nalézt na serveru Volby.cz 2011.

Voliči, jejichž kandidát z prvního kola nepostoupil, dávali ve druhém kole mnohem častěji hlas ideologicky příbuznému kandidátovi. Například ve 21 obvodech, kde do druhého kola postoupili kandidáti ČSSD, jim dalo hlas v průměru 75,72 % voličů KSČM. Naopak pouze průměrně 12,31 % voličů ČSSD ve třech obvodech, kde do druhého kola postoupil kandidát KDU-ČSL, a nikoli ČSSD, udělilo svůj hlas kandidátovi KDU-ČSL. Na základě údajů v této tabulce můžeme proto i třetí hypotézu považovat za potvrzenou.

Zbývá prověřit poslední zmíněnou hypotézu, podle níž bychom měli v obvodech, kde do druhého kola postoupili dva ideologicky podobní kandidáti, pozorovat nižší účast voličů neúspěšných kandidátů z opačného konce ideologického spektra. Tuto hypotézu můžeme testovat pouze částečně. Ve většině obvodů totiž postoupili do druhého kola kandidáti levice i pravice. Pouze ve čtyřech volebních obvodech se utkali kandidáti ideologicky podobných stran (vždy ODS a TOP 09). Nikde nedošlo k utkání dvou levicových kandidátů. Úplný přehled o stranické příslušnosti kandidátů ve druhém kole podává tabulka č. 6.

<sup>15</sup> Například v žádném obvodu, kde do druhého kola postoupil kandidát ODS, ale nikoli kandidát ČSSD, se o mandát neucházel kandidát jiné levicové strany. Nelze proto sledovat, zda voliči ČSSD dávali hlas spíše kandidátovi ODS, nebo levicovému protikandidátovi.

**Tabulka č. 6.** Stranická příslušnost kandidátů ve druhém kole

		Vítěz ve druhém kole								Celkem
		ČSSD	KDU-ČSL	NEZ	ODS	S.cz	TOP 09	VV	Nestraník	
Kandidáti ve druhém kole	ČSSD, KDU-ČSL	1	2	0	0	0	0	0	0	3
	ČSSD, NEZ	1	0	0	0	0	0	0	0	1
	ČSSD, ODS	8	0	0	6	0	0	0	0	14
	ČSSD, S.cz	0	0	0	0	1	0	0	0	1
	ČSSD, TOP 09	1	0	0	0	0	0	0	0	1
	ČSSD, VV	1	0	0	0	0	0	0	0	1
	ČSSD, Nestraník	0	0	0	0	0	0	0	1	1
	ODS, S.cz	0	0	0	0	1	0	0	0	1
	ODS, TOP 09	0	0	0	2	0	2	0	0	4

*Zdroj: Volby.cz 2011.*

Je možné sledovat, zda se voliči ČSSD, KSČM a SPOZ v obvodech, kde nepostoupil kandidát ČSSD, účastnili druhého kola méně často než voliči těchto stran v obvodech, kde kandidát ČSSD postoupil. Všechny údaje popisující uvedený jev představuje tabulka č. 7. V případě ČSSD nemáme žádné případy ke srovnání, protože jde o jedinou levicovou stranu, jejíž kandidáti postoupili do druhého kola, a zároveň o stranu, která měla ve druhém kole svého kandidáta téměř ve všech obvodech. Je nicméně patrné, že tam, kde proti sobě stáli kandidáti ODS a TOP 09, byla neúčast voličů ČSSD opravdu vysoká (81,71 %). U voličů KSČM jsou výsledky modelů zcela v souladu s očekáváním. Míra jejich účasti výrazně klesala, pokud do druhého kola nepostoupil kandidát ČSSD. Avšak pokud jde o voliče SPOZ, žádný takový vztah nepozorujeme. Hodnoty jejich neúčasti se v jednotlivých obvodech výrazně liší. Zatímco v obvodu Nový Jičín přišlo ve druhém kole k urnám plných 87,77 % voličů kandidáta SPOZ Zdeňka Krajčírě a většina z nich (78,81 %) podpořila Jaroslava Palase z ČSSD, v obvodu Liberec se druhého kola účastnilo pouze 15,64 % voličů kandidáta SPOZ Milana Drobneho. I zde ovšem 71,28 % z nich podpořilo Stanislava Eichlera z ČSSD proti bývalému předsedovi Senátu Přemyslu Sobotkovi z ODS. Hypotézu číslo čtyři každopádně nelze plnohodnotně prověřit, protože námi sledované senátní volby neobsahují ani jeden případ, kdy do druhého kola postoupili dva levicovní kandidáti.



