

Honba za vyšší návratností. Stojí opravdu za to? Chasing Higher Response Rate. Is It Really Worth It? Jiří Remr*

Abstrakt

Příspěvek obrací pozornost k míře návratnosti, jež představuje významnou charakteristiku mnohých evaluačních a výzkumných šetření. Příspěvek kromě samotného vymezení pojmu nabízí souhrn možných nápravných opatření, pokud návratnost nedosahuje očekávané míry, a shrnuje možné úpravy designu šetření směřující ke zvýšení návratnosti. Hlavní pozornost se přitom soustřeďuje na dodatečné výzvy k účasti na výzkumu. Jádrem statě je podrobný rozbor účinků dodatečných výzev nejen na samotnou míru návratnosti, ale také na strukturu zkoumaného vzorku formovanou dodatečnými výzvami, na vychýlení výzkumných odhadů, ke kterým může dojít v důsledku použití dodatečných výzev, a v neposlední řadě také na charakter klíčových zjištění, jež může být dodatečnými výzvami ovlivněn. V textu jsou jednotlivé závěry založeny na empirických datech dokumentujících spokojenost s informačním systémem BENEFIT7. Data pocházejí z výzkumného šetření, jež provedl Institut evaluací a sociálních analýz metodou CAWI.

Abstract

Text turns the attention to nonresponse that is the one of the key characteristic of both evaluation and research inquiries. Besides the review of most important definitions of nonresponse, an overview of selected approaches for improvement the response rate is presented. Moreover, the possible changes to research design are introduced and discussed. However,

the main attention is paid to follow-ups, their usefulness and impact on response rate and bias of the research estimates. Text also demonstrates how the follow-ups can influence the structure of the research sample and how it affects the character of achieved results. Specific attention is paid to answering the question of optimal number of follow-up waves. Presented conclusions are based on empirical data describing the satisfaction with information system BENFIT7; such data were collected by Institute for evaluations and social analyses with the use of CAWI.

Klíčová slova

Návratnost, vychýlení, reprezentativita, CAWI, dodatečné výzvy

Keywords

Response rate, bias; representativity, CAWI, follow-ups

1. Diskuse o návratnosti

Návratnost či míra návratnosti (*response rate*) představuje jednu z důležitých charakteristik výzkumných šetření, která odráží situaci, kdy nejsou k dispozici údaje od všech subjektů zahrnutých do výběrového souboru. Obvykle se rozlišuje *unit nonresponse* (tj. výpadek návratnosti¹ u celého případu) a *item nonresponse* (výpadek návratnosti u jednotlivé otázky). Unit nonresponse tedy dokumentuje stav, kdy se některé ze subjektů do výzkumu vůbec nezapojily, zatímco item nonresponse reflektuje skutečnost, kdy někteří respondenti neodpoví na danou konkrétní otázku. V této stati je pozornost věnována výpadku návratnosti u celého případu, tedy unit nonresponse.

Mnohé studie (de Leeuw, de Heer 2002; Atrostic et al. 1999; Groves, Couper 1998) prokázaly, že se míra návratnosti v posledních desetiletích snižuje. Za hlavní důvody tohoto poklesu jsou považovány zejména obavy obyvatel ze zneužití osobních údajů, nedůvěra oslovených v prováděná výzkumná šetření a nedostatek času na straně respondentů (Bates, Creigh-

¹ Anglický pojem non-response nemá adekvátní český ekvivalent, a proto v dalším textu používám pojem výpadek návratnosti navrhovaný Krejčím (2008).

ton 2000). Mezi důvody klesající návratnosti lze spatřovat také změny v obvykle používaném designu prováděných šetření, neboť některé metody sběru dat používané v posledních letech se vyznačují nižší mírou návratnosti než jiné². Ke klesající návratnosti může rovněž přispívat snižující se kvalita výběrové opory (Groves et al. 2002) dostupné pro účely provádění výzkumných a evaluačních šetření.

1.1 Podstata problému a jeho význam

Pro adekvátní posouzení návratnosti je zapotřebí odlišit ji od dalších souvisejících jevů a nastínit způsob, jak návratnost ve výzkumné a evaluační praxi korektně vykazovat. Návratnost nejčastěji představuje podíl respondentů, kteří se zapojili do výzkumu na celkovém počtu platných (*eligible*) jednotek (Bailar, Lanphier 1978 či Madow et al. 1983), přičemž za platnou jednotku lze považovat subjekty, které skutečně náleží do definované cílové skupiny³. Při posuzování návratnosti je však třeba obrátit pozornost také k tzv. pasivním kontaktům, tedy k jednotkám, které zbývají po odečtení respondentů z celkového souboru platných jednotek. Curtin, Presser a Singer (2005) v tomto ohledu rozlišují tři skupiny. První představují jednotlivci, kteří sice jsou platnými jednotkami, ale např. v důsledku nesprávných kontaktních údajů nebyla s těmito jednotkami navázána komunikace (*non-contacts*). Druhou skupinu představují ti, kteří sice byli osloveni, ale explicitně vyjádřili svou neochotu na daném výzkumu participovat (*refusals*). Třetí skupinu pak představují ostatní, kteří se výzkumu zúčastnit nemohli, např. z důvodu nemoci, kvůli jazykové bariéře či v důsledku zmeškání termínu pro vyplnění dotazníku. Jednotlivé způsoby výpočtu návratnosti se liší zahrnováním či vydělováním jednotlivých typů pasivních respondentů, nicméně převládá shoda na způsobu výpočtu návratnosti dle metodiky AAPOR⁴. Z hlediska účelu tohoto textu má

² Markantní je srovnání typické návratnosti mezi šetřeními prováděnými metodou CAWI (Computer Assisted Web Interviewing, tedy pomocí on-line dotazníků, které vyplňují jednotliví respondenti samostatně) a face-to-face interview. Zvyšující se podíl výzkumů prováděných metodou CAWI, pro které je nižší návratnost přirozená, tak může přispět ke klesajícímu trendu vykazované návratnosti.

³ To mnohdy nemusí korespondovat s výčtem jednotek ve výběrové opoře, neboť některé výběrové opory zahrnují i jednotky, které jsou zařazeny omylem nebo neodpovídají vymezení cílové skupiny (což se zjistí až během samotného dotazování, např. během screeningu).

⁴ AAPOR (American Association for Public Opinion Research) definuje celkem šest měř návratnosti; v tomto textu je návratnost kalkulována dle standardu RR2.

rozlišování jednotlivých skupin pasivních kontaktů zásadní význam, neboť se k němu váže jeden z důležitých závěrů vyplývajících z provedených analýz.

1.2 Faktory ovlivňující návratnost

Z rozlišení jednotlivých skupin pasivních kontaktů lze dovodit, že nízká návratnost (za níž lze považovat 10–15 %) je buď důsledkem odmítnutí ze strany oslovených jednotlivců požadované údaje poskytnout (*refusals*), nebo je důsledkem neuskutečněního kontaktu s potenciálními respondenty (*non-contacts*). Odmítnutí může být ovlivněno individuálními vlastnostmi tazatelů a jejich nedostatečnými komunikačními schopnostmi (pokud dané šetření prováděno metodou face-to-face dotazování); v případě on-line výzkumů (CAWI) může být odmítnutí vyvoláno grafickou podobou dotazníku či nedostatečnou přehledností příslušné aplikace (Groves et al. 2002). Vysoká míra odmítnutí, a tím vyvolaná nízká návratnost, může být rovněž odrazem neznalosti evaluátora o specifikách cílové skupiny nebo může poukazovat na neadekvátnost použitého výzkumného instrumentu. Mezi hlavní důvody neuskutečněního kontaktu patří nesprávné kontaktní údaje a chyby při doručení, kdy je např. e-mail zařazen do spamu nebo kdy je doručený e-mail adresátem ignorován.

