

Maria Králová

# Metody kauzální analýzy pro měření efektivity podpory prodeje



MASARYKOVA  
UNIVERZITA



Maria Králová

METODY KAUZÁLNÍ ANALÝZY  
PRO MĚŘENÍ EFEKTIVITY  
PODPORY PRODEJE

MUNI  
PRESS

MUNI  
ECON



Maria Králová

---

# **METODY KAUZÁLNÍ ANALÝZY PRO MĚŘENÍ EFEKTIVITY PODPORY PRODEJE**

Methods of causal analysis for measuring  
the sales support efficiency

**Masarykova univerzita**

Brno 2021

## KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Králová, Maria

Metody kauzální analýzy pro měření efektivity podpory prodeje = Methods of causal analysis for measuring the sales support efficiency / Maria Králová. – Vydání 1. – Brno : Masarykova univerzita, 2021. – 1 online zdroj

Anglické resumé

Obsahuje bibliografii a bibliografické odkazy

ISBN 978-80-210-9940-1 (online ; pdf)

\* 658.8 \* 339.187 \* 658.8:316.77 \* 519.86/.87 \* (048.8)

– prodej

– marketingová komunikace

– ekonomická efektivnost

– modelování a simulace

– kauzální analýza

– monografie

658 – Řízení a správa podniku [4]

Knihu recenzovali:

prof. Radim Jiroušek, DrSc.

doc. RNDr. Lenka Komárková, Ph.D.



Knih je šířena pod licencí CC BY-NC 4.0 (Attribution-NonCommercial 4.0)

© 2021 Masarykova univerzita

ISBN 978-80-210-9940-1 (online ; pdf)

ISBN 978-80-210-9939-5 (brožováno)

<https://doi.org/10.5817/CZ.MUNI.M210-9940-2021>

# Obsah

<b>1 Úvod</b>	<b>9</b>
1.1 Motivace pro odhady efektivity podpory prodeje . . . . .	9
1.2 Efektivita podpory prodeje . . . . .	10
1.2.1 Podpora prodeje . . . . .	10
1.2.2 Vyhodnocení efektu podpory prodeje . . . . .	11
1.3 Společnosti typu Groupon . . . . .	13
1.4 Využití kauzální analýzy v oblasti marketingu . . . . .	14
1.5 Kauzální analýza . . . . .	15
1.5.1 Randomizované kontrolované experimenty a observační studie . . . . .	16
1.5.2 Kauzální inference v kontextu observačních studií . . . . .	17
<b>2 Rubinův kauzální model</b>	<b>21</b>
2.1 Základní pojmy . . . . .	21
2.2 Potenciální odezvy . . . . .	22
2.3 Mechanismus přiřazování ošetření . . . . .	24
2.4 Parametry v rámci RCM . . . . .	27
2.5 Kauzalita před Rubinem . . . . .	29
<b>3 Metody pro regulární mechanismus</b>	<b>31</b>
3.1 Základní předpoklady . . . . .	31
3.2 Předpoklady identifikace . . . . .	33
3.3 Klasické odhady . . . . .	34
3.4 Subklasifikace . . . . .	35
3.5 Párování . . . . .	36
3.5.1 Podobnost při párování . . . . .	38
3.5.2 Párování na základě propenzitních skóre . . . . .	39
3.6 Využití propenzitních skóre . . . . .	41
3.6.1 Stratifikace pomocí propenzitních skóre . . . . .	41

---

3.6.2	Vážení propenzitními skóry . . . . .	42
3.6.3	Marginální strukturální modely . . . . .	45
<b>4</b>	<b>Metody pro neregulární mechanismus</b>	<b>49</b>
4.1	Instrumentální proměnné . . . . .	49
4.1.1	Model s binárním instrumentem a binárním ošetřením	50
4.1.2	Koncept potenciálních odezev v modelech s binární instrumentální proměnnou . . . . .	52
4.1.3	Předpoklady pro kauzální inferenci v modelech s instrumentální proměnnou . . . . .	53
4.1.4	Odhad efektu $W$ na $Y$ v modelu s binární instrumentální proměnnou . . . . .	56
4.1.5	Interpretace efektu $\tau_{LATE}$ . . . . .	60
4.1.6	Podmíněný instrument . . . . .	60
4.2	Regresní diskontinuita . . . . .	61
4.2.1	Úvod k modelům regresní diskontinuity . . . . .	61
4.2.2	Model ostré regresní diskontinuity . . . . .	62
4.2.3	Model neostré regresní diskontinuity . . . . .	65
4.2.4	Interní a externí validita . . . . .	68
4.2.5	Odhad efektu ošetření v modelech s ostrou i neostrou regresní diskontinuitou . . . . .	69
<b>5</b>	<b>Aplikace při regulárním mechanismu</b>	<b>71</b>
5.1	Vliv akce „Nakup nad“ na zisky . . . . .	71
5.1.1	Popis datového souboru . . . . .	77
5.1.2	Základní předpoklady identifikace efektu ošetření . . . . .	79
5.2	Párování na základě Mahalanobisovy metriky . . . . .	84
5.2.1	Vyváženost párování . . . . .	85
5.2.2	Odhady efektů . . . . .	92
5.3	Párování na základě propenzitních skóre . . . . .	94
5.3.1	Odhady propenzitních skóre . . . . .	96
5.3.2	Vyváženost párování . . . . .	97
5.3.3	Odhady efektů . . . . .	103
5.4	Vážení propenzitními skóry . . . . .	104
5.5	Interpretace odhadů, diskuze a omezení . . . . .	110



---

5.5.1	Co lze usuzovat na základě observačních dat? . . . . .	112
5.5.2	Co bylo zjištěno kauzální analýzou? . . . . .	113
5.5.3	Omezení pro aplikaci použitých metod . . . . .	117
<b>6</b>	<b>Aplikace při neregulárním mechanismu</b>	<b>121</b>
6.1	Vliv akce „Pokladovka“ na útraty I . . . . .	124
6.1.1	Popis datového souboru . . . . .	124
6.1.2	Předpoklady identifikace efektu ošetření pomocí instru- mentální proměnné . . . . .	125
6.1.3	Odhady efektu pomocí instrumentální proměnné . . . . .	127
6.2	Vliv akce „Pokladovka“ na útraty II . . . . .	134
6.2.1	Popis datového souboru . . . . .	134
6.2.2	Předpoklady identifikace efektu ošetření pomocí instru- mentální proměnné . . . . .	135
6.2.3	Odhady efektu pomocí instrumentální proměnné . . . . .	136
6.3	Interpretace odhadů, diskuze a omezení . . . . .	140
6.3.1	Interpretace odhadů a diskuze . . . . .	141
6.3.2	Omezení pro aplikaci metody instrumentální proměnné . . . . .	145
<b>7</b>	<b>Závěr</b>	<b>147</b>
	<b>Literatura</b>	<b>155</b>
	<b>Přílohy</b>	<b>165</b>
<b>A</b>	<b>Instrumentální proměnné</b>	<b>167</b>
A.1	Tradiční ekonometrický přístup . . . . .	167
A.2	Moderní přístup pomocí potenciálních odezev . . . . .	170
<b>B</b>	<b>Balíky pro kauzální analýzu v R</b>	<b>173</b>