**Tabulka č. 7.** Neúčast ve druhém kole v závislosti na ideologickém profilu kandidátů

		Druhé kolo											
		Levice vs. pravice			Pravice vs. pravice			Levice vs. jiný			Pravice vs. jiný		
		$p_u$	$\sigma_u$	$n_u$	$p_u$	$\sigma_u$	$n_u$	$p_u$	$\sigma_u$	$n_u$	$p_u$	$\sigma_u$	$n_u$
První kolo	ČSSD				81,71 %	0,097	4				36,91 %	0	1
	KDU-ČSL	62,75 %	0,119	11	34,06 %	0,097	2	79,63 %	0,103	2	66,69 %	0	1
	KSČM	59,13 %	0,197	18	83,67 %	0,113	4	48,71 %	0,336	3	64,37 %	0	1
	ODS	52,04 %	0,126	4				60,24 %	0,171	3			
	SNK ED	56,33 %	0	1	00,00 %	0	0	00,00 %	0	0	00,00 %	0	0
	SPOZ	54,53 %	0,232	14	55,15 %	0,262	3	64,24 %	0	1	00,00 %	0	0
	SSO	74,68 %	0,114	3	00,00 %	0	0	00,00 %	0	0	00,00 %	0	0
	SZ	51,52 %	0,073	2	64,88 %	0	1	00,00 %	0	0	00,00 %	0	0
	<b>TOP 09</b>	59,38 %	0,202	17				40,50 %	0,238	2	35,83 %	0	1

Zdroj: Autoři na základě výpočtu v programu Zelig (Imai a kol. 2007), použítá data lze nalézt na serveru Volby.cz 2011.

#### 4.2 Voličské přesuny ve vybraných obvodech

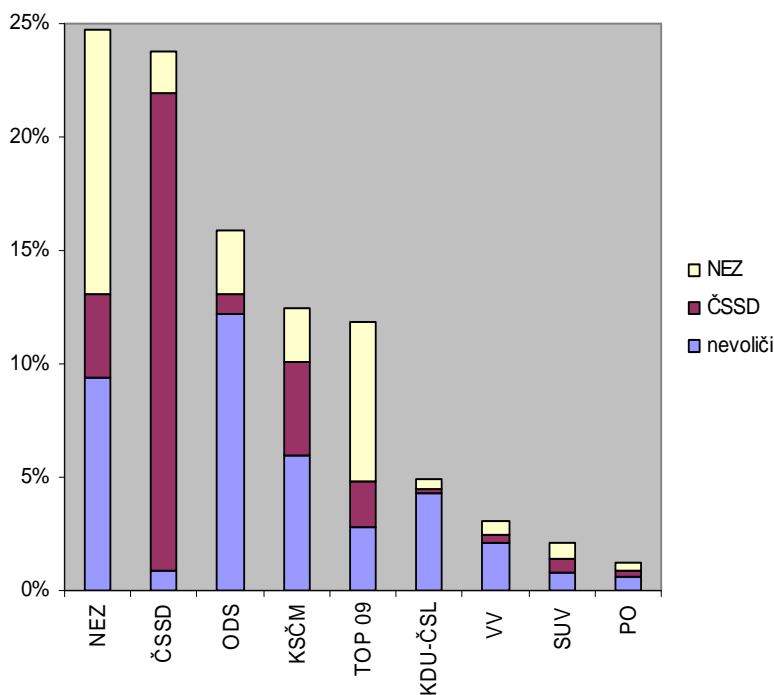
Rozsah této práce bohužel neumožňuje okomentovat výsledky modelu voličských přesunů ve všech 27 volebních obvodech, pojďme se přesto podívat alespoň na několik vybraných obvodů. Jde o obvody Ostrava-město, Nový Jičín, Brno-střed (číslo 58), Most a Ústí nad Labem. Právě na těchto obvodech se dají velmi dobře ilustrovat různé způsoby využití získaných informací o voličských přesunech. V případě obvodu Ostrava-město si můžeme povšimnout anomálního vzoru voličských přesunů, který by mohl být výsledkem medializace volební soutěže mezi oběma koly voleb. Na případě obvodu Nový Jičín vidíme, do jaké míry se voliči řídí doporučením kandidátů vypadlých po prvním kole. V obvodu Brno-střed dokážeme identifikovat voliče, kteří k zisku mandátu pomohli kandidátovi postupujícímu do druhého kola až ze druhého místa. Poslední dva zvolené obvody umožňují porovnat úspěšnost kandidátů regionálního hnutí Severočesí.cz proti kandidátům za ČSSD a za ODS.

Začneme v obvodu Ostrava-město, kde do druhého kola postoupila Liana Janáčková za hnutí Nezávislí, známá kontroverzními výroky namířenými proti Romům. Po prvním kole voleb vyzval předseda vlády Petr Nečas voliče, aby hlasovali proti kandidátům ČSSD, a zabránili tím této straně získat v Senátu většinu mandátů. Explicitně podpořil rovněž L. Janáčkovou. Vzápětí se stal terčem kritiky nejen ze strany opozice, ale i místní organizace ODS v Ostravě.<sup>16</sup> Jak se v této situaci zachovali voliči? Graf č. 1 ukazuje výsledky voleb po prvním kole. Obsah sloupců poskytuje informaci o tom, jak se voliči jednotlivých kandidátů zachovali ve druhém kole. Graf č. 2 podává stejné údaje pro druhé kolo. L. Janáčkovou po prvním kole opustila značná část

<sup>16</sup> Senátorka prý chtěla na Romů použít dynamit 2007, Lumpa Janáčkovou volit nebudeme, odmítá ostravská ODS Nečasovu výzvu 2010, Druhé kolo senátních voleb odstartovala roztržka mezi ODS a ČSSD 2010.