V návaznosti na uvedené rozlišení odmítnutí a neuskutečnění kontaktu následuje podrobnější rozbor jejich vlivu na návratnost. Groves a Couper (1998) v této souvislosti uvádějí faktory, které míru odmítnutí (a potažmo také nízkou návratnost) ovlivňují nejvíce. Rozlišují přitom vlivy exogenní vyplývající z kontextu, v němž je dané šetření realizováno a evaluátor ho nemůže nijak ovlivnit, od vlivů endogenních, které představují ovlivnitelné charakteristiky výzkumného uspořádání. Do první skupiny lze zařadit zejména vliv sociálního prostředí, tedy např. frekvenci prováděných šetření a sociodemografické a socioekonomické charakteristiky cílové skupiny (profesní zařazení respondentů, jejich pozice v zaměstnání apod.). Mezi významné faktory zvyšující odmítnutí dále patří nízká vnímaná relevance daného šetření, obava ze zneužití poskytnutých údajů, negativní postoj k zadavateli/realizátorovi výzkumu či nechuť vyjadřovat a sdílet vlastní názory (Lessler, Kalsbeek 1992). Odmítnutí může naopak snížit pozitivní motivace respondentů – Groves, Singer a Corning (2000) v této souvislosti upozorňují na vyšší pravděpodobnost zapojení jednotlivců, kteří mají konkrétní zájem na výsledcích prováděného šetření, a kteří jsou

přesvědčení o tom, že výsledky daného výzkumu mohou ovlivnit jejich pozici⁵.

Druhou skupinu determinant (tj. endogenní vlivy) tvoří konkrétní charakteristiky daného šetření, které lze souhrnně označit jako výzkumný design⁶. Vzhledem k tomu, že podrobný rozbor těchto faktorů přesahuje rámec tohoto textu, je pozornost zaměřena pouze na explikaci významu metod sběru dat. Platí, že každá ze zvolených metod s sebou nese specifické okolnosti, jež odmítnutí a potažmo návratnost ovlivňují. V případě face-to-face interview či CATI⁷ může oslovený odmítnout účast kvůli nevhodným okolnostem, za kterých byl tazatelem osloven (tedy např. kvůli časovému zaneprázdnění). Při CAWI může být důvodem odmítnutí nerosrozumitelnost pokynů či nevhodně zvolená aplikace pro dotazování. Na druhé straně lze odmítnutí snížit zdůrazněním důležitosti výzkumu již v prvotní zprávě o probíhajícím šetření, otevřenou komunikací s potenciálními respondenty zahrnující mj. zřetelně formulované cíle výzkumu, použitými incentivami či podporou ze strany autority, kterou oslovení uznávají.

Při výzkumech face-to-face lze konkrétní důvod výpadku návratnosti snadno zjistit, neboť tazatelé mohou zaznamenat, zdali konkrétní osobu na dané adrese nezastihli, či zdali se oslovený odmítl výzkumu zúčastnit. V rámci šetření bez tazatelů (tedy zejm. CAWI) však důvody nízké návratnosti spolehlivě rozlišit nelze.

1.3 Dopady nízké návratnosti a způsoby řešení

Řada zadavatelů, mnozí uživatelé výsledků provedených šetření i evaluátorů samotných posuzuje na základě dosažené návratnosti kvalitu výzkumu a úspěšnost terénních prací. Také diskuse v odborné literatuře se opakovaně vrací k otázce, jaká míra návratnosti je akceptovatelná a jaká rizika představuje nízká návratnost z hlediska přesnosti výzkumných odhadů. Z některých studií je přitom patrné, že nízká návratnost skutečně může snížit reprezentativitu (Lessler, Kalsbeek 1992) a zvýšit

⁵ V důsledku by tento pocit mohl ovlivnit nejen návratnost, ale i vychýlení výzkumných odhadů.

⁶ Jde zejména o techniku výběru, zvolenou metodu sběru dat a čas vymezený na terénní práce.

⁷ CATI – Computer Assisted Telephone Interviewing

pravděpodobnost vychýlení výzkumných odhadů. Na druhé straně však jiné výzkumy (mj. Cutrin, Presser a Singer 2005; Keeter et al. 2000; Schouten 2004; Merkle, Edelman 2002) ukazují, že i šetření s nízkou návratností mohou být reprezentativní a mohou na základě dosažených výsledků umožňovat formulaci závěrů o celé populaci. Tyto dvojaké poznatky dokazují, že míra návratnosti není sama o sobě vhodným indikátorem ani přesnosti (resp. nevychýlenosti) výzkumných odhadů, ani kvality provedení výzkumu. Před neadekvátními závěry dovozovanými z míry návratnosti varují také Groves, Fowler a Couper (2004), kteří vyvracejí častý omyl, kdy je míra návratnosti považována za ukazatel kvality výzkumu.

Nízká návratnost tedy nesnižuje sama o sobě hodnotu provedeného šetření a využitelnost jeho výsledků. Závažným nedostatkem, s nímž je nízká návratnost mnohdy nesprávně zaměňována, je vychýlení (*bias*)⁸, jehož výskyt je při nízké návratnosti pravděpodobnější (Biemer, Lyberg 2003; Thomsen et al. 2006). Tedy nikoliv nízká návratnost *per se*, nýbrž vychýlení výzkumného odhadu může vést k nesprávným závěrům o populaci (Groves, Peytcheva 2008; Järnbert, Öhrvall 2012). Proto by specifikaci návratnosti měl doprovázet také údaj o vychýlení výzkumných odhadů, např. ve formě analýzy rozdílů mezi charakteristikami respondentů a pasivních kontaktů. V podmínkách výzkumné a evaluační praxe je tedy žádoucí vykazovat nejen samotnou míru návratnosti, ale je také vhodné testovat schopnost daného výběrového souboru reprezentovat základní populaci a vhodným způsobem o ní informovat zadavatele a další uživatele.

Pokud je z provedené analýzy patrné, že výzkumné odhady jsou skutečně vychýleny, lze vzniklou situaci řešit využitím některého ze zavedených způsobů. Jde především o tyto postupy:

a) substituce

V tomto případě jde o náhradu subjektů, které se do výzkumu nezapojily, jinými jednotkami. V praxi se může jednat o dosběr, tedy o provedení sběru dat v dodatečně koncipovaném výběru kompenzujícím nedostatek primárního sběru či o extenzi cílové skupiny. V takovémto případě mo-

⁸ Vychýlení (*bias*) je rozdíl mezi výzkumným odhadem a parametrem základní populace, resp. skutečnou hodnotou (Brick, Bose 2001).

hou být do vzorku zahrnuti např. nejen žadatelé, kteří podali projekt v letech 2012–2013, ale také ti, kteří podali své projektové žádosti už v roce 2011.

b) korekce zjištění na základě odhadu výchylky

Porovnáním známých hodnot pro skupinu respondentů a pasivních kontaktů (např. podíl mužů a žen ve výběru oproti podílu mužů a žen v základní populaci) lze získat údaje o vychýlení získaných dat. A to jak o vychýlení vzorků oproti základní populaci, tak také o vychýlení pasivních kontaktů vůči souboru respondentů. Tyto údaje pak mohou být využity pro korekci ostatních zjištění, kde už takováto možnost srovnání není. Nastíněný postup je však poměrně rizikový a lze ho doporučit jen ve výjimečných případech⁹.

c) kompenzace pomocí statistických metod

V tomto případě jde o použití specifických postupů, jež mohou zmírnit dopady nízké návratnosti. Nejčastěji používaným postupem je aplikace post-stratifikační váhy, kdy je relativní zastoupení jedné podskupiny respondentů posíleno či oslabeno tak, aby lépe odráželo její zastoupení v základní populaci. Takovýto postup však nachází uplatnění spíše při řešení výpadku návratnosti u jednotlivé otázky (*item nonresponse*).