## Kapitola 1

# Úvod

### 1.1 Motivace pro odhady efektivity podpory prodeje

*„Half the money I spend on advertising is wasted; the trouble is, I don't know which half.“*

(John Wanamaker, 1838–1922)

*„Polovina peněz, které utrácím za reklamu, je vynaložena zbytečně; potíž je v tom, že nevím, která to je.“*

Náklady společností na aktivity spojené s podporou prodeje jsou v rozvinutých ekonomikách čím dál vyšší, ovšem podle výzkumu [Boston Consulting Group \(2015\)](#) 20 až 50 % podpůrných aktivit generuje zanedbatelný, žádný, či dokonce záporný zisk. Podle studie [Magid a Lodish \(1990\)](#) se dokonce vyplatilo pouze 16 % ze zkoumaných podpůrných aktivit. Studie profitability podpory prodeje od [Ailawadi, Harlam, Cesar a Trounce \(2006\)](#) zase ukázala, že ziskových aktivit na podporu prodeje je méně než 50 %. Přitom náklady výrobců na podporu prodeje neustále rostou, např. u výrobců produktů denní spotřeby a rychloobrátkového zboží činí dokonce dle [Gómez, Rao a McLaughlin \(2007\)](#) až 70 % rozpočtu na marketing.

Vyvstává tedy otázka, jestli a jak měří výrobci i prodejci efektivitu jednotlivých aktivit na podporu prodeje. Je zřejmé, že jejich vyhodnocení jen na základě intuice nebo tradice nestačí a pro posouzení výhodnosti je potřebný analytický přístup. Přitom dostupnost dat pro kompetentní analýzu přestává být pro společnosti problémem ([Blattberg a Briesch, 2012](#)). Problém tkví spíše v nedostatečném povědomí manažerů o možných metodách evaluace efektů podpůrných aktivit a nedostatečném využívání těchto metod.

Cílem této publikace je představit nový možný přístup k měření efektivity některých typů akcí na podporu prodeje (s ohledem na ukazatele

ziskovosti) pomocí metod kauzální analýzy a zároveň ukázat možnosti využití metod kauzální analýzy v oblasti marketingu.

Pro měření efektivity akcí na podporu prodeje se ukazuje jako zásadní odhad základních tržeb (podrobněji v 1.2.2), publikace představuje nový pohled na to, jak základní tržby odhadovat.

## 1.2 Efektivita podpory prodeje

### 1.2.1 Podpora prodeje

Podpora prodeje (sales promotion) je součást marketingové komunikace, jejímž cílem je v krátkodobém horizontu stimulovat nákup určitých produktů či služeb. [Kotler a Keller \(2011\)](#) definují podporu prodeje jako komplex různých motivačních nástrojů především krátkodobé povahy, jejichž účelem je podnítit rychlejší nebo větší nákupy určitých produktů nebo služeb. Od reklamy se podpora prodeje odlišuje v tom smyslu, že reklama nabízí spíše důvody ke koupi, zatímco podpora prodeje podněcuje ke koupi.

- Podle charakteru společnosti lze rozlišit následující směry působení podpůrných aktivit ([Blattberg a Briesch, 2012](#)):
  - od výrobce k prodejcům a dalším zprostředkovatelům (trade promotions)
  - od výrobce ke konzumentům (consumer promotions)
  - od prodejce ke konzumentům (retailer promotions)

V této publikaci budou řešeny úlohy spojené s odhadem efektů podpory prodeje ve směru působení od prodejce ke konzumentům.

- S ohledem na různé cíle i směr působení lze rozlišit různé formy podpory prodeje. Mezi nejčastější formy patří ([Blattberg a Briesch, 2012](#)) slevy, větší balení, kupony, akce typu „dva za cenu jednoho“, zboží na zkoušku, soutěže a loterijní hry či rabaty.
- Dopady akcí na podporu prodeje lze posuzovat nejen podle okamžitého či krátkodobého dopadu. Je potřeba zohlednit i dopad v delším časovém období. Proto podle trvání dopadu akce na tržby rozlišují [Gedenk, Neslin a Ailawadi. \(2010\)](#) krátkodobý efekt akce na podporu prodeje a dlouhodobý efekt. Krátkodobý efekt se vztahuje na období

trvání akce samotné a obvykle vede k nárůstu tržeb. Tento nárůst tržeb je ale částečně způsoben předzásobením, což může vést k propadu tržeb v následujícím období, proto je potřeba vyhodnocovat i dlouhodobý dopad této akce na tržby. V [Gedenk, Neslin a Ailawadi \(2010\)](#) je pak v časovém kontextu uvedena i podrobnější klasifikace různých dalších důsledků podpory prodeje (změna značky, změna nákupní kategorie, získání nových zákazníků, již zmíněné předzásobením, ale i loajalita značce apod.).

### 1.2.2 Vyhodnocení efektu podpory prodeje

Přestože je přirozené očekávat, že společnosti efekty podpory prodeje analyzují, ukazuje se, že vyhodnocování efektů je náročný a komplexní úkol, se kterým se prodejci obtížně vyrovnávají ([Ailawadi a Gupta, 2014](#)). Ač do podpory prodeje hodně investují, není jim zcela jasné, co jim tyto aktivity přinášejí. Dle studií [Boston Consulting Group \(2014\)](#) a [Boston Consulting Group \(2018\)](#) většina prodejců (mimo několika technologicky zdatných velkých společností) nezavedla do svých vyhodnocovacích procesů pokročilou analýzu, které věk velkých dat umožňuje. Často se místo analýzy dat spoléhají na taktiku z minulých let či doporučení dodavatelů, a to bez hlubšího vhledu a schopnosti podpůrné aktivity systematicky vyhodnocovat. Proto [Davenport, Harris a Morison \(2010\)](#) doporučují neupřednostňovat při rozhodování „dobrý pocit“, intuici, „dobrou víru“ či zvěsti a raději využívat vědecké metody založené na kvantitativních analýzách (ačkoliv i intuitivní přístup může do jisté míry přispět k porozumění kontextu). [Brynjolfsson, Hitt a Kim \(2011\)](#) ve své studii prokázali, že společnosti, jež svá rozhodnutí zakládají na datové analytice (přístup „DDD“, data driven decision making), jsou výkonnější. Proto je dle [McAfee a Brynjolfsson \(2012\)](#) potřeba, aby se společnosti více orientovaly na datovou analytiku firemních dat, jichž mají obvykle dostatek, a v managementu aplikovaly DDD přístup.

Příčiny potíží prodejců s kvantitativním vyhodnocováním dat na podporu prodeje spatřují [Breuer, Elliott a Rickert \(2013\)](#) ze společnosti McKinsey&Company v metodách vyhodnocování – pro sofistikované metody nemají prodejci kvalifikované analytiku a na druhé straně lehce uchopitelné jednoduché metody jsou ve svých výsledcích zavádějící. V této souvislosti

představují přínos takové metody, které jsou pro uživatele srozumitelné, přesto jsou v odhadování efektů spolehlivé.

