

Online příloha 1.

Některá další zjištění z analýzy

Užívání statistických technik v Sociologickém časopise

Pro zajímavost (ocenit to jistě historici české sociologie) v Tabulce T1¹ uvádíme, kdy poprvé bylo v Sociologickém časopise použito jednotlivých složitějších statistických technik v rámci sledovaného období posledních dvaceti let.

Z přehledu je zřejmé, že většina složitějších statistických technik (např. strukturní modely a logistická regrese) byla používána již v devadesátých letech.² Relativní novinkou je užívání hierarchických lineárních modelů (2006) a analýzy latentních tříd (2007). Následující Tabulka T2 pak sleduje, k jakým změnám v užívání složitějších statistických technik docházelo na stránkách Sociologického časopisu v čase.

Tabulka T1. Rok, kdy byla poprvé v Sociologickém časopise použita statistická technika v období 1995–2014

Technika	Rok prvního užití
nelineární regrese (zejm. logistická)	1995
konfirmační faktorová analýza	1995
strukturní modely	1995
loglineární modely	1997
logitové modely	1998
korespondenční analýza	1998
explorační faktorová analýza	1999
diskriminační analýza	1999
seskupovací analýza	2001
mnohorozměrné škálování	2004
hierarchické lineární modely	2006
analýza latentních tříd	2007

¹ Pro odlišení jsou tabulky a grafy v příloze číslovány jako T1, T2, respektive G1, G2 atd.

² V mnoha případech ještě před rokem 1995, ostatně například první metodologický text věnovaný strukturnímu modelování otiskl Sociologický časopis již v roce 1989 [Matějů 1989].

Tabulka T2. Vývoj užívanosti složitějších statistických technik v Sociologickém časopise (1995–2014) dle pětiletých období (procento užití technik v kvantitativně orientovaných článcích)

Technika	1995–1999	2000–2004	2005–2009	2010–2014
korelace	20,6	13,2	15,2	0,0
regrese lineární	23,5	23,7	23,9	22,7
regrese nelineární	17,6	7,9	34,8	27,3
EFA	5,9	10,5	13,0	6,8
CFA	5,9	0,0	0,0	0,0
SEM	14,7	2,6	8,7	13,6
mnohor. škálování	0,0	2,6	2,2	0,0
procenta	38,2	18,4	6,5	15,9
průměr	2,9	0,0	2,2	0,0
analýza rozptylu	11,8	7,9	8,7	2,3
t-test	8,8	10,5	0,0	9,1
loglineární modely	2,9	0,0	2,2	4,5
kontin. tabulky	8,8	21,1	10,9	15,9
diskr. analýza	2,9	2,6	2,2	0,0
logitové modely	5,9	15,8	0,0	0,0
koresp. analýza	2,9	2,6	0,0	0,0
seskupování	0,0	5,3	0,0	13,6
analýza přežití	0,0	2,6	0,0	0,0
hier. modely	0,0	0,0	8,7	11,4
latentní třídy	0,0	0,0	2,2	0,0
jiná	0,0	0,0	0,0	2,3

Poznámka: EFA = explorační faktorová analýza, CFA = konfirmační faktorová analýza, SEM = strukturní modelování.

Pokud se zaměříme na nejpoužívanější techniky, pak lze vysledovat určité tendence. Používání korelací v posledních pěti letech vymizelo (obdobná je situace u analýzy rozptylu a procent), lineární regrese se užívá zhruba stále ve stejné míře. Nelineární regresní modely (převážně binární logistická regrese) se v posledním desetiletí užívá častěji než v předchozím desetiletí.

U strukturních modelů lze vypořizovat jistou renesanci v užívání (poměrně často se užívaly v druhé polovině devadesátých let a objevují se častěji i v poslední sledované pětiletce). Další zajímavostí je vymizení logitových modelů v posledních sledovaných deseti letech, což lze snadno vysvětlit ukončením

publikační aktivity Blanky Řehákové, která tuto techniku v české sociologii jako jediná užívala. Celkově nelze konstatovat, že by docházelo k nárůstu nebo poklesu v užívání složitějších statistických technik, a lze hovořit spíše o stagnaci. Lze tak uzavřít, že se v průběhu posledních dvaceti let mnoho nového v oblasti technik používaných v Sociologickém časopise neobjevilo.