Vzhledem k tomu, že každý z uvedených postupů s sebou nese řadu rizik, je prvořadým zájmem vychýlení (resp. nízké návratnosti, jež může vychýlení způsobit) předcházet. Jednotlivá preventivní opatření pak mohou být zaměřena jak na zvyšování pravděpodobnosti, že žádoucí subjekt bude osloven (tedy, že mu bude doručena výzva k účasti na výzkumu), tak také na posilování motivace k zapojení do výzkumu. Keeter et al. (2000) upozorňují, že zvýšená návratnost je častokrát vykoupena mimořádným úsilím vynaloženým během terénních prací. Konkrétní opatření směřující ke zvyšování návratnosti přitom mohou mít mnoho různých forem, z nichž nejvýznamnější jsou podrobněji popsány v následujícím textu.

a) Nulová varianta, která častokrát spontánně vyplývá ze situace, kdy problematice návratnosti není věnována zvláštní pozornost. Z mnohých evaluačních zpráv z ČR i ze zahraničí je patrné, že dosažená návratnost bývá přijímána jako fakt beze snahy, zájmu či potřeby přijímat kon-

krétní opatření směřující k jejímu zvýšení. V mnoha jiných případech je však nulová varianta výsledkem záměrného rozhodnutí založeného např. na analýze dokazující, že se pasivní kontakty neodlišují od respondentů. Existují navíc okolnosti, za kterých může být aplikace postupů pro zvýšení návratnosti považována za nekorektní a eticky či dokonce právně spornou. V praxi může být např. vysoký počet dodatečných výzev mnohými adresáty vnímán jako obtěžování.

b) Dalším z opatření, jež v této souvislosti zmiňují např. Lessler a Kalsbeek (1992), je redesign sběru dat. Jejich doporučení směřuje zejména k uplatnění mixed-mode¹⁰, ke zvýšení flexibility časování terénních prací, k volbě adekvátních incentív a k provedení speciálního školení tazatelů zaměřeného na minimalizaci odmítnutí.

c) Častým způsobem zvyšování návratnosti je opakovaná žádost o participaci na výzkumu. V případě face-to-face výzkumů má typicky podobu opakovaných návštěv tazatele v místě bydliště či pracoviště respondentů, zatímco při výzkumech prováděných metodou CATI jde o opakovaná volání, nezřídka přesahující 10 pokusů. V rámci CAWI výzkumů jsou obvykle používány dodatečné výzvy k vyplnění dotazníku (tzv. *follow-ups*) určené subjektům ve výběrovém souboru, které do stanoveného data daný dotazník nevyplnily¹¹.

S ohledem k zaměření této statě se další pozornost soustřeďuje na způsoby a formy uplatnění dodatečných výzev v rámci CAWI šetření. Dodatečné výzvy mohou být koncipovány různým způsobem; může být pro ně použit identický komunikační kanál, kterým byl distribuován samotný dotazník nebo může být záměrně zvolen jiný způsob komunikace (např. v rámci CAWI výzkumu mohou být dodatečné výzvy provedeny telefonicky). Jednotlivá šetření používající dodatečné výzvy se dále mohou lišit počtem vln a jejich časovým rozestupem, např. Dillman (1972) v této souvislosti navrhuje na základě vlastních pozorování realizovat tři vlny dodatečných výzev následované po první pozvánce k participaci na výzkumu během sedmi týdnů. Dodatečné výzvy se mohou lišit také obsahovou

¹⁰ Jde o specifický typ uspořádání sběru dat, který poskytuje respondentovi možnost volby mezi několika metodami sběru dat.

¹¹ Aplikace dodatečných výzev však není vázána výlučně na CAWI; stejně dobře mohou být dodatečné výzvy použity i při jiných metodách sběru dat.

⁹ Více o možnostech korekce zjištění pojednává např. Lessler a Kalsbeek (1992).

náplní: může se jednat o upozornění na blížící se termín (Dohremwend 1970), o oznámení o prodloužení termínu (Ferber, Sudman 1974) či o urgenci. Výjimečně není ani opakované zaslání dotazníku či linku k on-line dotazníku, a dokonce se lze setkat i s dodatečně rozesílaným ujištěním o potřebnosti spolupráce (Gunn, Rhodes 1981).

Na účinnost dodatečných výzev a na okolnosti jejich uplatnění ve výzkumu ukazuje řada konkrétních šetření (mj. Dillman 1991; Mehta, Sivadas 1995; Sheehan, Hoy 1999), přičemž mnohé prameny (Groves, Cialdini a Couper 1992) zdůrazňují, že dodatečné výzvy mají zásadní vliv na výslednou míru návratnosti. Lohr (2010) v tomto ohledu upozorňuje, že tlak na zvyšování velikosti výběrového souboru bez adekvátní analýzy pasivních kontaktů a zacílení konkrétních aktivit na tuto skupinu sice může zvýšit celkový počet případů, nicméně ponechává stranou otázku dopadu dodatečných výzev na vychýlení výzkumných odhadů.

V případech, kdy se jednotlivci bezprostředně reagující na úvodní oslovení (promptní respondenti) neliší od ostatních, mohou dodatečné výzvy zvýšit návratnost, aniž by tím došlo k vychýlení výzkumných odhadů. Jiná situace však nastává, pokud se z nějakého důvodu promptní respondenti ve svých postojích a chování významně liší od ostatních. Potom dodatečné výzvy mohou kromě zvýšení návratnosti také snížit vychýlení, neboť do vzorku budou zahrnuti i ti, kteří na úvodní oslovení nereagovali. Považuji však za nutné upozornit, že mohou nastat i situace, kdy se vychýlení výzkumných odhadů v důsledku použití dodatečných výzev naopak zvýší.

Ačkoliv platí, že zvyšování návratnosti je nejlepším způsobem, jak snížit vychýlení (Brick, Bose 2001), je třeba mít na paměti, že každé z opatření má kromě svých výhod také konkrétní rizika.

2. Popis analyzovaných dat

Výsledky prezentované v této stati vznikly na základě samostatné dílčí úlohy vycházející z projektu „Stanovení metodologie zjišťování hodnot vybraných indikátorů výsledků pro Operační program Technická pomoc 2014–2020 a zjištění výchozích hodnot v roce 2014“ realizovaného Institutem evaluací a sociálních analýz pro Ministerstvo pro místní rozvoj v roce

2014. Původním cílem citovaného projektu bylo provést šetření ve čtyřech cílových skupinách, nicméně údaje prezentované v tomto článku se s ohledem k předmětu zde prezentované dílčí analýzy zaměřují jen na jednu z nich, a to na žadatele/příjemce. Obsahovým cílem projektu bylo hodnocení spokojenosti vybraných aktérů s vymezenými oblastmi; šetření realizované v cílové skupině žadatelů/příjemců se kromě celkové spokojenosti s informačním systémem zaměřilo také na spokojenost s pracovním prostředím informačního systému BENEFIT7, spokojenost s daty obsaženými v systému a spokojenost s technickou podporou.

Šetření bylo realizováno metodou CAWI, která umožnila respondentům postupovat při vyplňování dotazníku jejich vlastním tempem a navíc absence tazatele vedla k vyšší otevřenosti respondentů a snížila tak výskyt tzv. sociální desirability. Na druhé straně bylo při aplikaci metody CAWI třeba počítat s vyšším rizikem výběrového vychýlení v důsledku samovýběru. V daném případě byly obavy spojovány s vyšším zapojením mimořádně nespokojených či naopak velmi spokojených jednotlivců, na úkor žadatelů/příjemců s nevyhraněnými či neutrálními postoji.

Celková doba trvání terénních prací činila 15 dní a vlastní sběr dat probíhal ve třech vlnách. Po úvodním rozeslání pozvánky k zapojení do výzkumu byl respondentům ponechán časový prostor v délce šesti dní. Po uplynutí této šestidenní lhůty byla těm, kteří nereagovali, zaslána opakovaná výzva k zapojení do prováděného výzkumu (první dodatečná výzva) a byla stanovena nová lhůta pro vyplnění dotazníku v délce čtyř dní. Poté byla rozeslána další výzva k participaci (druhá dodatečná výzva) a po dalších 4 dnech byly linky k on-line dotazníku deaktivovány a proces sběru dat byl ukončen.