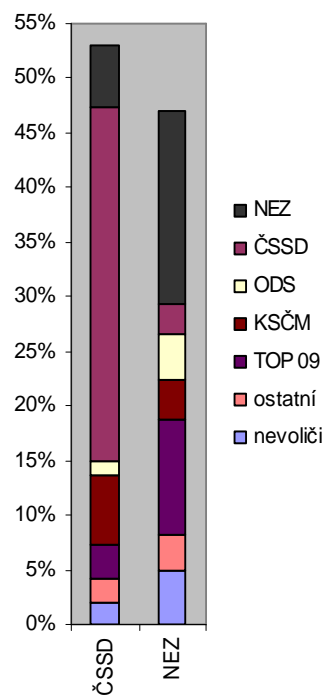
voličů. Většina z nich se druhého kola neúčastnila a část dokonce podpořila protikandidáta Antonína Maštalíře z ČSSD. Jeho vlastní voliči byli ovšem mimořádně disciplinovaní – téměř všichni se účastnili druhého kola a podpořili ho znovu. V obvodu Ostrava-město došlo v roce 2010 mimo jiné i k největšímu „zvratu“ mezi prvním a druhým kolem. Zatímco L. Janáčková v prvním kole těsně zvítězila, o týden později byla poměrně přesvědčivě poražena. A. Maštalíř získal mandát díky disciplinovanosti voličů ČSSD a rovněž díky podpoře části voličů neúspěšné kandidátky KSČM.

**Graf č. 1.** Ostrava-město –  
podíly hlasů v prvním kole  
a volba ve druhém kole



Zdroj: *tamtéž*.

**Graf č. 2.** Ostrava-město –  
podíly hlasů ve druhém kole  
a volba v prvním kole

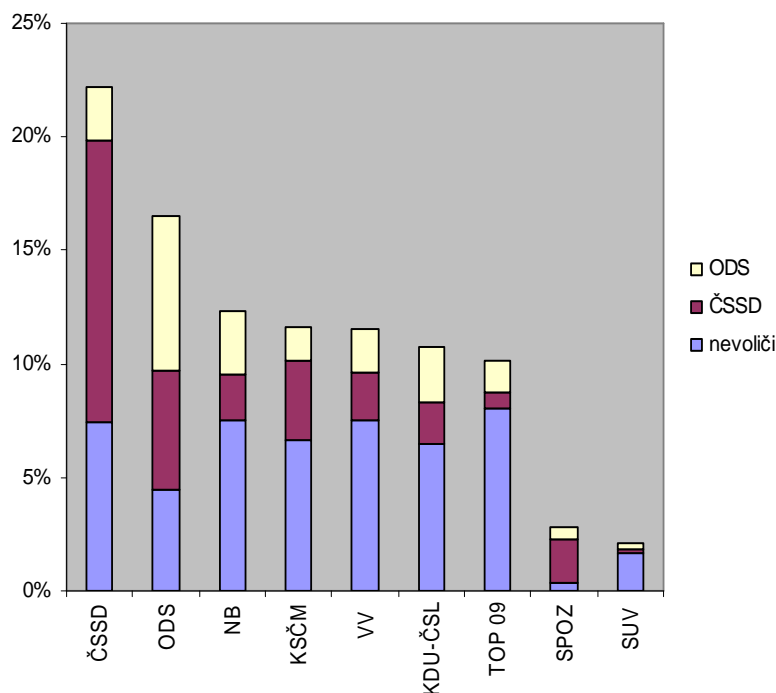


Zdroj: *tamtéž*.

V obvodu Nový Jičín můžeme sledovat, do jaké míry se voliči řídí doporučením svých kandidátů. Již před volbami se zde Milan Bureš z ODS a Drahomír Strnadel kandidující za hnutí Naše Beskydy dohodli na vzájemné podpoře. Pak se ale ukázalo, že Milan Bureš je bývalým členem KSČ. D. Strnadel proto vyzval své voliče, aby dali ve druhém kole hlas kandidátu ČSSD Zdeňku Bestovi. Zajímavá byla i situace u VV. Zatímco strana oficiálně podporovala M. Bureše, sama kandidátka VV Sylva Kováčiková vyjádřila svoji podporu Z. Bestovi (Neúspěšní senátorští kandidáti zmátli finalisty 2010). Zdá se, že většina voličů kandidáta za sdružení Naše Beskydy

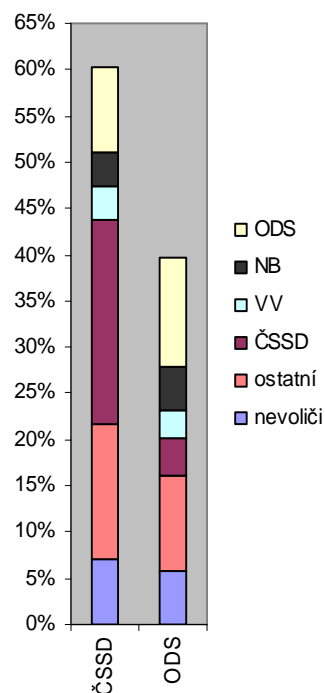
i VV na výzvy svých kandidátů příliš nedbala a druhého kola se neúčastnila. Zbytek se rozdělil víceméně rovnoměrně mezi kandidáty ČSSD a ODS. Za pozornost stojí poměrně vysoká míra přechodu voličů M. Bureše (ODS) k jeho protikandidátovi. Z. Besta si každopádně udržel náskok z prvního kola a získal mandát.

**Graf č. 3.** Nový Jičín –  
 podíly hlasů v prvním kole  
 a volba ve druhém kole



Zdroj: *tamtéž*.