Institucionální kontext české vědy a možné dopady na změny kvality kvantitativních analýz

V rámci vlivu institucionálního kontextu české vědy má smysl se zaměřit na dvě oblasti. První je výrazná změna, která nastala ve financování vědy zejména s ukončením výzkumných záměrů a zavedením financování dle bodů přidělených vědeckým výsledkům zavedeným do databáze RIV (srov. popis dále). S ohledem na zvýšený tlak na publikování skrze tento systém financování by bylo možné očekávat výrazný nárůst vědecké produkce a snížení její kvality.

Druhým sledovaným fenoménem bude vliv publikačních standardů, které nastavují oborové asociace, respektive jednotlivé časopisy. V oblasti našeho zkoumání (Československá psychologie, Pedagogika a Sociologický časopis) nedošlo ve sledovaném období k žádné aktivitě příslušných českých (československých) oborových asociací. Psychologové mohli respektovat doporučení APA (5. manuál z roku 2001, respektive 6. manuál z roku 2010), pedagogové pak doporučení AERA (z roku 2006). Kromě toho vyšly v časopise Československá psychologie dva články Urbánka [2007, 2008] a posléze tento časopis zveřejnil též metodologické požadavky na výzkumné studie, které na zmíněné dva články obsahově navazují. Bylo by tedy možné předpokládat, že zejména v Československé psychologii bude cca roku 2009 docházet ke zlepšování kvality kvantitativní produkce.

Před prezentací závěrů stručně připomeňme historii hodnocení a financování vědy v ČR. Nejvýraznější změnu přinesl dosud platný zákon č. 130/2002 Sb., o podpoře výzkumu, experimentálního vývoje a inovací z veřejných prostředků a o změně některých souvisejících zákonů (zákon o podpoře výzkumu, experimentálního vývoje a inovací), který v návaznosti na Rejstřík informací o výsledcích (dále jen „RIV“) zakotvil základní principy hodnocení a financování české vědy. Instituce vykazují své vědecké a výzkumné výsledky skrze RIV a dle těchto výsledků (zpravidla několika let zpětně) získávají finanční prostředky. Metodika se mezi jednotlivými roky zpravidla proměňovala, ale princip byl v zásadě stejný. Každý výsledek vědy a výzkumu byl obodován a dle počtu získaných bodů získala příslušná instituce finanční podporu. Detailnější parametry nastavilo Usnesení vlády ze dne 23. června 2004 č. 644 k hodnocení výzkumu a vývoje a dále dokument nazvaný Reforma systému výzkumu, vývoje a inovací v ČR, schválený usnesením vlády pod č. 287 dne 26. března 2008. Obecně platí, že zhruba od roku 2008 začaly výzkumné a vědecké instituce v ČR výrazně sledovat své vědecké a výzkumné výsledky a začal mj. sílit tlak na publikování. Dodejme, že dostupná analýza výsledků vědy a výzkumu v ČR, zpracovaná Součkem [2012], konstatuje,