2.1 Cíl analýzy

Meritorní cíle analýzy prezentované v této stati jsou dva; první cílem je ověřit reprezentativitu vzorku vzniklého v podmínkách CAWI výzkumu v cílové skupině žadatelů/příjemců. Druhým cílem je prozkoumat efekty použití dodatečných výzev. V rámci druhého z cílů se analytická pozornost zaměřila na změny návratnosti po jednotlivých výzvách, na difference ve struktuře vzorku, na rozdíly ve způsobu vyplnění dotazníku (*response pattern*) a na změny v charakteru zjištění. Hypotézy, které byly v rámci jednotlivých analytických úloh formulovány, jsou:

- a) Výběrový soubor je reprezentativní ve vztahu k základní populaci žadatelů/příjemců (resp. výběrový soubor není vůči základní populaci vychýlen).
- b) Dodatečné výzvy vedou ke zvýšení celkové návratnosti.
- c) Dodatečné výzvy nemají vliv na strukturu vzorku (resp. rozdíly mezi respondenty, kteří reagují na jednotlivé výzvy, nejsou statisticky významné).
- d) Rozdíly respondentů v přístupu k vyplňování dotazníků se mezi jednotlivými výzvami neliší.
- e) Dodatečné výzvy nemají vliv na podstatu meritorních zjištění.

2.2 Metodika

V rámci specifikovaného výzkumného šetření byl proveden test účinnosti dodatečných výzev a analýza jejich efektů, nicméně významnou otázkou bylo rovněž ověření reprezentativity vzorku respondentů. V této souvislosti je třeba předeslat, že pojem reprezentativity má v relevantní literatuře několik různých interpretací¹², z nichž některé odkazují ke „zmenšenině populace“ či k „absenci selektivních sil“. V kontextu provedeného šetření byla reprezentativita chápána jako absence vychýlení výzkumných odhadů.

Při ověřování reprezentativity se obvykle porovnávají výzkumné odhady se známými parametry základní populace, např. Thomsen a Siring (1983) takto porovnávali údaje získané v rámci jimi realizovaného šetření s údaji ze sčítání lidu. V praxi však často nastává situace, kdy vychýlení není možné tímto způsobem exaktně změřit vzhledem k tomu, že nejsou známy skutečné hodnoty parametrů základní populace. Nejinak tomu bylo i v případě šetření provedeného Institutem evaluací a sociálních analýz. V rámci konkrétních okolností sběru dat však bylo v reálných podmínkách možné vytvořit dva na sobě nezávislé výběry. Vzhledem k tomu, že návratnost v rámci prováděného šetření¹³ dosáhla nižší než očekávané míry, byl ze základního souboru vybrán ještě druhý výběrový soubor, jehož cílem bylo zvýšit počet platných případů pro následné analýzy.

¹² Jejich přehled viz Kruskal a Mosteller (1979).

¹³ Počet oslovených respondentů byl záměrně relativně nízký vzhledem ke snaze minimalizovat zátěž zkoumané populace – v případě žadatelů/příjemců používajících informační systém BENEFIT7 je totiž riziko v úvodu zmíněné přezkoumanosti poměrně vysoké.

Tímto opatřením byl jednak naplněn původní cíl, tedy zvětšení velikosti zkoumaného vzorku, a navíc tak vnikla mimořádná příležitost k ověření reprezentativity pomocí srovnání výzkumných odhadů založených na takto vzniklých datových souborech a k verifikaci hypotézy o přítomnosti, resp. absenci vychýlení.

Nulová hypotéza v tomto případě předvídá podobnost výzkumných odhadů založených na obou výběrech, resp. předjímá, že případné rozdíly v jednotlivých zjištěních nejsou statisticky významné. Z podobnosti obou výběrů lze dovodit, že by výzkumné odhady nevykazovaly statisticky významné difference ani při porovnání s parametry základní populace (pokud by tyto parametry byly známy a pokud by bylo možné daný test skutečně provést). Alternativní hypotéza počítá s existencí statisticky významných odlišností mezi výzkumnými odhady založenými na jednotlivých výběrech. Pokud by byly tyto difference prokázány, byl by tím indikován výskyt vychýlení. Na základě dostupných dat by sice nebylo možné určit, který z obou výběrů je vychýlený oproti základní populaci, nicméně z hlediska interpretace dosažených výsledků by i takovýto závěr byl užitečný a smysluplný.

V rámci obou výběrových souborů byly osloveny všechny zařazené jednotky. Postup provedení sběru dat byl pro oba výběrové soubory identický; jediná odlišnost se týkala načasování, kdy sběr dat ve výběrových souborech neprobíhal paralelně, nýbrž sukcesivně.

Prvotní výzva obsahovala úvodní dopis s vysvětlením účelu a významu prováděného šetření, odkaz s možností ověřit si věrohodnost prováděného šetření a apel na zapojení do výzkumu; incentive nebyly v tomto případě nabízeny.

3. Zjištěné výsledky

Dosažené výsledky jsou prezentovány v pěti částech, které korespondují s výše formulovanými hypotézami. Prvotní úlohou bylo posoudit reprezentativitu výběrového souboru a odhadnout míru vychýlení. Následné úlohy se zaměřily na vliv dodatečných výzev na celkovou návratnost, na hodnocení rozdílu ve struktuře výsledného vzorku vyvolaných dodatečnými výzvami, na posouzení rozdílu ve způsobu vyplnění on-line do-

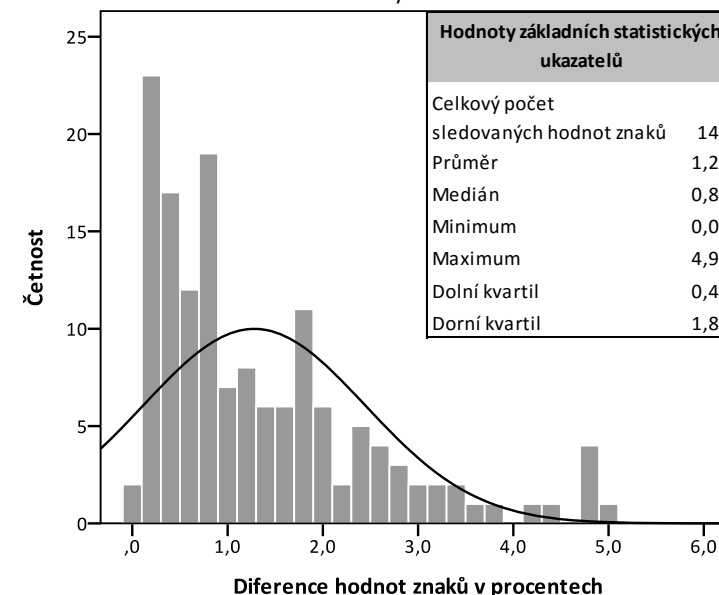
tazníku mezi respondenty participujícími na výzkumu v rámci jednotlivých výzev a v neposlední řadě také na analýzu hlavních zjištění a změn jejich charakteru vyvolaných dodatečnými výzvami.

3.1 Ověření reprezentativity

V rámci provedeného šetření bylo zjišťováno celkem 58 proměnných (bez pomocných a konstruovaných znaků) a pro každou z nich byly testovány rozdíly v dosažených hodnotách mezi oběma výběrovými soubory. Z hlediska dalších analýz je třeba uvést, že toto srovnání bylo provedeno na základě dat po obou vlnách dodatečných výzev. Základním výstupem exploratorní analýzy je histogram, který dokumentuje četnost rozdílů v hodnotách sledovaných proměnných mezi oběma výběrovými soubory¹⁴. Je patrné, že hodnoty všech sledovaných znaků se odchylují od normálního rozložení a jsou sešikmené směrem k nižším diferenčním hodnotám. V průměru se jednotlivé hodnoty znaků mezi oběma soubory liší o 1,28 procentního bodu (střední hodnota přitom činí 0,85 procentního bodu). Nejmenší pozorovaný rozdíl činí 0 procentního bodu (jde tedy o identické hodnoty zjištěné v rámci obou výběrových souborů), zatímco největší difference činí 4,9 procentního bodu. Tři čtvrtiny ze všech 146 porovnávaných hodnot se mezi oběma soubory liší nejvýše o 1,8 procentních bodů.