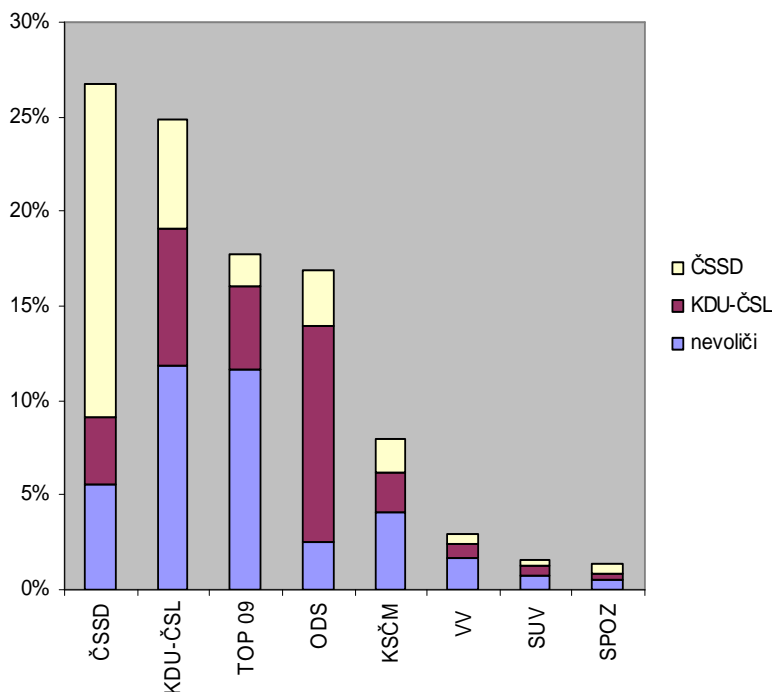
**Graf č. 4.** Nový Jičín –  
 podíly hlasů ve druhém kole  
 a volba v prvním kole



Zdroj: *tamtéž*.

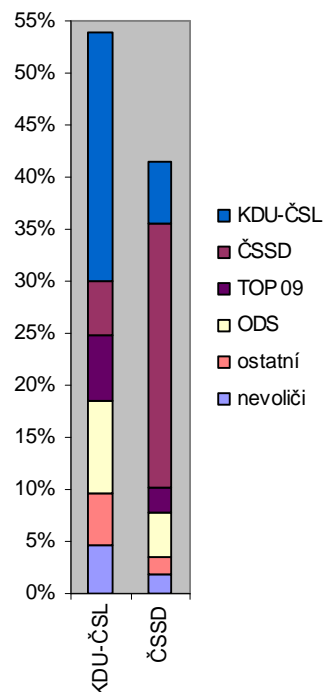
Zajímavá situace nastala i v obvodu Brno-město číslo 58, kde do druhého kola postoupil z prvního místa Václav Božek (ČSSD), mandát ale nakonec získal Stanislav Juránek z KDU-ČSL, a to s poměrně velkým náskokem. Jací voliči mu umožnili „předběhnout“ jeho protikandidáta? Ačkoli mnoho jeho vlastních voličů nepřišlo ve druhém kole k urnám, získání mandátu mu zajistili zejména voliči ODS a TOP 09.

**Graf č. 5.** Brno-střed (obvod číslo 58) –  
 podíly hlasů v prvním kole  
 a volba ve druhém kole



Zdroj: *tamtéž*.

**Graf č. 6.** Brno-střed (obvod číslo 58) –  
 podíly hlasů ve druhém kole  
 a volba v prvním kole

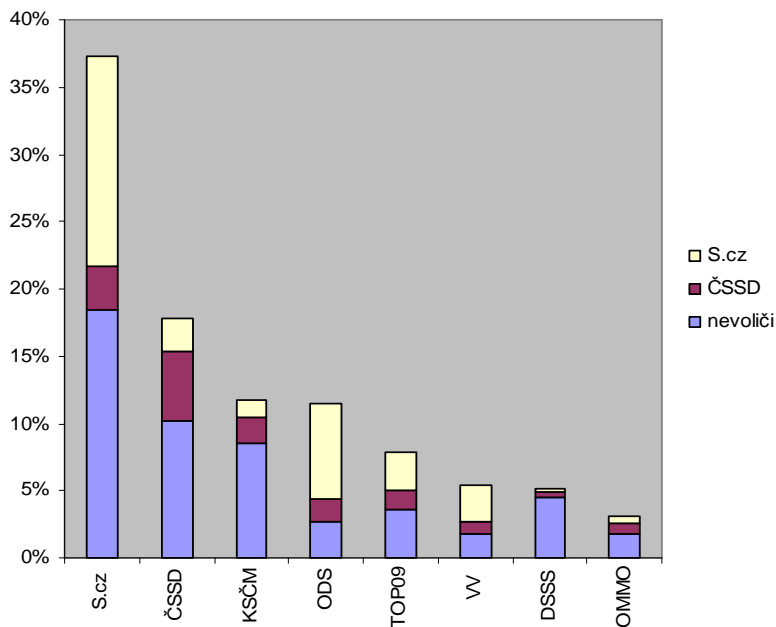


Zdroj: *tamtéž*.