že nejvíce bodů bylo získáno za články v časopisech s impakt faktorem (cca 60 % v letech 2006–2010), nezanedbatelný byl i podíl bodů za články v recenzovaných časopisech (cca 15 %). Pro úplnost dodejme, že podíl knih či kapitol v knize na celkovém počtu hodnocených výsledků byl lehce pod 10 %. S ohledem na skutečnost, že v oblasti sociálních a příbuzných věd jsou jiné než časopisecké či knižní výstupy velice nepravděpodobným výstupem vědy a výzkumu, lze *cum grano salis* konstatovat, že v případě sociálních a příbuzných věd směřoval po zavedení popsaného systému hodnocení největší tlak právě na publikování v časopisech, zejména těch s impakt faktorem, protože zde se zpravidla získávalo více bodů. Dále je nutno poznamenat, že mnohé instituce převzaly systém hodnocení institucí na základě získaných bodů interně a začaly dle těchto kritérií hodnotit jednotlivá pracoviště, případně i jednotlivé pracovníky. Tím logicky tlak na publikování ještě zesílil. Lze si položit otázku, zda tento tlak mohl mít negativní dopady na publikování, které začalo být méně kvalitní. Samozřejmě je též otázkou, zda lze tyto skutečnosti prokázat na provedeném výzkumu tří předních českých časopisů. Cílem tohoto exkurzu bude tedy sledovat vývoj kvantitativně laděné produkce v Československé psychologii, Pedagogice a Sociologickém časopise v čase z hledisek, která byla rozebrána v předchozích analýzách (využívání statistických testů pro nevhodná data, mechanické či jiné nevhodné interpretace získaných výsledků, věcná interpretace výsledků a užívání měr věcné významnosti). Na počátku je nutné otevřeně přiznat, že tato analýza rozhodně nepodá komplexní odpověď na otázku dopadů systému hodnocení a financování vědy, ale může poukázat na některé efekty.

Pokud by mělo platit, že změna financování české vědy započatá v roce 2009 změnila strategie publikování vědců, pak by měl být ve třech sledovaných časopisech přinejmenším patrný nárůst produkce (1) a zhoršení kvality článků (2).

Ad 1. Z Grafu 2 v článku plyne, že minimálně v oblasti kvantitativně orientovaných článků se od roku 2009 žádný očekávaný boom neodehrál, naopak v případě Československé psychologie je patrný setrvalý pokles počtu článků od roku 2010. Samozřejmě lze nalézt i alternativní vysvětlení, které nahrává tezi o vlivu nového systému financování české vědy. Toto vysvětlení by říkalo, že schopnější autoři začali více publikovat v zahraničí a přestali publikovat v českých časopisech, protože zahraniční publikace byly více bodované (na základě sledování produkce tří českých časopisů je tato teze neověřitelná).

Ad 2. Zajímavější je pohled na proměnu kvality kvantitativní produkce ve třech sledovaných časopisech. Grafy 5, 8, 9 a 13 (v článku) ukazují vývoj nesprávného používání statistických testů, mechanické aplikace statistických testů, záměny statistické a věcné významnosti a používání měr věcné významnosti. V žádné ze zkoumaných oblastí nelze od roku 2009 (ve srovnání s předchozím pětiletým obdobím) vysledovat systematické zhoršení a nelze tedy konstatovat, že by změna systému financování vědy v ČR měla přímý a uvedenými nástroji měřitelný negativní dopad na kvalitu kvantitativní vědecké produkce publikované ve třech zkoumaných časopisech.

Naopak je patrný pokles mechanické aplikace statistických testů a zejména pak nižší míra absence měr věcné významnosti v Československé psychologii. Samozřejmě je možné se přít, zda je to způsobeno publikovanými manuály APA, nebo články Urbánka [2007, 2008] a doprovodnými metodologickými požadavky na výzkumné studie. Ve srovnání s dalšími dvěma časopisy (v zásadě beze změn) je ovšem zřejmé, že v časopise Československá psychologie dochází ke zlepšování v dodržování standardů kvantitativní produkce, chtělo by se konstatovat, novému systému financování vědy navzdory.

Vliv autorství na kvalitu kvantitativních analýz

Kromě pokusu zmapovat dopad kontextu české vědy na užívání statistické a věcné významnosti v kvantitativní produkci v českých sociálních a příbuzných vědách je možné na základě provedeného výzkumu časopisů zaměřit pozornost na vliv jedinců (autorů). Pro možnost detailnějšího vhledu se zde omezíme jen na výsledky publikované v Sociologickém časopise. Možnosti výzkumných otázek jsou zde mnohé, nicméně omezíme se jen na dvě, s tím, že první je inspirována výzkumným zjištěním Ziliaka a McCloskey [2008] a další pak plyne z výsledků v této práci již dříve prezentovaných analýz. Konkrétní znění výzkumných otázek je následující:

1. Ovlivňuje počet autorů (spoluautorů) kvalitu zpracování kvantitativního výzkumu a kvalitu prezentace a interpretace výsledků (z pohledu využití koncepcí statistické a věcné významnosti)?
2. Liší se kvalita zpracování kvantitativního výzkumu a kvalita prezentace a interpretace výsledků (z pohledu využití koncepcí statistické a věcné významnosti) dle toho, zda autor v oboru začíná (juniovní pracovník), nebo zda jde o již zkušeného autora (seniorní pracovník)?