Při vyhodnocování přínosů podpory prodeje pro tržby je zásadní znalost toho, jaké by tržby byly, kdyby se konkrétní aktivita na podporu prodeje nekonala. Tyto tržby jsou v literatuře označovány jako základní tržby (baseline sales) (Blattberg, Kim a Ye, 1996) a jejich srovnáním se skutečnými tržbami po dobu konání aktivity na podporu prodeje (nebo v jinak definovaném období) se odhaduje efekt této aktivity. Vzhledem k tomu, že hypotetické tržby – tedy tržby, jaké by byly, kdyby se podpůrná aktivita nekonala – jsou nepozorovatelné, je možné je pouze odhadovat.

Bohužel ve srovnání s tím, jak masivně je podpora prodeje používána, metody odhadu základních tržeb nejsou v literatuře dostatečně popsány a používané metody se soustřeďují zejména na porovnávání tržeb mezi obdobími před zahájením podpůrné aktivity, během konání této aktivity a po jejím ukončení (pre-post analysis) (Nagyová, 1999). Období před aktivitou podpory prodeje může být definováno různě, a tak se při odhadování základních tržeb vychází z tržeb v týdnu před zahájením aktivit na podporu prodeje, z tržeb ze stejného období v předchozím roce nebo z průměru tržeb za předešlý rok a podobně. Je zřejmé, že tyto tržby jsou ovlivněny mnoha dalšími vlivy včetně jiných aktivit na podporu prodeje a jako odhad základních tržeb jsou velmi nepřesné. Proto např. Ailawadi, Harlam, Cesar a Trounce (2006) na základě práce Abraham a Lodish (1993) odhadují základní tržby pomocí klouzavých průměrů prodejů v těch týdnech, které byly časově blízko k týdnu s podpůrnou aktivitou a ve kterých žádná podpůrná aktivita neprobíhala. Pauwels, Hanssens a Siddarth (2002) a Ataman, Mela a Van Heerde (2008) odhadují základní tržby již pomocí pokročilých dynamických lineárních modelů, ovšem podle Jetta a Rengifo (2011) jsou tyto odhady vhodné spíše pro specifické akademické aplikace.

Ať jsou již v rámci pre-post analýzy základní tržby odhadovány jakkoliv, všechny jsou více či méně problematické kvůli vychýlení, které je způsobeno množstvím nekontrolovaných vlivů, jež se v čase mění. Tyto vlivy mohou tržby zvyšovat či snižovat, a tak jsou odhady základních tržeb pro následné srovnávání s tržbami, které byly pozorovány v období konání podpůrné aktivity, zkreslené. Pre-post analýza tak například nemá pod kontrolou chování konkurence v období před marketingovou akcí, během ní a po ní,

změny portfolia produktů v čase, dlouhodobý dopad předchozích kampaní, vnější ekonomické vlivy a další.

Jiný přístup k odhadování základních tržeb představují experimenty<sup>1</sup>, ovšem ne pro každou aktivitu na podporu prodeje je experimentální design použitelný. Navíc společnosti nerady podstupují riziko, že by v rámci experimentálního designu přišly o možné zisky z kontrolní skupiny. Proto jsou metody pro odhad efektu podpory prodeje založené zejména na observačních datech pro praktické aplikace užitečné.

V této publikaci bude odhad efektu vybraných aktivit na podporu prodeje modelován pomocí kauzální analýzy na observačních datech české společnosti typu Groupon provozující slevový portál.

### 1.3 Společnosti typu Groupon

Společnosti typu Groupon provozují slevové portály, kde obvykle se slevou přeprodávají zboží a služby od mnoha dalších výrobců či poskytovatelů služeb. Fungují na principu „sleva dne“ (deal-of-the-day). Zajímavých slev pro nakupující, tedy uživatele svého portálu, mohou dosáhnout tehdy, když zajistí dostatečný počet nákupů tak, aby se sleva (jež má charakter především množstevní slevy) výrobcí či poskytovateli služeb vyplatila. Proto musí své uživatele neustále povzbuzovat ke sledování nových nabídek zboží a služeb a následně k jejich koupi. K tomuto účelu slouží aktivity na podporu prodeje, které by měly zvedat tržby, budovat věrnost aktuálních uživatelů a získávat uživatele nové.

Zatímco v případě podpory prodeje u výrobců jde zejména o posílení značky, u prodejců jde o nárůst tržeb v celé kategorii či kategoriích produktů. („Brand switching“ může způsobit nárůst tržeb za produkt v akci na úkor prodejů podobných produktů ze stejné kategorie produktů, které však v akci nejsou, proto jde prodejcům o nárůst tržeb za celou kategorii produktů.) Slevové portály pak musejí pomocí podpory prodeje zajistit zejména dostatečný počet nakupujících, aby filozofie množstevních slev fungovala. Zisk slevových portálů je pak tvořen rozdílem mezi cenou zaplacenou uživatelem a již zlevněnou cenou přeprodávaného zboží či služeb. Marže je tedy již tak malá, takže další náklady na podporu

---

<sup>1</sup> Kontext experimentálních a observačních dat je blíže popsán v podkapitole [1.5.1](#).

prodeje mohou být pro ziskovost společností provozujících slevové portály kritické.

V této publikaci bude analyzován nárůst tržeb (či dalších ukazatelů ziskovosti) v důsledku provedení dvou typů aktivit na podporu prodeje u české společnosti provozující slevový portál. Tato společnost bude označována jako společnost  $F$ . Analyzované podpůrné aktivity jsou podrobně popsány v kapitolách 5 a 6. Odhad základních tržeb (a z nich odvozeného efektu akcí na podporu prodeje) nebude odvozen z tržeb předcházejících období konání podpůrných aktivit, jak je obvyklé v pre-post analýzách. Místo toho budou tržby – jaké by byly, kdyby se podpora prodeje nekonala – simulovány pomocí kauzální analýzy pro stejný čas, kdy se podpora prodeje konala. Tím bude odstraněn problém dalších vlivů, které se v čase mění, působí na tržby, a které pre-post analýzy nemají pod kontrolou.

## 1.4 Využití kauzální analýzy v oblasti marketingu

Ač jsou v oblasti ekonometrie kauzální modely známé, v oblasti marketingu a managementu je povědomí o nich velmi nízké a autorce není známo, že by byly využity pro odhady základních tržeb.

Dosavadní opomíjení kauzálních metod v oblasti marketingu a podpory prodeje reflektují také zakladatel kauzální analýzy D. B. Rubin a R. P. Waterman ve společné práci (Rubin a Waterman, 2006) ve speciálním vydání časopisu *Statistical Science, A Special Issue on Statistical Challenges and Opportunities in Electronic Commerce Research*. V práci mimo jiné upozorňují na nesprávné využití klasických statistických metod (regrese, metody data miningu apod.) k odhadování efektů reklamy a podpory prodeje a uvádějí možnosti, jak související problémy řešit pomocí metod kauzální analýzy.