¹⁴ N je v tomto případě rovno celkovému počtu porovnávaných hodnot; celkem bylo porovnáno 146 hodnot.

Graf 1: Rozložení diferenciací sledovaných hodnot znaků



Zdroj: autor

Následná analýza se zaměřila na testování statistické významnosti rozdílů mezi odhady obou výběrových souborů¹⁵. Z výsledků vyplynulo, že rozdíly jednotlivých znaků mezi oběma výběrovými soubory nejsou v naprosté většině statisticky významné. Výjimky v tomto ohledu představují znaky měřící vnímanou důležitost dílčích charakteristik informačního systému. Z uvedeného histogramu (graf 1) je patrné, že difference v hodnotách těchto znaků přesahují čtyři procentní body. Tyto difference jsou však ojedinělé, inkonzistentní pokud jde o směr vychýlení (některé hodnoty jsou v druhé vlně vyšší oproti první vlně, zatímco u jiných hodnot je tomu přesně naopak) a lze je proto považovat za nahodilé jevy.

Vzhledem k tomu, že jde o ekvivalentní výběry pořízené z výběrové opory zahrnující všechny jednotky základní populace, je na základě dosažených výsledků možné usuzovat, že oba výběrové soubory jsou nevychýlené.

¹⁵ Testovány byly difference mezi všemi 58 proměnnými; rozdíly mezi nominálními znaky byly testovány pomocí chí-kvadrátového testu, v případě ordinálních a kardinálních proměnných byl použit Mann–Whitneyův U-test.

lené. Lze rovněž shrnout, že získané vzorky reprezentují základní populaci a výzkumné odhady skutečně odrážejí parametry základní populace.

3.2 Změna návratnosti po jednotlivých výzvách

Je patrné, že dodatečné výzvy zvýšily celkovou návratnost zásadním způsobem, neboť v rámci prvního z výběrových souborů odpovědělo dokonce více respondentů až na dodatečné výzvy (503 respondentů) než na úvodní oslovení (451 respondentů). Efekt dodatečných výzev tedy v případě prvního výběrového souboru činí 112 procent; návratnost se v tomto případě totiž zvýšila z výchozích 16,70 % až na výsledných 35,32 %. V rámci druhého výběrového souboru činí celkový efekt dodatečných výzev 76 %, kdy se návratnost zvýšila z výchozích 18,34 % na výsledných 32,32 %. Z tabulky 1 jsou patrné také absolutní počty vyplněných dotazníků. Je tedy patrné, že dodatečné výzvy návratnost skutečně významně zvyšují.

Tabulka 1: Změny v návratnosti po jednotlivých výzvách

	první výběrový soubor			druhý výběrový soubor		
	absolutní četnost (N)	kumulativně vyjádřená návratnost v %	změna v %	absolutní četnost (N)	kumulativně vyjádřená návratnost v %	změna v %
celkový počet oslovených	2,701	---	---	2,704	---	---
úvodní oslovení	451	16,70	---	496	18,34	---
první dodatečná výzva	327	28,80	+ 12,11	285	28,88	+ 10,54
druhá dodatečná výzva	176	35,32	+ 6,52	93	32,32	+ 3,44
kumulovaný efekt výzev	503	112	+18,63	378	76	+ 13,98

Zdroj: autor

Rovněž je patrná klesající mezní účinnost dodatečných výzev, neboť v rámci prvního výběrového souboru zvýšila první vlna dodatečných výzev návratnost o 72 %, zatímco druhá dodatečná výzva přispěla

ke zvýšení o 23 %. V druhém výběrovém souboru zvýšila první výzva návratnost o 57 % a druhá výzva pak návratnost zvýšila o dalších 12 %.

Tyto hodnoty jsou vyšší, než kolik uvádějí některé prameny; např. Groves, Cialdini a Couper (1992) pozorovali ve své studii zvýšení návratnosti v průměru o 11 procent na každou vlnu.¹⁶ Každopádně je však patrné, že dodatečné výzvy mají odezvu (byť postupně klesající) a vedou ke zvýšení celkové návratnosti.

3.3 Diference ve struktuře vzorku

V rámci verifikace hypotézy o vlivu dodatečných výzev na reprezentativitu se analýza zaměřila na podrobný rozbor výzkumných odhadů (hodnot všech znaků) v rámci obou výběrových souborů. Následující analýza se obrací ke struktuře souborů založených na odpovědích respondentů reagujících na jednotlivé výzvy. Pozornost se přitom soustřeďuje na hlavní popisné znaky, jakými jsou typ instituce, v níž respondenti pracují, frekvence práce s informačním systémem BENEFIT7 a v neposlední řadě také na základní sociodemografické znaky jakými jsou pohlaví, věk a nejvyšší dosažené vzdělání respondentů.

Z výsledků je patrné, že pracovníci z veřejné správy reagovali nejčastěji promptně – tedy bezprostředně na základě úvodní pozvánky k zapojení do výzkumu; naproti tomu zaměstnanci v podnikatelském sektoru reagovali nejčastěji na základě první a druhé dodatečné výzvy. Dále platí, že pracovníci nestátních neziskových organizací reagovali typicky až po druhé dodatečné výzvě. Rovněž jsou zřejmé rozdíly dané frekvencí práce s informačním systémem, kdy promptně odpověděli zejména ti, kteří se systémem BENEFIT7 pracují často (tj. každý den či několikrát týdně), zatímco ostatní reagovali až na první a druhou dodatečnou výzvu. Žádné další změny ve struktuře podsouborů nejsou až na uvedené difference statisticky významné.

¹⁶ Podobně také Enander, Sajti (1999); Batagelj, Lozar a Vehovar (1998) či Flemming, Sonner (1999) ve svých výsledcích uvádějí, že dodatečné výzvy zvýšily v jimi prováděných šetřeních návratnost přibližně o třetinu.

Tabulka 2: Struktura vzorku po jednotlivých výzvách

		Dodatečné výzvy			Celkem
		úvodní oslovení	první dodatečná výzva	druhá dodatečná výzva	
Pohlaví respondenta:	muž	39,4 %	36,7 %	33,0 %	37,5 %
	žena	60,6 %	63,3 %	67,0 %	62,5 %
Věk respondenta:	méně než 30 let	15,0 %	14,3 %	17,4 %	15,1 %
	30–49 let	65,1 %	68,0 %	65,1 %	66,2 %
	50 a více let	19,9 %	17,7 %	17,4 %	18,7 %
Nejvyšší dosažené vzdělání respondenta	základní, střední	16,3 %	12,8 %	12,4 %	14,5 %
	vysokoškolské	83,7 %	87,2 %	87,6 %	85,5 %
Frekvence práce s informačním systémem BENEFIT7:	minimálně několikrát za týden	23,8 %	23,1 %	16,5 %	22,5 %
	několikrát za měsíc/čtvrtletí	56,2 %	62,7 %	66,5 %	60,1 %
	méně často	20,1 %	14,1 %	17,0 %	17,4 %
Typ organizace, v níž respondent pracuje:	veřejná správa	32,8 %	28,8 %	26,6 %	30,4 %
	podnikatelský subjekt	28,4 %	34,7 %	31,7 %	31,2 %
	nestátní nezisková organizace	21,2 %	19,0 %	27,5 %	21,3 %
	ostatní	17,6 %	17,5 %	14,2 %	17,1 %
Celkem		100,0 %	100,0 %	100,0 %	100,0 %
		947	612	269	1.828

Zdroj: autor

Přestože ani další dva případy diferencí nejsou statisticky významné, stojí za zmínku. Uvedená zjištění totiž korespondují se závěry jiných šetření (Willke et al. 1999; Lohr 2010), která dokládají rozdíly podle pohlaví, kdy na dodatečné výzvy reagují spíše ženy. Dosažené výsledky v tomto ohledu dokládají vyšší podíl žen reagujících na druhou dodatečnou výzvu (67 %) oproti struktuře vzorku promptních respondentů, v němž ženy dosahují podílu 61 %. Z dat je dále patrné, že starší respondenti reagují na prvotní výzvu ve větší míře, a tak dodatečné výzvy podporují zapojování mladších jednotlivců. Konkrétně se ukazuje, že zatímco na úvodní výzvu reagovalo 20 % respondentů starších 50 let, na dodatečné výzvy se do výzkumu nezapojilo více než 18 % z této věkové skupiny. V případě respondentů do 30 let činí jejich podíl v rámci souboru jednotlivců reagujících na první oslovení 15 %, zatímco v druhé vlně dodatečných výzev jejich podíl činí 17 %.