K nabídnutým výzkumným otázkám je vhodné doplnit podrobnější vysvětlení.

Ad 1. Ziliak a McCloskey [2008] na základě výzkumu článků publikovaných v *American Economic Review* ukázali, že texty, které mají více autorů, jsou z hlediska statistického zpracování horší, tj. autoři se dopouštějí více chyb. Jako nejméně problematické tak byly vyhodnoceny texty, které měly autora jediného. Analogicky tedy bude provedeno srovnání kvality statistického zpracování článků publikovaných v Sociologickém časopise v posledních dvaceti letech dle počtu autorů.

Ad 2. Odlišnost publikační práce juniorních a seniorních pracovníků by mohla pomoci vysvětlit nárůst počtu nesprávných či nevhodných používání statistické významnosti v průběhu sledovaného období (viz popisy výše). Lze očekávat, že juniorní pracovník, který vstřebal poučky kvantitativní metodologie a statistiky na počátku své vědecké kariéry, ještě postrádá potřebný odstup a uplatňuje tak striktně postup dle učebnicových pouček, které ovšem zpravidla

nejsou určeny pro přímé vědecké využití. Teprve v průběhu kariéry s nárůstem zkušeností (zejména skrze četbu odborných textů a jejich oponování v recenzních řízeních) dochází k získání potřebného nadhledu a zažití postupů *de lege artis*. Pro účely analýzy rozdílů mezi juniorními a seniorními pracovníky byla určena mez 35 let věku (vychází z hranice, která je v české vědě běžně užívána). S ohledem na trvání recenzního řízení tak byl výzkumník do 36. roku svého věku v roce publikace článku označen jako juniorní a poté již jako seniorní.

Začneme s první zkoumanou otázkou, tj. zda ovlivňuje počet (spolu)autorů kvalitu zpracování kvantitativního výzkumu a kvalitu prezentace a interpretace výsledků (z pohledu využití koncepcí statistické a věcné významnosti). Pro tyto účely byly zjednodušeny dříve uvedené kategorie a bylo zkoumáno:

- a) zda data, která autor užívá, umožňují zpracování skrze statistické testy,
- b) zda autor využívá v článku minimálně v jedné analýze mechanické aplikace statistických testů,
- c) zda autor v článku minimálně v jedné analýze zaměřuje slovně či fakticky věcnou a statistickou významnost výsledků,
- d) zda autor neopomněl věcně interpretovat minimálně u jedné analýzy věcně získané výsledky a
- e) zda autor neopomněl minimálně u jedné analýzy použít některou z měr věcné významnosti.

Ze 162 analyzovaných článků Sociologického časopisu vydaných v letech 1995–2014 měla celá polovina jednoho autora, 37 % dva spoluautoři a zbývajících 13 % tři a více autorů (více než tři autory jen 3,5 %, maximální počet byl šest spoluautorů u jediného textu). S ohledem na toto rozložení jsou pro analýzu dat použity tři kategorie: jeden autor, dva spoluautoři a tři a více spoluautorů. Toto rozložení zajišťuje, že i v nejméně početné kategorii článků se třemi a více spoluautory je k dispozici dostatečný výzkumný soubor (21 článků). Odpověď na otázku č. 1 nabízí souhrnně Tabulka T3.