Příkladem mála prací, jež využívají modely kauzální analýzy v oblasti marketingu, je např. aplikace modelů regresní diskontinuity v oblasti vyhodnocení vybraných marketingových aktivit (Hartmann, Nair a Narayanan, 2011) či využití párování na základě propenzitních skóru a rozdílů v rozdílech při vyhodnocování chování zákazníků v reakci na streamovací služby (Datta, Knox a Bronnenberg, 2018). Obecně jsou problémy spojené s endogenitou v modelech strategického managementu popsány v Hamilton a Nickerson (2003).



## 1.5 Kauzální analýza

„*Felix, qui potuit rerum cognoscere causas.*“

(Vergilius, 29 př. Kr.)

„*Štasten, kdo mohl poznat příčiny věcí.*“

Identifikace kauzálních souvislostí a odvozování kauzálních inferencí pomocí experimentálních dat je probádaná oblast usuzování s dlouhou tradicí a širokým uplatněním zejména v přírodních a technických vědách. V případě společenských věd se kauzální inference na základě experimentálních dat uplatňuje zejména v těch oblastech, kde je možné plánovat randomizované experimenty, nebo v případech, kdy dojde k *přirozeným experimentům* (natural experiments) (Imbens a Rubin, 2015). V managementu, marketingu a v řadě dalších ekonomických disciplín jsou ale randomizované experimenty buď zcela nemožné, nebo těžko proveditelné a k dispozici jsou často pouze data z observačních studií. Kauzální usuzování na základě čistě statistických nástrojů pak není možné, a často se tedy alespoň sledují různé míry asociace. Statistická i ekonometrická literatura je bohatá na mnohé technicky různě náročné míry asociace vytvořené zejména na základě různých koeficientů korelace. Ovšem žádnou míru asociace nelze interpretovat jako „důkaz“ kauzální souvislosti, a tak může být vysoká míra asociace pouze podnětem k hlubšímu zkoumání možné kauzality.

Kauzální vliv proměnné  $W$  na proměnnou  $Y$  se zcela jistě projeví jejich velkou (resp. aspoň nenulovou) mírou asociace. Na druhé straně však velká míra asociace mezi proměnnými  $W$  a  $Y$  nutně neznamená kauzalitu. Může být způsobena jinou proměnnou  $X$  (či jinými proměnnými), která kauzálně působí jak na proměnnou  $W$ , tak na proměnnou  $Y$  na pozadí toho, co je pozorovatelné. A ačkoliv spolu proměnné  $W$  a  $Y$  nemusejí kauzálně souviset, v důsledku závislosti obou proměnných na proměnné  $X$  je jejich vzájemná míra asociace vysoká. V takovémto případě jde o tzv. falešnou korelaci (spurious correlation) proměnných  $W$  a  $Y$  a proměnná  $X$  se označuje jako matoucí či *zavádějící* proměnná (confounding variable, confounder). Samotné míry asociace jsou tedy v kontextu kauzálního usuzování nezajímavé a zejména v případě observačních studií je nutné hledat takové přístupy, které umožní kauzální usuzování nad rámec toho, co je možné řešit pomocí statistiky samotné.

Pomocí kauzální inference lze zodpovědět otázku, co **by se stalo** se závislou proměnnou  $Y$  v důsledku **hypotetické** změny nějaké vysvětlující proměnné  $W$ . Tato vysvětlující proměnná  $W$  se v kontextu kauzální analýzy nazývá *ošetření* (treatment) a závislé proměnné  $Y$  se říká *odezva* (response, outcome). Ačkoliv je označení „treatment“ převzato z aplikací v medicíně, reprezentuje jakoukoliv kategoriální či spojitou proměnnou, která může mít vliv (effect) na odezvu. Jedná se tedy např. o intervenci, vystavení subjektů působení nějakého vlivu apod. Odezva i ošetření může být každá jak kategoriální, tak i spojitá proměnná.

Studium kauzální analýzy v kontextu různých ekonomických disciplin je výzkumně živá oblast (Moneta a Russo, 2014). V této publikaci bude kauzální analýza využita v oblasti marketingu pro odhadování vlivu akcí na podporu prodeje (ošetření  $W$ ) na zisky (odezva  $Y$ ) či další z pohledu společnosti  $F$  zajímavé odezvy, přičemž bude kontrolován vliv možných zavádějících proměnných. Analogicky pak lze kauzální analýzu využít i v dalších podobných úlohách z oblasti managementu, marketingu a dalších ekonomických věd.

### 1.5.1 Randomizované kontrolované experimenty a observační studie

Pro odhad příčinného vlivu vysvětlující proměnné  $W$  (ošetření; intervence; vystavení subjektů působení něčeho) na odezvu  $Y$  jsou za zlatý standard považovány *randomizované kontrolované experimenty* (Greenland, Pearl a Robins, 1999).

Je-li například ošetření  $W$  binární proměnná (lze zobecnit i na víceúrovňovou kategoriální proměnnou či kategorizovanou spojitou proměnnou), lze pak srovnáním odezvy  $Y$  u skupiny subjektů s aktivním ošetřením ( $W = 1$ ) se skupinou subjektů s kontrolním ošetřením (t.j. bez ošetření,  $W = 0$ ) získat představu o vlivu (efektu) proměnné  $W$  na  $Y$ . Takto získaný rozdíl v odezvách obou skupin lze považovat za efekt ošetření pouze tehdy, když jsou obě skupiny subjektů (skupina s aktivním i skupina s kontrolním ošetřením) zcela srovnatelné s ohledem na zavádějící proměnné. Pokud nejsou v tomto ohledu srovnatelné, pak by rozdíl v odezvě mezi skupinami mohl být způsoben i jinými příčinami než ošetřením  $W$ . Proto jsou subjekty v rámci randomizovaného kontrolovaného experimentu přiřazo-

vány do skupiny s aktivním a skupiny s kontrolním ošetřením **náhodně**. Jinými slovy subjekt, který byl zařazen do skupiny s aktivním ošetřením, mohl být se stejnou pravděpodobností (nebo výzkumníkem nastavenou pravděpodobností) zařazen i do skupiny s kontrolním ošetřením.

V kontrolovaných experimentech tedy o příslušnosti subjektu ke skupině rozhoduje výzkumník (a nemusí nutně použít randomizovaný mechanismus přiřazení subjektu ke skupině). V *observačních studiích* naopak rozhoduje o své příslušnosti ke skupině subjekt sám (nebo se subjektem svázaný mechanismus). Výzkumník je pak pouze v roli pozorovatele. Rozdíl v odezvě  $Y$  mezi skupinou s aktivním ošetřením a skupinou s kontrolním ošetřením nemusí být tedy způsoben jenom efektem ošetření  $W$  (pokud vůbec existuje), ale i vlivem zavádějících proměnných  $X = (X_1, \dots, X_k)'$ , které ovlivňují jak příslušnost ke skupině (proměnnou  $W$ ), tak zároveň i odezvu (proměnnou  $Y$ ). Zavádějící proměnné pak představují (hlavně při jejich odlišné distribuci v jednotlivých skupinách) jeden z hlavních problémů spojených s observačními studii. Proto je v observačních studiích při zkoumání vlivu ošetření  $W$  na odezvu  $Y$  nutné kontrolovat i působení těchto zavádějících proměnných.