Z těchto zjištění lze tedy dovodit závěr, že dodatečné výzvy v konkrétním situačním kontextu přispívají ke změně struktury vzorku a mohou tedy mít vliv na vychýlení výzkumných odhadů. Data prokazují, že podskupina promptně reagujících respondentů se statisticky významně odlišuje od těch, kteří reagují až na základě dodatečných výzev (a to v případě takových znaků jakými jsou frekvence práce s informačním systémem BENEFIT7 a typ organizace, v níž respondent pracuje). K podobnému závěru dospěly také mnohé další studie, které prokazují signifikantní rozdíl mezi promptními a ostatními respondenty (Dunkelberg, Day 1973; Traugott 1987; Voigt, Koepsell a Daling 2003). Uvedená zjištění tedy ukazují na význam provádění dodatečných výzev nejen z hlediska zvyšování návratnosti, ale také z hlediska zpřesňování výzkumných odhadů.

3.4 Diference ve způsobu vyplnění dotazníku

Další analytická pozornost byla věnována rozdílům ve způsobu vyplnění on-line dotazníku (*response pattern*). V tomto ohledu byly sledovány dvě charakteristiky dokumentující přístup respondentů k vyplnění dotazníku, a to výpadky návratnosti u otázek (*item nonresponse*), kde bylo možné se poskytnutí odpovědi vyhnout, a celkový čas věnovaný vyplňování dotazníku.

a) Výpadek návratnosti u jednotlivých otázek

Předmětem srovnání je počet respondentů, kteří neuvedli žádnou odpověď na položené dotazy. V rámci výzkumného instrumentu byly položeny celkem tři samostatné otevřené otázky, v jejichž případě byla on-line aplikace naprogramována tak, že je respondent mohl ponechat bez odpovědi. Počty respondentů, kteří neuvedli žádnou odpověď, jsou patrné z tabulky 3. Platí, že ve skupině respondentů, kteří reagovali bezprostředně na úvodní oslovení, vyplnilo všechny otevřené otázky 37,5 % jednotlivců. Podíl takto pečlivých respondentů ve skupinách respondentů reagujících na dodatečné výzvy klesl o přibližně čtyři procentní body.

Tabulka 3: Podíly respondentů, kteří neodpověděli na otevřené otázky

		Dodatečné výzvy			Celkem
		úvodní oslovení	první dodatečná výzva	druhá dodatečná výzva	
počet nevyplněných otevřených otázek	žádná (všechny vyplněné)	37,5%	34,2%	33,6%	35,7%
	jedna	12,8%	11,6%	13,4%	12,4%
	dvě	15,9%	16,1%	14,7%	15,8%
	tři	33,8%	38,2%	38,2%	36,0%
Celkem		100,0%	100,0%	100,0%	100,0%

Zdroj: autor

Z konkrétních dat je tedy zřejmá vyšší míra výpadků návratnosti u dotazníků vyplněných v rámci dodatečných výzev ve srovnání s dotazníky vyplněnými promptně. Bylo by tedy možné shrnout, že vzhledem k většímu počtu respondentů, kteří ponechali otevřené otázky nevyplněné, jsou dotazníky získané v rámci dodatečných výzev jako celek vyplněny méně pečlivě. Platnost tohoto závěru by měl doložit kratší čas strávený vyplňováním dotazníku. Další analýza se proto zaměřila na porovnání celkové doby vyplňování.

b) Čas strávený vyplňováním dotazníku

V návaznosti na předchozí závěr je důvodné očekávat, že čas věnovaný vyplňování dotazníku je v rámci dodatečných výzev kratší než čas, který vyplňování věnovali promptní respondenti. Z dosažených výsledků je však patrné, že se doba vyplňování dotazníku mezi jednotlivými vlnami liší o cca 10 vteřin, což při průměrné době vyplňování 12 minut není statisticky (a ani věcně) významný rozdíl. Úspora času, ke které dochází vynecháním některých (otevřených) otázek tedy nevede k celkově kratší době vyplnění dotazníku.

Tabulka 4: Čas věnovaný vyplňování dotazníku

	průměrná doba	směrodatná odchylka
úvodní oslovení	11:56	11:14
první dodatečná výzva	11:49	12:16
druhá dodatečná výzva	12:03	12:02
Celkem	11:55	11:44

Zdroj: autor

Navíc je na rozdíl od situace u meritorních otázek směrodatná odchylka v rámci dodatečných výzev vyšší ve srovnání se skupinou promptních respondentů (přibližně o jednu minutu), což ukazuje na větší rozptýl dat v případě respondentů zapojujících se až na základě dodatečných výzev.

3.5 Změny v charakteru zjištění

Ukázalo se, že dodatečné výzvy nepůsobí na výzkumné odhady jedno- směrně. Zatímco v některých případech hodnotu získanou promptními odpověďmi snižují, v jiných případech ji naopak zvyšují. Efekt dodatečných výzev je navíc v případě některých proměnných inkonzistentní (resp. nelineární), neboť druhá dodatečná výzva vychyluje získané odhady v opačném směru než první dodatečná výzva. V této souvislosti je třeba uvést, že zjištěné změny nejsou statisticky významné a nelze tak u hodnocených proměnných přičíst zjištěné změny vlivu dodatečných výzev¹⁷. Popisované jevy jsou z úsporných důvodů dokumentovány pouze na souhrnných indikátorech.

Tabulka 5: Změny ve výsledcích

		Průměrná hodnota	Směrodatná odchylka
indikátor spokojenosti s pracovním prostředím	Úvodní oslovení	64,69	14,765
	První dodatečná výzva	64,48	13,969
	Druhá dodatečná výzva	62,97	13,548
	Celkem	64,36	14,305
indikátor spokojenosti s daty	Úvodní oslovení	68,10	15,132
	První dodatečná výzva	68,67	14,273
	Druhá dodatečná výzva	67,99	13,579
	Celkem	68,30	14,597
indikátor spokojenosti s technickou podporou	Úvodní oslovení	69,85	17,603
	První dodatečná výzva	70,77	15,154
	Druhá dodatečná výzva	69,88	13,901
	Celkem	70,18	16,268
indikátor celkové spokojenosti se systémem BENEFIT7	Úvodní oslovení	61,58	20,615
	První dodatečná výzva	60,70	20,128
	Druhá dodatečná výzva	59,94	19,612
	Celkem	61,02	20,289

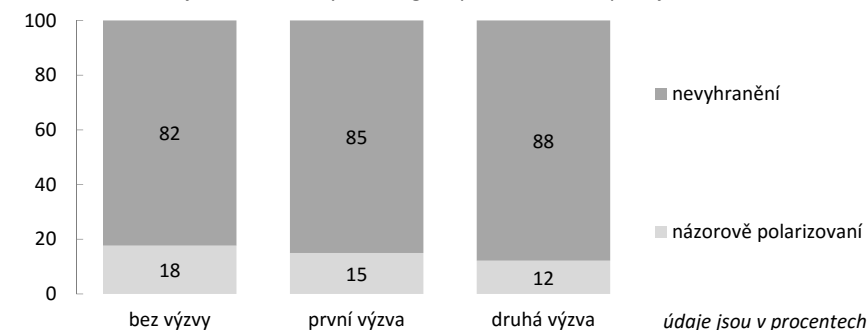
Zdroj: autor

¹⁷ Jde tedy o nevýznamný jev, jehož efekt by sice bylo možné predikovat pomocí regresního modelu, nicméně praktický smysl by takováto analýza neměla, neboť jde o unikátní, nezobecnitelný vliv týkající se konkrétních znaků.