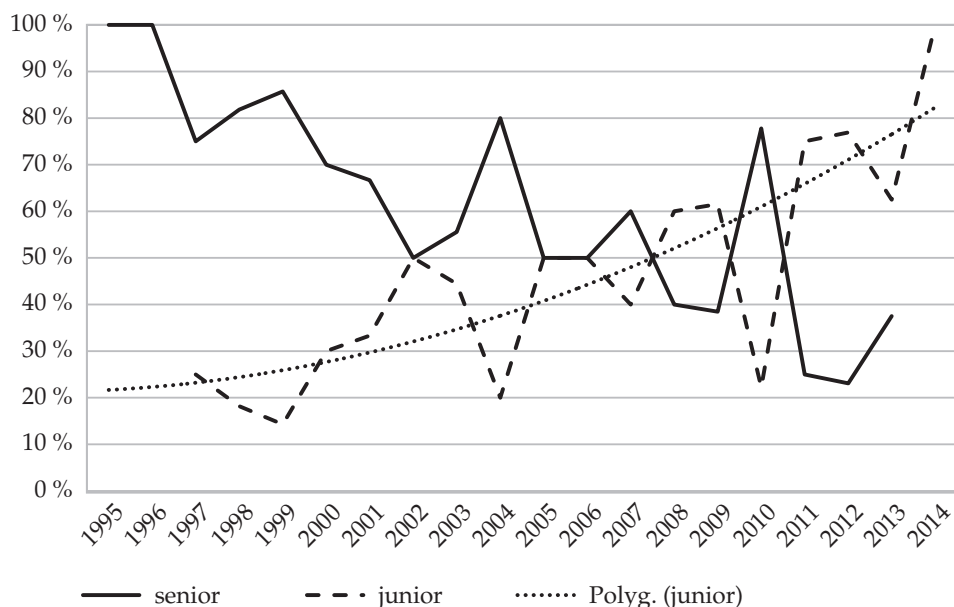
Tabulka T3. Problémy při využívání statistické a věcné významnosti dle počtu (spolu)autorů v Sociologickém časopise v letech 1995–2014 (v %)

Počet (spolu)autorů/fenomén	1	2	3 a více
Použití st. testů pro nevhodná data	19,8	23,3	33,3
Mechanická aplikace st. testů (min. 1× v článku)	49,4	36,7	23,8
Záměna věcné a statistické významnosti (min. 1× v článku)	8,6	5,0	0
Chybějící věcná interpretace (min. 1× v článku)	17,3	5,0	0
Chybějící míra věcné významnosti (min. 1× v článku)	38,3	40,0	19,0

Připomeňme, že dle výsledků Ziliaka a McCloskey [2008] v případě, že publikuje více autorů, je pravděpodobnější chybné využití koncepce statistické či věcné významnosti. Ziliak a McCloskey to odůvodňují tím, že pokud publikuje více autorů společně, pak každý spoléhá na druhého a nepřijímá dostatek vlastní odpovědnosti (jde o *ex post facto* vysvětlení, které nemá žádnou oporu v jejich empirickém materiálu). Když nahlédneme na výsledky platné pro posledních dvacet let v Sociologickém časopise, lze závěry Ziliaka a McCloskey zpochybnit, respektive mnohé výsledky v Tabulce T3 jim přímo protirečí. Jediný fenomén, kde platí, že publikuje-li více autorů, pak je větší riziko nesprávného užití statistické významnosti, je patrný u používání statistických testů pro data, která toto neumožňují. U textů se třemi a více spoluautory je tento fenomén přítomen ve třetině případů, pokud byl autor jediný, pak jen v pětině. Nutno ale poznamenat, že při detailnějším vzhledu (není zobrazen v Tabulce T3) platí, že bez ohledu na počet autorů je podíl textů, kde je využito testů pro zpracování dat z nenáhodných výběrů, v zásadě vyrovnaný. U více spoluautorů je pouze častější tendence užívat statistické testy pro data, která pocházejí z censu. S ohledem na diskusi uvedenou výše v příslušném exkurzu je tato praxe některými statistiky a metodology doporučována a nelze ji zcela odmítnout. U všech ostatních sledovaných fenoménů platí, že v textech, kde je více autorů (zejména tři a více) se s nimi setkáváme méně často. Nejvýraznější rozdíly jsou u mechanické aplikace statistických testů a u chybějících měr věcné významnosti. Proti tezi Ziliaka a McCloskey o neodpovědnosti spoluautorů můžeme tedy postavit konkurenční tezi o pozitivním efektu spoluautorství z hlediska užívání statistické a věcné významnosti (lidová moudrost by velela užít přísloví: „Více očí více vidí.“). Samozřejmě definitivní rozhodnutí, která z tezí o vlivu spoluautorství je pravdivá, by poskytl obsáhlý reprezentativní výzkum vědeckých publikací napříč vědními obory. Je docela dobře možné, že v některých vědních oborech může být platná teze o pozitivním efektu a jinde o negativním efektu vícečetného autorství.