V kapitolách 5 a 6 bude v roli ošetření  $W$  účast uživatele společnosti  $F$  v akci na podporu prodeje, v roli odezvy  $Y$  např. zisk společnosti připadající na zákazníka po dobu konání akce a v roli zavádějících kovariát<sup>2</sup> různé charakteristiky zákazníka, např. jak často otevírá propagační emaily, jak často navštěvuje stránky společnosti  $F$ , jestli čerpá výhody věrnostního programu, jak často nakupuje apod. Tyto proměnné totiž mohou ovlivňovat jak ochotu zákazníka účastnit se akce na podporu prodeje, tak i zisky, které pro společnost  $F$  po dobu konání akce generuje.

### 1.5.2 Kauzální inference v kontextu observačních studií

Povaha výzkumu kauzálních vztahů v mnohých ekonomických disciplínách neumožňuje provádět randomizované kontrolované experimenty, a tak jsou observační studie často jediným zdrojem informací pro výzkum v těchto aplikačních oblastech. Ve zkoumání kauzality v kontextu observačních studií převažuje v současné statistické i ekonometrické literatuře přístup

<sup>2</sup> Podrobněji je význam termínu kovariáta v kontextu kauzální analýzy popsán v kapitole 2.1.

definovaný Rubinem v sedmdesátých letech minulého století: Rubin (1973a), Rubin (1973b), Rubin (1974), Rubin (1977), který zavádí formální definici kauzálního vlivu (efektu) na základě *potencionálních odezev* (2.2). Název „Rubinův kauzální model“ (RCM) byl poprvé použit v osmdesátých letech minulého století v Holland (1986). Ačkoliv není idea potenciálních odezev nová, rostoucí zájem o kauzální inferenci zejména v kontextu observačních studií vedl v posledních letech k velkému nárůstu objevů a článků na toto téma. Jejich přehled je podán např. v Imbens (2004) a v Imbens a Wooldridge (2009).

V observačních studiích není jednotlivým subjektům datového souboru přiřazováno ošetření náhodně a ani není možné pozorovat odezvu jednotlivého subjektu  $i \in \{1, 2, \dots, N\}$  zároveň jak s ošetřením<sup>3</sup> ( $Y_i | W = 1$ ), tak bez ošetření ( $Y_i | W = 0$ ). Proto nelze kauzální vliv určit pouze na základě dat samotných (tedy výhradně na základě statistických metod)!

K identifikaci kauzálního vlivu je nutné formulovat a obhájit i tzv. předpoklady identifikace (podrobně v 3.2), které na základě hlubokého porozumění zkoumanému fenoménu odstraní veškerá zkreslení efektu např. zavádějícími proměnnými či dalšími zdroji zkreslení. Podstatnou součástí kauzální inference je tudíž přesvědčivý design výzkumu. Až následně je důležitá technická stránka řešení, tedy metody, které v rámci obhajitelného designu výzkumu vedou na základě dostupných dat k vhodným estimátorům zkoumaných efektů, což je již úkol statistiky.

Zkreslení efektu lze odstranit (s ohledem na použitý design výzkumu a předpoklady popsané dále) např. pomocí metod párování (matching), metody instrumentálních proměnných (instrumental variables), regresní diskontinuity (regression discontinuity), metody rozdílu v rozdílech (difference in differences) a dalších metod, z nichž některé budou blíže popsány v následujících kapitolách.

Většina anglických termínů vztahujících se k uvedeným metodám nemá zavedené české ekvivalenty, navíc ani anglická terminologie není zcela jednotná. Pro snadnější orientaci v používaných termínech je proto v tabulce 1.1 nabídnut jejich stručný přehled. Přehled používaných zkratk obsahuje tabulka 1.2.

---

<sup>3</sup> Pro jednoduchost je uvažováno binární ošetření, tedy  $W_i \in \{0, 1\}$ .

**Tabulka 1.1:** Přehled původních anglických termínů používaných v oblasti kauzální analýzy a jejich české ekvivalenty. Termíny *always-taker*, *complier* a *never-taker* lze volně přeložit jako „subjekty, kterým je bez ohledu na podnět instrumentu vždy přiřazeno aktivní ošetření“, „subjekty, kterým je ošetření přiřazeno v souladu s podnětem instrumentu“ a „subjekty, kterým je bez ohledu na podnět instrumentu vždy přiřazeno kontrolní ošetření“. V zájmu srozumitelnosti textu se budou v této publikaci používat původní anglické termíny.

Anglický termín	Český ekvivalent
<i>always-taker</i>	<i>bez českého ekvivalentu</i>
caliper	měřítko pro okolí
complier	<i>bez českého ekvivalentu</i>
confounding variable	zavádějící proměnná
counterfactual	kontrafaktuál
difference in differences (DiD)	rozdíly v rozdílech
encouragement design	design s povzbuzením
forcing variable <i>synonymum</i> : running variable	účinná proměnná
imbalance	nevyváženost
matching	párování
never-taker	<i>bez českého ekvivalentu</i>
one-to-one matching	párování „jeden na jednoho“
overlap; <i>synonyma</i> : common support, positivity assumption	překryv
pre-treatment variables	proměnné před ošetřením
regression discontinuity (RD)	regresní diskontinuita
selection bias	výběrové vychýlení
selfselection	samovýběr
Stable Unit Treatment Value Assumption (SUTVA)	předpoklad stabilní hodnoty ošetření subjektu
trimming	ořezávání
truncation	osekávání
unconfoundedness; <i>synonyma</i> : ignorability, selection on observables, exogeneity, conditional independence	nezavádějící mechanismus přiřazování

**Tabulka 1.2:** Přehled používaných zkratek

Zkratka	Celý název	Odkaz na stránku
IPTW	(Inverse Probability Treatment Weighting	42
ITT	Intent to Treat	51
MSM	Marginal Structural Model	45
RCM	Rubin Causal Model	21
RD	Regression Discontinuity	61
SMD	Standardized Mean Difference	86
SUTVA	Stable Unit Treatment Value Assumption	23

V kapitolách 2 až 4 budou dále stručně podány vybrané metody kauzální analýzy vhodné pro měření efektivity podpory prodeje. V kapitolách 5 až 7 budou formulovány výzkumné otázky spojené s měřením efektivity konkrétních akcí na podporu prodeje a následně bude podán podrobný návod k implementaci vhodných metod kauzální analýzy pro řešení formulovaných problémů. Součástí bude i pečlivé ověření předpokladů těchto metod a interpretací výsledků.

Praktická implementace těchto metod vyžaduje použití softwaru – ideálně takového, který již má procedury vhodné pro kauzální analýzu připravené. Pro tyto účely je vhodný jazyk **R**, pomocí kterého byly všechny analýzy a grafy (i většina ilustračních obrázků) v této publikaci provedeny. Přehled potřebných rozšiřujících balíčků jazyka **R** pro používané metody je v tabulce B.1 v příloze B.