Postupné snižování hodnoty je patrné v případě indikátoru spokojenosti s pracovním prostředím, kdy dodatečné výzvy snížily průměrné hodnocení z 64,69 % na celkových 64,36 %, a indikátoru celkové spokojenosti se systémem BENEFIT7, kdy došlo ke snížení z 61,58 % na celkových 61,02 %. Inkonzistentní vliv výzev na výzkumné odhady je patrný v případě indikátorů spokojenosti s technickou podporou a spokojenosti s daty.

Z podrobnější analýzy se dále ukazuje, že v rámci dodatečných výzev (bez ohledu na to, zdali jde o první či druhou výzvu) je směrodatná odchylka vždy nižší než v případě promptních odpovědí¹⁸. Na možnou příčinu snižující se směrodatné odchylky poukazuje další z provedených analýz, která dokumentuje postupné zapojování názorově nevyhraněných respondentů. Z dostupných dat lze vysledovat zřetelnou tendenci, kdy názorově polarizovaní respondenti (tj. ti, kteří jsou velmi spokojeni nebo velmi nespokojeni) reagují promptně, zatímco nevyhranění (tedy spíše spokojení nebo spíše nespokojení) se do výzkumu zapojují až na základě dodatečných výzev.

Graf 2: Součet krajních a středových kategorií při hodnocení spokojenosti



Zdroj: autor

Z nastíněných rozdílů v charakteru zjištění je patrné, že promptně reagují spíše respondenti s vyhraněnými názory, resp. lidé, kteří mají o daném tématu konkrétní představu. Naopak jednotlivci s ambivalentními či indiferentními postoji zapojení do výzkumu odkládají a zapojují se až na zá-

¹⁸ V této souvislosti je třeba upozornit, že každá vlna byla vždy hodnocena samostatně. Snižování směrodatné odchylky (a potažmo i rozptylu) tedy nemůže být způsobeno vyšším počtem případů.

kladě dodatečných výzev. Výsledný vzorek by tak při neprovedení dodatečných výzev nadhodnocoval postojové extrémy. Podobně jako v části věnované diferencím ve struktuře vzorku je i zde patrný efekt dodatečných výzev při zpřesňování výzkumných odhadů. Na mizivý rozdíl v čase potřebném pro vyplnění dotazníku, který je nastíněn výše, lze tedy z této perspektivy pohlížet jako na indikátor vyšší kognitivní zátěže, kterou vyplnění dotazníku představuje pro jedince s nevyhraněnými postoji (resp. s nižším povědomím o předmětu výzkumu).

4. Závěry

Na základě dosažených výsledků je tedy možné konstatovat, že nízká míra návratnosti nemusí sama o sobě nutně znamenat, že jsou výzkumné odhady vychýlené a zjištěné údaje nereprezentativní. Z výsledků mnou prováděných analýz je patrné, že vzorek promptních respondentů (tj. v rámci provedeného šetření návratnost 17–18 %) dokáže reprezentovat základní populaci stejně dobře jako vzorek doplněný o respondenty reagující na dodatečné výzvy. I když některé prameny poukazují na přijatelnou míru návratnosti (mj. Babbie 2007), nelze shrnout či uvést, jaká míra návratnosti je adekvátní či dostatečná, neboť záleží mj. na povaze pasivních jednotlivců. Pokud je např. skupina pasivních jednotlivců velmi heterogenní, tak i 90% návratnost může být nedostatečná, a naopak v mnohých jiných případech nemusí 20% návratnost představovat sebe-menší problém. Přesto je údaj o návratnosti důležitou informací, která by měla být uváděna spolu s dalšími informacemi při prezentaci výsledků z prováděných šetření.

Dosažené výsledky rovněž potvrzují známý fakt, že velikost zkoumaného souboru nemá na reprezentativitu přímý vliv. Nicméně vyšší počet jednotek ve vzorku poskytuje lepší analytické možnosti, neboť s vyšším počtem případů lze provádět detailnější analýzy či konstruovat přesněji vymezené segmenty. Konkrétní míru návratnosti je tedy vhodné posuzovat v kontextu homogenity či heterogenity základní populace; pokud je zkoumaná populace fragmentována, je použití nástrojů pro zvýšení návratnosti vhodné.

Dalším významným závěrem je, že dodatečné výzvy nezpůsobily v podmínkách provedeného šetření vychýlení dat, resp. že zjištěné odchylky

jsou nahodilé a nemají charakter systematického vychýlení. Nicméně v dílčích parametrech je patrná strukturální distorze výběrového souboru (v rámci provedeného šetření dle věku a dle frekvence práce s informačním systémem). V tomto ohledu jsou dosažené výsledky podobné s výsledky např. Willkeho et al. (1999), kteří uvádějí, že v jejich šetření byla struktura respondentů reagujících na dodatečné výzvy jiná ve srovnání se skupinou reagující na úvodní oslovení; k podobnému závěru dospěl mj. také Lohr (2010).

Navíc platí, že efekt výzev nemusí být jen pozitivní (jednosměrný); jinými slovy nemusí platit přímá úměra, že čím více výzev, tím lepší data. Dodatečné výzvy totiž mohou strukturu výběrového souboru vychýlit, neboť reaktance na opakované výzvy se liší podle typu respondenta. Z provedeného výzkumu se mj. ukazuje, že na opakované výzvy reagují ve větší míře ženy; podobně také názorově nevyhranění (ambivalentní) respondenti se zapojují zejména až na základě dodatečných výzev. Dodatečné výzvy tedy mohou ovlivnit rozptyl, a potažmo i přesnost výsledků. Při jejich použití je tedy vřele doporučeno důkladně analyzovat odlišnosti výzkumných odhadů mezi jednotlivými vlnami dodatečných výzev.

Počet dvou dodatečných výzev následujících po úvodním oslovení se ukázal optimálním kompromisem mezi snahou o maximalizaci návratnosti a neúnosnou obtrusivitou z pohledu respondentů. Je třeba připomenout, že dle platných etických kodexů upravujících provádění výzkumných šetření (AAPOR 2015) je zapojení do výzkumu dobrovolné a nelze ho vynucovat. Prvotní oslovení následované dvěma opakovanými výzvami dobrovolnost respondentů participovat na výzkumu respektuje a představuje pro oslovené jednotlivce přijatelný „nátlak“. V případě CAWI nejsou náklady na dodatečné výzvy podstatným výdajem, nicméně více než dvě vlny lze v rámci dané cílové skupiny považovat za kontraproduktivní. Jak je patrné z konkrétních výsledků provedeného šetření, nezvyšuje se míra návratnosti konstantně. Na základě porovnání účinnosti jednotlivých dodatečných výzev lze odhadnout, že třetí vlna by patrně nepřinesla více než 1–3 % odpovědí (cca 50 respondentů), což je zcela marginální efekt. Lze doporučit, aby problematice optimálního počtu dodatečných výzev a prokázání jejich účinků byla věnována další pozornost evaluátorů v rámci připravovaných šetření.