Na závěr srovnáme vzorce voličských přesunů v obvodech Most a Ústí nad Labem, kde byli zvoleni dva senátoři za regionální hnutí Severočesí.cz. Zatímco v Mostě se jejich kandidátka Alena Dernerová utkala s kandidátem ČSSD, v Ústí nad Labem stál Jaroslav Doubrava (S.cz) proti kandidátovi ODS.

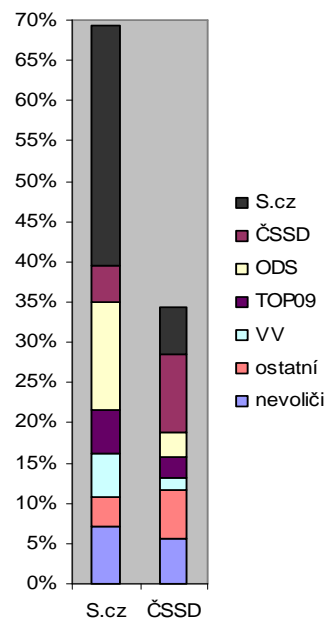
V obou obvodech to totiž byli především voliči neúspěšných kandidátů ČSSD a ODS z prvního kola, kteří podpořili kandidáta S.cz proti jeho soupeři, a pomohli mu tím k zisku mandátu. Je nicméně pravdou, že v obvodu Most byl náskok A. Dernerové velmi vysoký už v prvním kole, takže by mandát pravděpodobně získala i bez této podpory. V obvodu Ústí nad Labem se ovšem tento jev plně projevil, vždyť Jaroslav Doubrava (S.cz) postoupil ze druhého místa a bez podpory voličů ČSSD by mandát zřejmě nezískal.

**Graf č. 7. Most –**  
 podíly hlasů v prvním kole  
 a volba ve druhém kole



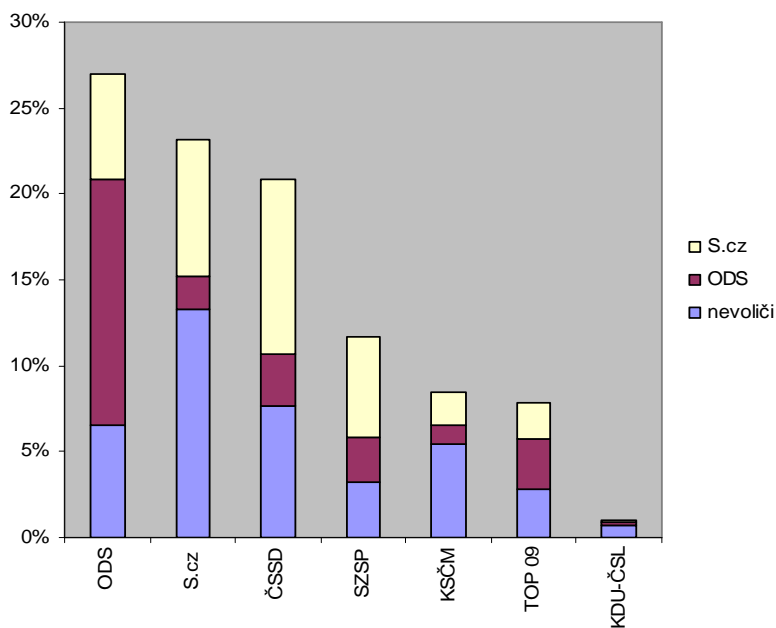
Zdroj: tamtéž.

**Graf č. 8. Most –**  
 podíly hlasů ve druhém kole  
 a volba v prvním kole



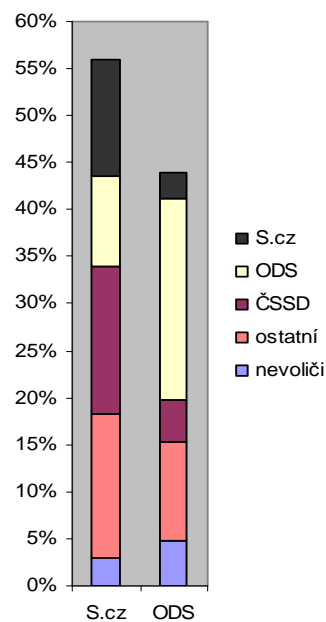
Zdroj: tamtéž.

**Graf č. 9. Ústí nad Labem –**  
 podíly hlasů v prvním kole  
 a volba ve druhém kole



Zdroj: tamtéž.

**Graf č. 10. Ústí nad Labem –**  
 podíly hlasů ve druhém kole  
 a volba v prvním kole



Zdroj: tamtéž.

## 5. Závěr

Cílem této práce bylo statistické modelování přesunů voličské podpory jednotlivých kandidátů mezi prvním a druhým kolem senátních voleb v roce 2010 pomocí bayesovského modelu ekologické inference. Výstupy modelu byly použity k testování čtyř hypotéz týkajících se strategických účinků volebního systému používaného v senátních volbách a použitými daty byly oficiální výsledky voleb na úrovni 4.860 volebních okrsků.

Ekologická inference je metodou odvozování informací o chování jednotlivců nevratně ztracených při procesu agregace dat, a to v situaci, kdy nejsou k dispozici data na individuální úrovni. Práce podala výklad o historickém vývoji této oblasti kvantitativní metodologie od prvotního podchycení problému ekologického klamu v odborné literatuře přes rané metody ekologického usuzování, jako je Goodmanova ekologická regrese či metoda hranic, až po moderní inferenční modely, které využívají bayesovskou pravděpodobnostní matematiku a umožňují modelovat vzájemné vztahy mezi vícekategoriálními proměnnými.