## Kapitola 2

# Rubinův kauzální model (RCM)

### 2.1 Základní pojmy

Rubinův přístup ke kauzálnímu modelování je v této kapitole představen pro situaci, kdy ošetření, tedy proměnná  $W$ , nabývá pouze dvou hodnot,  $W \in \{0, 1\}$ . Aktivnímu ošetření odpovídá  $W = 1$  a kontrolnímu ošetření  $W = 0$ . (Může jít o účast/neúčast podniku v dotačním programu, či jeho držení/nedržení ISO certifikátu; nebo může jít o registrovaného/anonymního zákazníka konkrétního e-shopu či proškoleného/neproškoleného zaměstnance apod.) Rubinův přístup lze ale jednoduše zobecnit i na situace s víceúrovňovým ošetřením.

Efekt ošetření  $W$  na odezvu  $Y$  bude posuzován na základě souboru  $N$  subjektů indexovaných pomocí  $i \in 1, \dots, N$ , u kterých lze pozorovat odezvu  $Y_i$  i obdržené ošetření  $W_i$ . Symbol  $N_1$  označuje počet těch subjektů, které obdržely aktivní ošetření, a symbol  $N_0$  označuje počet těch subjektů, které obdržely kontrolní ošetření.  $\mathbf{Y} = (Y_1, \dots, Y_N)'$  označuje vektor odezev pro celý soubor.

Kromě ošetření a odezvy lze u subjektů dále pozorovat i hodnoty kovariát  $X = (X_1, \dots, X_k)'$ , přičemž symbol  $\mathbf{X}$  bude označovat  $N \times k$  rozměrnou matici hodnot kovariát všech subjektů souboru. Kovariátou může být například velikost podniku či obor podnikání, pokud se jedná o soubor podniků; pokud se jedná o soubor zaměstnanců či zákazníků, pak kovariátou může být věk, dosažené vzdělání, pohlaví apod. V těchto případech jde o trvalé charakteristiky subjektů. Za kovariáty budou považovány i tzv. *proměnné před ošetřením* (pre-treatment variables), které jsou v čase proměnlivé. Proměnnou před ošetřením může být tedy například konkrétní finanční ukazatel podniku **před** zavedením nějaké intervence.

Za kovariáty (v souladu s konvencemi kauzální analýzy) nelze považovat proměnné po ošetření (post-treatment), proto se v této publikaci

pojem kovariáty vztahuje výlučně k proměnným stabilního charakteru a proměnným před ošetřením.

## 2.2 Potenciální odezvy

V kontextu RCM je kauzální efekt ošetření  $W$  na odezvu  $Y$  u  $i$ -tého subjektu<sup>4</sup> definován jako rozdíl mezi oběma potenciálně možnými odezvami, které **by** mohly nastat, kdyby subjekt obdržel aktivní i kontrolní ošetření. Tedy pro každý jednotlivý subjekt  $i \in 1, \dots, N$  lze hypoteticky uvažovat dvě *potenciální odezvy*:  $Y_i(0)$ , což je odezva, kterou **by**  $i$ -tý subjekt nabyl při kontrolním ošetření, a  $Y_i(1)$ , kterou **by**  $i$ -tý subjekt nabyl při aktivním ošetření<sup>5</sup>.

Hlavní problém kauzální inference pak spočívá v tom, že z dvojice  $(Y_i(0), Y_i(1))$  je pozorovatelná vždy nejvýše jedna odezva. Tu „druhou“ z dvojice potenciálních odezev  $i$ -tého subjektu nazýváme její *kontrafaktuál* (counterfactual). Tedy je-li pozorovatelná například  $Y_i(1)$ , pak jejím kontrafaktuálem je  $Y_i(0)$ , která v datovém souboru představuje chybějící hodnotu (a analogicky pro pozorovatelnou  $Y_i(0)$ ). Problém kauzální inference tak tkví v chybějících hodnotách, jelikož kauzální efekt ošetření u  $i$ -tého subjektu  $Y_i(1) - Y_i(0)$  kvůli zmíněné chybějící hodnotě kontrafaktuálu není pozorovatelný. Kauzální inferenci tak lze v jistém smyslu chápat i jako predikční úlohu – co by se stalo s odezvou, kdyby subjekt obdržel jinou úroveň ošetření než pozorovanou. Je tedy potřeba predikovat kontrafaktuál.

Pozorovanou odezvu  $Y_i$  lze vyjádřit jako

$$Y_i = Y_i(W_i) = Y_i(0)(1 - W_i) + Y_i(1)W_i = \begin{cases} Y_i(0) & \text{pro } W_i = 0 \\ Y_i(1) & \text{pro } W_i = 1 \end{cases} \quad (2.2.1)$$

S ohledem na modely moderní statistické a ekonometrické kauzální analýzy je tedy zásadní rozlišovat mezi pozorovanou odezvou  $i$ -tého subjektu  $Y_i$  a dvojicí potenciálních odezev  $(Y_i(0), Y_i(1))$ .

Velkou a na první pohled zřejmou výhodou tohoto přístupu je skutečnost, že možná heterogenita v efektu ošetření na odezvu napříč subjekty  $i$  je

---

<sup>4</sup> V rámci jak experimentu, tak observační studie.

<sup>5</sup> V literatuře používané alternativní označení pro  $Y_i(0)$  je i  $Y_i^0$ , nebo  $Y_{0i}$ . Analogicky pro  $Y_i(1)$ .



podchycena ihned na začátku modelování. (Není tedy nutné předpokládat „fixní efekty“, jako v klasické regresi.) Dále pak pro definování samotného efektu není potřeba specifikovat konkrétní statistický model, který by efekt (efekty) odhadoval.

Vztah 2.2.1 implicitně předpokládá,

- 1.) že potenciální odezvy  $i$ -tého subjektu nejsou ovlivněny ošetřením jakéhokoliv jiného subjektu. Tedy symbolicky:

$$Y_i(W_1, W_2, \dots, W_N) = Y_i(W'_1, W'_2, \dots, W'_N) \quad \text{když } W_i = W'_i$$

- 2.) a dále, že neexistuje více verzí pro jednotlivé úrovně ošetření.<sup>6</sup>

Souhrnně se těmito předpoklady říká *předpoklad stabilní hodnoty ošetření subjektu* (Stable Unit Treatment Value Assumption – SUTVA<sup>7</sup>) (Rubin, 1980).

Pokud je některá ze dvou podmínek SUTVA porušena, nelze jednoznačně definovat potenciální odezvy, tudíž nelze definovat kauzální efekty a ani jejich odhady. Ovšem skutečnost, zda SUTVA v uvažované studii (či experimentu) platí, nelze prokázat pouze na základě dat. Platnost tohoto předpokladu lze obhájit pouze ze znalosti zkoumané problematiky. Proto je potřeba zejména v kontextu observačních studií koncipovat studii tak, aby byly předpoklady SUTVA obhajitelné.