Zdroje

- [1] Aapor (2015) *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 8th edition. AAPOR
- [2] Atrostic, B. K., Bates, N., Burt, G., Silberstein, A., Winters, F. (1999) "Non-response in Household Surveys: New Measures and Insights", a paper presented at the International Conference on Survey Nonresponse, Portland, Oregon.
- [3] Babbie, E. R. (2007) *The Practice of Social Research*. Belmont: Wadsworth.
- [4] Bailar, B. A., Lanphier, C. M. (1978) *Development of Survey Methods to Assess Survey Practices*. Washington, D. C.: American Statistical Association.
- [5] Batagelj, Z., Lozar, K., Vehovar, V. (1998) *Respondent's Satisfaction in WWW Surveys*. Paper presented at the International Conference on Methodology and Statistics, Preddvor, Slovenia.
- [6] Bates, N., Creighton, K. (2000) *The Last Five Percent: What Can We Learn from Difficult Interviews?* In: Proceedings of the Annual Meetings of the American Statistical Association, 13–17 August.
- [7] Biemer, P., Lyberg, L. (2003) *Introduction to Survey Quality*. Wiley Series in Survey Methodology.
- [8] Brick, J. M., Bose, J. (2001) Analysis of potential nonresponse bias. Proceedings of the Joint Statistical Meetings, American Statistical Association, Atlanta.
- [9] Brownlee, K. A. (1957) A Note of the Effects of Nonresponse on Surveys. *American Statistical Association Journal*, 29–32.
- [10] Curtin, R., Presser, S., Singer, E. (2005) Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opinion Quarterly*, Volum 69, pp. 87–98.
- [11] de Leeuw, E., de Heer, W. (2002) *Trends in Household Survey Nonresponse: A Longitudinal and International Perspective*. In: Survey Nonresponse. New York: Wiley, pp. 41–54.
- [12] Deming, W. E. (1953) *On the Probability Mechanism to Attain an Economic Balance Between the Resultant Error of Non-Response and the Bias of Non-Response*. *Journal of the American Statistical Association*, 48, 743–772.
- [13] Dillman, D. A., (1991) The Design and Administration of Mail Surveys. In: *Annual Review of Sociology*. Vol. 17, pp. 225–249.
- [14] Dillman, D. A., (1972) Increasing Mail Questionnaire Response in Large Samples of the General Public. *Public Opinion Quarterly*. Vol. 36, pp. 257–257.

- [15] Dohremwend, B. S. (1970) An Experimental Study of Payments to Respondents. *Public Opinion Quarterly*. Vol. 34, pp. 621–624.
- [16] Donald, M. N. (1960) Implications of Nonresponse for the Interpretation of Mail Questionnaire Data. *Public Opinion Quarterly*. Vol. 24, pp. 99–114.
- [17] Dunkelberg, W. C., and Day, G. S. (1973) Nonresponse Bias and Callbacks in Sample Surveys In: *Journal of Marketing Research*. Vol 10, pp. 160–168.
- [18] Enander, J., Sajti, A. (1999) Online Survey of Online Customers, Value-Added Market Research through Data Collection on Internet. In: Proceedings of the ESOMAR Worldwide Internet Conference Net Effects. Amsterdam: ESOMAR, pp. 35–52.
- [19] Ferber, R., Sudman S. (1974) Effects of Compensation in Consumer Expenditure Surveys. *Annals of Economic and Social Measurement*. Vol. 3, pp. 319–331.
- [20] Flemming, G., Sonner, M. (1999) Can Internet Polling Work? Strategies for Conducting Public Opinion Surveys Online, paper presented at the 1999 AAPOR Conference, St. Petersburg, Florida.
- [21] Groves, R. M., Fowler, F. J., Couper, M. P. (2004) *Survey Methodology*. Hoboken: John Wiley and Sons, Inc.
- [22] Groves, R. M., Couper, M. (1998) *Nonresponse in Household Interview Surveys*. John Wiley and Sons, Inc.: New York, NY.
- [23] Groves, R. M., Dillman, D. A., Eltinge, J. L., Little, R. A. (2002) *Survey Nonresponse*. New York: John Wiley and Sons, Inc.
- [24] Groves, R., Cialdini, R., Couper, M. (1992) “Understanding the Decision to Participate in a Survey”, *Public Opinion Quarterly*, Vol. 56, No. 4, Winter, pp. 475–95.
- [25] Groves, R., Peytcheva, P. (2008) “The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis.” *Public Opinion Quarterly* 72(2):167–89.
- [26] Groves, R., Singer, E., Corning, A. (2000) Leverage-saliency Theory of Survey Participation: Description and an Illustration. *Public Opinion Quarterly*, Vol. 64, pp. 299–308.
- [27] Gunn, W., Rhodes, I. (1981) Physician Response Rates to a Telephone Survey: Effects of Monetary Incentive Level. In: *Public Opinion Quarterly*, Vol. 45, pp. 109–115.
- [28] Järnbert, M., Öhrvall, R. (2012) *Effects of Nonresponse on Survey Estimates of Political Participation*. Paper presented at the European Conference on Quality in Official Statistics. Athens, Greece

- [29] Keeter, S., Miller, C., Kohut, A., Groves, R. M. and Presser, S. (2000) *Consequences of reducing nonresponse in a national telephone survey*. *Public Opinion Quarterly*. Vol. 64, pp. 125–148.
- [30] Kruskal, W., Mosteller, F. (1979) Representative sampling III: Current statistical literature. *International Statistical Review*, 47, 245–265.
- [31] Krejčí, J. (2008) *Kvalita sociálněvědních výběrových šetření v České republice*. Praha: Sociologické Nakladatelství.
- [32] Lohr, S. L. (2010) *Sampling: Design and Analysis*. Boston: Brooks/Cole.
- [33] Lessler, J. T., Kalsbeek, W. D. (1992) *Nonsampling Error in Surveys*. John Wiley and Sons, Inc.: New York, NY.
- [34] Madow, W. G., Nisselson, H. and Olkin, I. (eds) (1983) *Incomplete Data in Sample Surveys*. Vol 3. New York Academic Press.
- [35] Mehta, R., Sivadas, E. (1995) Comparing Response Rates and Response Content in Mail Versus Electronic Mail Surveys. In: *Journal of the Market Research Society*, Vol. 37, pp. 429–439.
- [36] Merkle, D. M., and Edelman, M. (2002) Nonresponse in exit polls: A comprehensive analysis. In *Survey Nonresponse* (Eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A Little). New York: John Wiley & Sons, Inc., 243–258.
- [37] Särndal, C. E., Lundström, S. (2005) *Estimation in Surveys with Nonresponse*. Chichester: Wiley.
- [38] Sheehan, K. B., Hoy, M. G. (1999) Using E-mail to Survey Internet Users in the United States: Methodology and Assessment. In: *Journal of Computer Mediated Communication*, Vol 4, p. 3.
- [39] Schouten, B. (2004) Adjustment for bias in the Integrated Survey on Household Living Conditions (POLS) 1998. Discussion paper 04001, CBS, Voorburg
- [40] Steeh, C. (1981) Trends in Response Rates. *Public Opinion Quarterly*, 45, 40–57.
- [41] Thomsen, I. and Siring, E. (1983) On the Causes and Effects of Nonresponse: Norwegian Experiences In: W. G. Madow and I Olkin (eds), *Incomplete Data in Sample Surveys*, Vol. 3. New York; Academic Press.
- [42] Thomsen, I., Kleven, O., Wang, J. H., Zhang, L. C. (2006) *Coping with Decreasing Nonresponse Rates in Statistics Norway*. Oslo, Statistics Norway.
- [43] Traugott, M. W. (1987) The Importance of Persistence in Respondent Selection for Preelection Surveys In: *Public Opinion Quarterly*, 51, 48–57.

- [44] Voigt, L. F., Koepsell, T. D. and Daling, J. R. (2003) Characteristics of Telephone Survey Respondents According to Willingness to Participate In: *American Journal of Epidemiology*, 157, 66–73.
- [45] Willke, J., Adams, C., Girnius, Z. (1999) *Internet Testing. A Landmark Study of the Differences Between Mail Intercept and On-line Interviewing in the United States*. In: *Proceeding of the Worldwide Internet Seminar 1998 in Paris, France, Amsterdam: ESOMAR*, pp. 145–157.