Při modelování voličských přesunů mezi prvním a druhým kolem senátních voleb v roce 2010 bylo zjištěno, že druhého kola se účastnili především voliči, kteří k urnám přišli již v prvním kole. Ukázalo se, že voliči ve druhém kole nejčastěji odevzdávali hlasy stejnému kandidátovi jako v kole prvním, jestliže se mu podařilo do druhého kola postoupit. Pokud tomu tak nebylo, pak voliči ve druhém kole podpořili kandidáty ideologicky podobných stran. Tento vztah je zřetelný zejména u voličů KSČM podporujících ve druhém kole kandidáty ČSSD. Rovněž voliči ODS a TOP 09 dávali ve druhém kole ve velké míře hlas kandidátu jedné z těchto dvou stran, pokud jejich vlastní kandidát nepostoupil.

Bohužel se nepodařilo prověřit hypotézu, že se voliči politických stran jednoho pólu ideologického spektra neúčastní druhého kola, pokud do něj postoupí dva kandidáti z opačného pólu ideologického spektra. Podle tohoto vzoru se ve sledovaných volbách chovali zejména voliči kandidátů KSČM, jejichž účast ve druhém kole výrazně poklesla, pokud do něj nepostoupil kandidát ČSSD. Celkově však senátní volby v roce 2010 nenabízí dostatečný počet vhodných případů pro otestování platnosti této hypotézy.

Problém zásadní statistické neurčitelnosti, jemuž metody ekologické inference čelí, je kvůli nevratnosti ztráty informací při procesu agregace dat nepřekonatelný. Konstrukce dostatečně přesného a robustního modelu by nicméně představovala nesmírně významný milník ve vývoji sociálních věd. Na problém špatné dostupnosti či přímo absence dat na individuální

úrovni totiž sociální vědci naráží téměř při každém výzkumu. Proto je podle našeho názoru velmi důležité rozvíjet toto odvětví kvantitativní metodologie i v českých podmínkách.

### Seznam pramenů a literatury

- Allport, Floyd H. (1924): The Group Fallacy in Relation to Social Science, *American Journal of Sociology*, Vol. XXIX, No. 6, pp. 688-703.
- Černák, Daniel (2010): Ekologická inference a ekologické zkreslení, *Socioweb*, roč. VIII, č. 2, on-line text <<http://www.socioweb.cz/index.php?disp=teorie&shw=427&lst=117>>. [cit. 06. 05. 2011]
- Volební server České statistické úřadu, on-line zdroj <<http://www.volby.cz/>>. [cit. 01. 08. 2011]
- Deming, Edwards W. – Stephan, Frederick F. (1940): On a Least Squares Adjustment of a Sampled Frequency Table When Expected Marginal Totals Are Known, *Annals of Mathematical Statistics*, Vol. XI, No. 4, pp. 427-444.
- Duncan, Otis D. – Davis, Beverly (1953): An Alternative to Ecological Correlation, *American Sociological Review*, Vol. XVIII, No. 6, pp. 665-666.
- Elff, Martin – Gschwend, Thomas – Jonhston, Ron J. (2008): Ignoramus, Ignorabimus? On Uncertainty in Ecological Inference, *Political Analysis*, Vol. XVI, No. 1, pp. 70-92.
- Ferree, Karen (1999): *Iterative Approaches to R x C Ecological Inference Problems: Where They Can Go Wrong*, Working paper, on-line text <<http://polmeth.wustl.edu/media/Paper/ferre99.pdf>>. [cit. 2001-05-08]
- Freedman, David A. – Ostland, Michael – Roberts, Michael R. – Klein, Stephen P. (1999): Response to King's Comment, *Journal of the American Statistical Association*, Vol. XCIV, No. 440, pp. 355-357.
- Gehlke, C. E. – Biehl, Katherine (1934): Certain Effects of Grouping upon Size of the Correlation Coefficient in Census Tract Material, *Journal of the American Statistical Association Supplement*, Vol. XXIX, No. 185, pp. 663-664.
- Gelman, Andrew – Park, David K. – Ansolabehere, Stephen – Price, Phillip N. – Minnite, Lorraine C. (2001): Models, Assumptions and Model Checking in Ecological Regressions, *Journal of the Royal Statistical Society*, Vol. CLXIV, No. 1, pp. 101-108.
- Goodman, Leo (1953): Ecological Regressions and the Behavior of Individuals, *American Sociological Review*, Vol. XVIII, No. 6, pp. 663-666.
- Goodman, Leo (1959): Some Alternatives to Ecological Correlation, *American Journal of Sociology*, Vol. LXIV, No. 6, pp. 610-624.
- Glynn, Adam – Wakefield, Jon (2009): Ecological Inference in the Social Science, *Statistical Methodology*, Vol. VII, No. 3, pp. 307-322.
- Graunt, John (1662): Natural And Political Observations Mentioned in a Following Index, And Made upon the Bills of Morality, London, John Martyn and James Allestry; cit. in: King, Gary (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton, Princeton University Press, p. 4.
- Herron, Michael C. – Sekhon, Jasjeet S. (2005): Black Candidates And Black Voters: Assessing the Impact of Candidate Race on Uncounted Vote Rates, *The Journal of Politics*, Vol. LXVII, No. 1, pp. 154-177.
- Cho, Wendy T. (1998): If the Assumption Fits: A Comment on the King Ecological Inference Solution, *Political Analysis*, Vol. VII, pp. 143-163.