Oblíbeným odhadovaným parametrem v kauzální analýze je *průměrný efekt ošetření* (average treatment effect)  $\tau_{ATE}$  definovaný pro celou populaci:

$$\tau_{ATE} = E[Y(1) - Y(0)] \quad (2.2.2)$$

Za intuitivní odhad  $\tau_{ATE}$  se v rámci náhodného výběru nabízí výběrový průměrný efekt ošetření:

$$\hat{\tau}_{ATE} = \frac{1}{N} \sum_{i=1}^N [Y_i(1) - Y_i(0)] = \frac{1}{N} \sum_{i=1}^N Y_i(1) - \frac{1}{N} \sum_{i=1}^N Y_i(0) \quad (2.2.3)$$

<sup>6</sup> Tento předpoklad se někdy nazývá předpoklad konzistence Robins (1986).

<sup>7</sup> SUTVA a předpoklad nezávislosti náhodných chyb v regresních modelech není totéž! Například metoda nejmenších čtverců by i při porušené nezávislosti náhodných chyb odhadla efekt ošetření nevychýleně, ovšem při porušeném SUTVA by byly odhady efektu metodou nejmenších čtverců vychýlené.

Ovšem vzhledem k nepozorovatelným kontrafaktuálům pozorovaných odezev je i  $\hat{\tau}_{ATE}$  nepozorovatelné. Za této situace je lákavé nahradit v 2.2.3

- výraz  $\frac{1}{N} \sum_{i=1}^N Y_i(1)$  výrazem  $\frac{1}{N_1} \sum_{i=1}^N W_i Y_i$ , jenž průměruje pouze pozorované odezvy při aktivním ošetření
- a výraz  $\frac{1}{N} \sum_{i=1}^N Y_i(0)$  výrazem  $\frac{1}{N_0} \sum_{i=1}^N (1 - W_i) Y_i$ , jenž průměruje pouze pozorované odezvy při kontrolním ošetření.

Zjišťovat však vliv efektu ošetření pouze porovnáním průměrovaných odezev mezi skupinou, která dostala aktivní ošetření, a skupinou, která dostala kontrolní ošetření, může být velmi zavádějící, jelikož není známo, jakým mechanismem bylo jednotlivým subjektům aktivní, či kontrolní ošetření přiřazeno.

Například efekt účasti zapojení v dotačním programu na ziskovost podniku nelze odhadovat pouze porovnáváním ziskovosti mezi skupinou podniků zapojených do programu a skupinou nezapojených. Je totiž možné, že tyto dvě skupiny podniků se liší i v hodnotách relevantních kovariát (např. velikost podniku), které mají vliv jak na zapojení podniku do dotačního programu, tak na ziskovost. Odhadování efektu dotace na ziskovost pouze porovnáním odezvy v obou skupinách je tedy s ohledem na skrytý vliv kovariát (zde velikost podniku) irelevantní.

Proto je nezbytnou součástí kauzální analýzy v rámci RCM také identifikace mechanismu, jenž rozhoduje, kterým subjektům bude přiřazeno aktivní a kterým kontrolní ošetření.

## 2.3 Mechanismus přiřazování ošetření

$N$ -dimenzionální vektory potenciálních odezev jednotlivých subjektů  $i = 1, \dots, N$  pro aktivní ošetření, potencionálních odezev pro kontrolní ošetření a samotné ošetření jsou označeny následně:

$$\mathbf{Y}(1) = (Y_1(1), \dots, Y_N(1)), \mathbf{Y}(0) = (Y_1(0), \dots, Y_N(0)), \mathbf{W} = (W_1, \dots, W_N).$$

Mechanismus přiřazování ošetření je procedura, která rozhoduje, který subjekt dostane aktivní a který kontrolní ošetření. Tato procedura tedy také rozhodne, která z potenciálních odezev bude ta skutečně pozorovaná. Technicky se jedná o funkci, jež libovolnému vektoru ošetření  $\mathbf{W}$  přiřadí,

podmíněně na všech kovariátách a potenciálních odezvách, jeho pravděpodobnost. Přesněji, za předpokladu, že  $N$  subjektů reprezentuje celou populaci, je *mechanismus přiřazování* taková pravděpodobnostní funkce  $P(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(1), \mathbf{Y}(0))$  s oborem hodnot  $\langle 0, 1 \rangle$ , pro kterou platí:<sup>8</sup>

$$\sum_{\mathbf{W} \in \{0,1\}^N} P(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(1), \mathbf{Y}(0)) = 1$$

pro libovolné  $\mathbf{X}, \mathbf{Y}(1), \mathbf{Y}(0)$ , přičemž na pořadí, v němž jsou subjekty ve vektorech uspořádány, nezáleží.

U mechanismu přiřazování je (v kontextu kauzální analýzy) žádoucí, aby měl následující vlastnosti:

1. Mechanismus přiřazování je *individualistický*:  
Pravděpodobnost, že uvažovanému subjektu bude přiřazeno aktivní ošetření, nezávisí ani na kovariátách ani na odezvách ostatních subjektů.
2. Mechanismus přiřazování je *probabilistický*:  
Pravděpodobnost přiřazení aktivní úrovně ošetření subjektu má u každého subjektu nenulovou pravděpodobnost.
3. Mechanismus přiřazování je *nezavádějící* (unconfounded):<sup>9</sup>  
Mechanismus přiřazování není závislý na potenciálních odezvách, tedy

$$P(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(1), \mathbf{Y}(0)) = P(\mathbf{W} \mid \mathbf{X}) \quad (2.3.1)$$

Pokud má mechanismus přiřazování všechny tři zmíněné vlastnosti, pak jsou obě skupiny subjektů – skupina subjektů s aktivním ošetřením a skupina subjektů s kontrolním ošetřením – za předpokladu kontroly na kovariáty velmi podobné (ideálně stejné) a označují se jako *vyvážené* (balanced). Potom lze efekt ošetření skutečně odhadnout přímým porovnáním odezev mezi oběma skupinami.

<sup>8</sup> Symbol  $\mathbf{W} \in \{0,1\}^N$  označuje všechny možné  $N$ -tice, jejichž prvky jsou jedničky a nuly, kterých může vektor  $\mathbf{W}$  nabýt. (Takových  $N$ -tic je  $2^N$ .)

<sup>9</sup> Rozložení dvojice potenciálních odezev  $(Y(1), Y(0))$  je pro každou úroveň ošetření stejné za předpokladu kontroly na kovariáty.

Pokud v důsledku zavádějícího mechanismu přiřazování ošetření s ohledem na kovariáty nejsou tyto skupiny podobné, jedná se o *výběrové vychýlení* (selection bias), jež následně vede k vychýlenému odhadu efektu ošetření. Toto vychýlení je způsobeno skutečností, že subjekty si „samy“ vybírají úroveň ošetření. V této souvislosti se také používá pojem *samovýběr* (selfselection).

Například je-li počet kvalifikovaných zaměstnanců podniku v roli kovariáty, která se podílí na mechanismu přiřazování ošetření (účast/neúčast v dotačním programu) a zároveň ovlivňuje ziskovost podniku, pak ve skupině podniků s dotací skončí jiný „druh“ podniků než ve skupině bez dotace. Skupiny pak nejsou vyvážené a odhad efektu účasti v dotačním programu na ziskovost podniku získaný prostým porovnáním ziskovosti obou skupin podniků je vychýlený.