Analogicky k analýze vlivu počtu autorů bude provedena analýza rozdílů mezi juniorními a seniorními výzkumníky. Budou použity stejné analytické kategorie, není tedy třeba tyto opakovat. Je zajímavé nahlédnout podíl kvantitativní produkce, která byla napsána juniorními, respektive seniorními pracovníky (u textů s více autory bylo využito zařazení dle hlavního autora článku), a zejména vývoj tohoto podílu v čase. Základní přehled nabízí Graf G1. Pro lepší nahlédnutí trendů byla přidána trendová křivka (kvadratický trend) pro podíl textů juniorních výzkumníků. Když odhlédneme od extrémů (v roce 1995 nebyl žádný text od juniorního výzkumníka a v roce 2014, kdy bylo textů velice málo, nebyl žádný od seniorního výzkumníka), je patrný poměrně jasný trend nárůstu podílu textů od juniorních výzkumníků a logický pokles podílu seniorních výzkumníků. Vysvětlení (bez datové opory) jsou v zásadě dvě. Jednak během sledovaného období dochází v české sociologii ke střídání generací a střední věk téměř chybí (po roce 1990 zažila česká sociologie poměrně výrazný odliv, kdy pracovníci nejruznějších státních výzkumných institucí buď zakládali soukromé

Graf G1. Vývoj podílu seniorních a juniorních hlavních autorů kvantitativních článků v Sociologickém časopise v letech 1995–2014



výzkumné agentury, nebo do nich odcházeli pracovat). Druhé možné vysvětlení je internacionalizace české sociologie, díky níž seniorní pracovníci začali publikovat v zahraničních časopisech a ponechali prostor v Sociologickém časopise mladším autorům.

Na první pohled nejvýraznější rozdíl nalezneme u používání statistických testů pro nevhodná data – ve třetině textů juniorních výzkumníků byly takto testy použity, u seniorních výzkumníků činil podíl jen 17 %. Za pozornost stojí ještě rozdíl u mechanické aplikace statistických testů, která je také častější v textech juniorních výzkumníků. Lze tedy vyslovit tezi o okouzlení statistickými testy u juniorních výzkumníků. Ostatní rozdíly již nejsou věcně tak výrazné (maximálně 5 procentních bodů). Zajímavostí je (zde lze usuzovat na pozitivní vliv nedávného vzdělávání juniorních výzkumníků), že v textech juniorních výzkumníků méně často absentují míry věcné významnosti pro vypočtené výsledky. Téma rozdílů mezi výzkumníky a vlivu vzdělání (jak během studia, tak poté) bude námětem dalšího výzkumu autora práce. Pozornost bude zaměřena na znalosti výzkumníků z oblasti statistické významnosti a na proces přípravy budoucích výzkumníků na vysokých školách (kurikula, učebnice a vyučující).

Meritem analýzy je ovšem skutečnost, zda se liší užívání statistické a věcné významnosti seniorních a juniorních výzkumníků. Odpověď nabízí Tabulka T4.

Tabulka T4. Problémy při využívání statistické a věcné významnosti dle věku hlavního autora v Sociologickém časopise v letech 1995–2014 (v %)

Hlavní autor/fenomén	Juniorní	Seniorní
Použití st. testů pro nevhodná data	31,9	17,2
Mechanická aplikace st. testů (min. 1× v článku)	46,4	37,6
Záměna věcné a statistické významnosti (min. 1× v článku)	7,2	5,4
Chybějící věcná interpretace (min. 1× v článku)	13,0	8,6
Chybějící míra věcné významnosti (min. 1× v článku)	33,3	38,7