- Cho, Wendy T. – Gaines, Brian J. (2004): The Limits of Ecological Inference: The Case of Split-ticket Voting, *American Journal of Political Science*, Vol. XLVIII, No. 1, pp. 152-171.
- Chytilék, Roman (2005a): České senátní volby. Podněty, výsledky, alternativy, in: Břetislav Dančák – Petr Fiala – Vlastimil Havlík (eds.): *Evropeizace: Nové téma politologického výzkumu*, Brno, Mezinárodní politologický ústav Masarykovy univerzity, s. 105-115.
- Chytilék, Roman (2005b): Volby 2004 a strategické účinky volebních systémů, *Středoevropské politické studie*, roč. VI, č. 1, s. 13-25.
- Chytilék, Roman – Šedo, Jakub – Lebeda, Tomáš – Čaloud, Dalibor (2009): *Volební systémy*, Praha, Portál.
- Chytilék, Roman – Eibl, Otto (2011): České politické strany v politickém prostoru, *Sociologický časopis*, roč. XLVII, č. 1, s. 61-88.
- Imai, Kosuke – King, Gary – Lau, Olivia (2007): *Zelig: Everyone's Statistical Software*, on-line source <<http://gking.harvard.edu/zelig>>. [cit. 06. 05. 2011]
- Johnston, Ron – Gschwend, Thomas – Pattie, Charles (2004): *On Estimates of Split-ticket Voting: EI And Emax*, Working paper, on-line text <<http://polmeth.wustl.edu/media/Paper/jgp04.pdf>>. [cit. 06. 05. 2011]
- King, Gary (1997): *A Solution to the Ecological Inference Problem. Reconstructing Individual Behavior from Aggregate Data*, Princeton, Princeton University Press.
- King, Gary – Rosen, Ori – Tanner, Martin A. (1999): Binomial-Beta Hierarchical Models for Ecological Inference, *Sociological Methods and Research*, Vol. XXVIII, No. 1, pp. 61-90.
- King, Gary – Tanner, Martin A. – Rosen, Ori (2004, eds.): *Ecological Inference: New Methodological Strategies*, Cambridge, The University of Cambridge.
- King, Gary – Rosen, Ori – Tanner, Martin A. – Wagner, Alexander (2008): Ordinary Economic Voting Behavior in the Extraordinary Election of Adolf Hitler, *Journal of Economic History*, Vol. LXVIII, No. 4, pp. 951-956.
- Kopstein, Jeffrey S. – Wittenberg, Jason (2003): Who Voted Communist? Reconsidering the Social Bases of Radicalism in Interwar Poland, *Slavic Review*, Vol. LXII, No. 1, pp. 87-109.
- Kopstein, Jeffrey S. – Wittenberg, Jason (2004): Ethnic Diversity, Democracy, and Electoral Extremism: Lessons from Interwar Poland and Czechoslovakia, Working paper, on-line text <[http://politics.as.nyu.edu/docs/IO/4734/wittenberg\\_s05.pdf](http://politics.as.nyu.edu/docs/IO/4734/wittenberg_s05.pdf)>. [cit. 06. 05. 2011]
- Kopstein, Jeffrey S. – Wittenberg, Jason (2009): Does Familiarity Breed Contempt? Inter-ethnic Contact And Support for Illiberal Parties, *The Journal of Politics*, Vol. LXXI, No. 2, pp. 414-428.
- Kopstein, Jeffrey S. – Wittenberg, Jason (2010a): Beyond Dictatorship And Democracy: Rethinking National Minority Inclusion and Regime Type in Interwar Eastern Europe, *Comparative Political Studies*, Vol. XLIII, No. 8-9, pp. 1089-1118.
- Kopstein, Jeffrey S. – Wittenberg, Jason (2010b): Deadly Communities: Local Political Milieus And the Persecution of Jews in Occupied Poland, *Comparative Political Studies*, Vol. XLIV, No. 3, pp. 259-283.
- Kopstein, Jeffrey S. – Wittenberg, Jason (2011a): *Between State Loyalty and National Identity. Electoral Behaviour in Inter-war Poland*, Working paper, on-line text <<http://witty.berkeley.edu/Proofs%20Polin%2024.pdf>>. [cit. 06. 05. 2011]
- Kopstein, Jeffrey S. – Wittenberg, Jason (2011b): *Did Ethnic Balance Matter? Elections in Interwar Poland*, Working paper, on-line text



- <[http://individual.utoronto.ca/kopstein/journals/Between\\_Nationalization\\_and\\_Internationalization.pdf](http://individual.utoronto.ca/kopstein/journals/Between_Nationalization_and_Internationalization.pdf)>. [cit. 06. 05. 2011]
- Lau, Olivia – Moore, Ryan. T. – Kellermann, Michael (2007): EiPack: R x C Ecological Inference And Higher-dimension Data Management, *R News*, Vol. VII, No. 2, pp. 43-47.
- Lebeda, Tomáš – Vlachová, Klára – Řeháková, Blanka (2009): *První volby do Senátu: analýza voleb do Senátu Parlamentu České republiky v roce 1996*, Praha, Sociologický ústav Akademie věd České republiky.
- Leemann, Lucas – Leimgruber, Philipp (2009): *Ecological Inference and 113 Votes*, Conference Swiss Political Science Association, St. Gallen.
- Lumpa Janáčkovou volit nebudeme, odmítá ostravská ODS Nečasovu výzvu, 2010, iDNES, on-line text <<http://zpravy.idnes.cz/>>. [cit. 06. 05. 2011]
- Lyons, Pat (2008): Ekologické usuzování: Explorace metody latentní struktury za využití volebních dat z ČR, *Data a výzkum – SDA Info*, roč. 2, č. 1, s. 49-75.
- Malcová, Karolína (2010): Analýza senátních voleb pohledem mikroúrovně vybraných obvodů neúspěšných pro ODS, *Politologická revue*, roč. V, č. 2, s. 77-101.
- Malíř, Jirí – Marek, Pavel (2005, eds.): *Politické strany II. 1938–2004*, Brno, Doplněk.
- Matějů, Petr – Vlachová Klára (1997): Volby do Senátu. Voliči, strany, osobnosti a ideologie, *Parlamentní zpravodaj*, č. 4-5; cit. in: Tomáš Lebeda – Klára Vlachová – Blanka Řeháková (2009): *První volby do Senátu: analýza voleb do Senátu Parlamentu České republiky v roce 1996*, Praha, Sociologický ústav Akademie věd České republiky, s. 12.
- Navrátil, Vojtěch (2010): Výsledky druhého kola senátních voleb 1996–2008: Kdo, proti komu, odkud a jak, *Středoevropské politické studie*, roč. VII, č. 1.
- O'Loughlin, John (2002): The Electoral Geography of Weimar Germany: Exploratory Spatial Data Analyses (ESDA) of Protestant Support for the Nazi Party, *Political Analysis*, Vol. X, No. 3, pp. 217-243.
- Ogburn, William – Goltra, Inez (1919): How Women Vote: A Study of an Election in Portland, Oregon, *Political Science Quarterly*, Vol. XXXIV, No. 3, pp. 413-433.
- Park, Won-ho (2004): *Estimation of Voter Transition Rates and Ecological Inference*, Conference Western Political Science Association, Portland, Oregon.
- Park, Won-ho (2008): *Ecological Inference and Aggregate Analysis of Elections*, Dissertation Thesis, Ann Arbor, University of Michigan.
- Petty, William (1690): *Political Arithmetic*, London, Robert Clavel and Henry Mortlock; cit. in: Gary King (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton, Princeton University Press, p. 4.
- Petty, William (1691): *The Political Anatomy of Ireland*, London, D. Brown and W. Rogers; cit. in: Gary King (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton, Princeton University Press, p. 4.
- Robinson, William (1950): Ecological Correlation And the Behavior of Individuals, *American Sociological Review*, Vol. XV, No. 3, pp. 351-357; cit. in: Pat Lyons (2008): Ekologické usuzování: Explorace metody latentní struktury za využití volebních dat z ČR, *Data a výzkum – SDA Info*, roč. II, č. 1, s. 50.
- Rosen, Ori – Jiang, Wenxin – King, Gary – Tanner, Martin A. (2001): Bayesian And Frequentist Inference for Ecological Inference: The R x C Case, *Statistica Neerlandica*, Vol. LV, No. 2, pp. 134-156.

- Schuessler, Alexander A. (1999): Ecological Inference, *Proceedings of the National Academy of Science*, Vol. No. 96, pp. 10578-10581.
- Senátorka prý chtěla na Romy použít dynamit*, 2007, iDNES, on-line text <<http://zpravy.idnes.cz/>>. [cit. 06. 05. 2011]
- Mattos, de Silva – Veiga, Álvaro (2002): *The Binomial-beta Hierarchical Model for Ecological Inference: Methodological Issues and Fast Implementation via the ECM Algorithm*, Working paper, on-line text <<http://polmeth.wustl.edu/media/Paper/matto01.pdf>>. [cit. 06. 05. 2011]
- Simpson, Edward H. (1951): The Interpretation of Interactions in Contingency Tables, *Journal of the Royal Statistical Society*, Vol. XIII, pp. 238-241; cit. in: Adam Glynn – Jon Wakefield (2009): Ecological Inference in the Social Sciences, *Statistical Methodology*, Vol. VII, No. 3, p. 310.
- Thronthike, E. L. (1939): On the Fallacy of Imputation the Correlations Found for Groups to the Individuals or Smaller Groups Composing Them, *American Journal of Psychology*, Vol. LII, pp. 122-124.
- Tomz, Michael – Tucker, Joshua A. – Wittenberg, Jason (2002): An Easy and Accurate Regression Model for Multiparty Electoral Data, *Political Analysis*, Vol. X, No. 1, pp. 66-83.
- Wakefield, Jon (2008): Ecologic Studies Revisited, *Annual Review of Public Health*, Vol. XXIX, No. 1, pp. 75-90.
- Yule, G. Undy – Kendall, Maurice (1950): *An Introduction to the Theory of Statistics*; cit. in: Gary King (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton, Princeton University Press, p. 4.