V kauzální analýze jde tedy o to zajistit, aby byly obě skupiny s ohledem na kovariáty vyvážené. Nejjednodušší způsob, jak toho dosáhnout, je design kontrolovaného randomizovaného experimentu, který přiřazuje jednotlivým subjektům úroveň ošetření náhodně. Tento mechanismus přiřazování ošetření je individualistický, probabilistický a je nezavádějící (unconfounded), což plyne přímo z designu kontrolovaného randomizovaného experimentu. (Design randomizovaných experimentů lze dále ještě různě variovat například na zcela randomizovaný experiment, párový randomizovaný experiment či zcela stratifikovaný randomizovaný experiment apod., ale tyto designy se v ekonomických vědách až na výjimky nevyužívají.)

Bohužel v případě observačních studií není mechanismus přiřazování známý, a tedy ani není známo, zda je individualistický, probabilistický a nezavádějící. (Proto může prosté porovnání odezev mezi skupinami s různými úrovněmi ošetření vést k vychýlenému odhadu efektu ošetření.) Pokud tyto tři vlastnosti lze **předpokládat** (tedy neplynou přímo z designu experimentu tak jako u randomizovaných experimentů), pak mluvíme o *regulárním* mechanismu přiřazování. V opačném případě nazýváme mechanismus přiřazování *neregulárním*.

V rámci observačních studií je problematický zejména předpoklad třetí vlastnosti „nezavádějícího mechanismu přiřazování“. Proto lze metody, které umožní odhadnout efekty ošetření i v rámci observačních studií,

rozdělit na ty, které u mechanismu přiřazování umožní předpokládat, že je nezavádějící, a dále na ty, u kterých toto předpokládat nelze.

V prvním případě se jedná o mechanismy přiřazení, u kterých lze odhad efektu kontrolovat na zavádějící proměnné. Pokud všechny zavádějící proměnné dokážeme identifikovat (tvoří pak vektor kovariát  $X$ ) a platí tedy  $P(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(1), \mathbf{Y}(0)) = P(\mathbf{W} \mid \mathbf{X})$ , pak kauzální efekt můžeme odhadovat „téměř jako“ v randomizovaných experimentech. Za těchto okolností lze k odhadování efektu ošetření použít i regresi, regresi s využitím vah propenzitních skóru nebo párování (matching).

V literatuře lze pro předpoklad nezavádějícího mechanismu přiřazování najít i další termíny, např. „ignorability assumption“, „conditional independence“ (Lechner, 2001), „selection on observables“ (Heckman, Ichimura a Todd, 1997) či „exogenita“ (Imbens, 2004). Termín „unconfoundedness“ (nezavádějící mechanismus přiřazování) je zaveden v Rubin (1990).

Pokud u mechanismu přiřazování jeho „nezavádějící“ charakter předpokládat nelze, tedy přiřazení ošetření souvisí s potenciálními odezvami a kontrola na zavádějící proměnné není k dispozici, pak lze odhady efektů získat jen obtížně a jen ve specifických situacích. Při platnosti dalších omezujících podmínek je možno použít některou z následujících metod: metodu „instrumentálních proměnných“ (Instrumental Variables), „regresní diskontinuitu“ (Regression Discontinuity) či metodu „rozdílů v rozdílech“ (Difference in differences). Další metody jsou uvedeny v přehledu Imbens a Wooldridge (2009) a popsány např. v Manski (2007).

## 2.4 Parametry v rámci RCM

V mnoha aplikacích je zcela nerealistické předpokládat homogenní efekt ošetření na celou populaci. Navíc ani nemusí být z pohledu aplikace zajímavé zkoumat efekt ošetření na celou populaci. Například je jistě z praktického pohledu užitečnější ptát se, jak by byly ziskové ty podniky, které dostaly dotaci, kdyby dotaci nedostaly, než otázka, jak by byly ziskové všechny podniky, kdyby dotaci dostaly a kdyby všechny dotaci nedostaly. Výhodou RCM je, že na rozdíl od tradičních přístupů k odhadům efektů, které předpokládají homogenní efekt ošetření na celou populaci (a tedy odhadují jediný parametr), umožňuje heterogenitu efektu ošetření v odhadech zohlednit.

I když je kauzální efekt ošetření na odezvu konkrétního  $i$ -tého subjektu nepozorovatelný, lze odhadovat alespoň různě definované průměrné efekty, které mohou být z pohledu zkoumání vlivu ošetření na odezvu zajímavé. Nejčastěji se odhadují průměrné efekty buď v rámci celé populace ( $\tau_{ATE}$ , zmíněn v podkapitole 2.2), nebo pro různě omezené subpopulace (např.  $\tau_{ATT}$ ,  $\tau_{ATC}$ ).

V této kapitole budou definovány pouze teoretické průměrné efekty a jejich výběrové analogie, v dalších kapitolách bude uvedeno, jak je prakticky odhadnout.  $N$  pozorovaných subjektů reprezentuje náhodný výběr z odpovídající populace a níže uvedené střední hodnoty odpovídají efektům v této populaci.

- Střední hodnota efektu ošetření v rámci celé populace (Population Average Treatment Effect) je definována:

$$\tau_{ATE} = E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

a v rámci náhodného výběru ji odhaduje nepozorovatelné

$$\frac{1}{N} \sum_{i=1}^N [Y_i(1) - Y_i(0)].$$

- Střední hodnota efektu ošetření v rámci subpopulace subjektů s aktivním ošetřením (Population Average Treatment Effect on the Treated) je definována:

$$\begin{aligned} \tau_{ATT} &= E[Y(1) - Y(0) \mid W = 1] = \\ &= E[Y(1) \mid W = 1] - E[Y(0) \mid W = 1] \end{aligned}$$

a v rámci náhodného výběru ji odhaduje nepozorovatelné

$$\frac{1}{N_1} \sum_{i=1}^N W_i [Y_i(1) - Y_i(0)].$$

- Střední hodnota efektu ošetření v rámci subpopulace subjektů s kontrolním ošetřením (Population Average Treatment Effect on the Contoled) je definována:

$$\begin{aligned} \tau_{ATC} &= E[Y(1) - Y(0) \mid W = 0] = \\ &= E[Y(1) \mid W = 0] - E[Y(0) \mid W = 0] \end{aligned}$$

a v rámci náhodného výběru ji odhaduje nepozorovatelné

$$\frac{1}{N_2} \sum_{i=1}^N (1 - W_i) [Y_i(1) - Y_i(0)].$$

Je zřejmé, že pokud spolu odezva  $Y$  a mechanismus přiřazování ošetření  $W$  souvisejí, pak například  $\tau_{ATE} \neq \tau_{ATT}$ .

Někdy mohou být (s ohledem na charakter aplikační úlohy) zajímavé i odhady efektů na jemněji rozložených subpopulacích podle různých úrovní (vybraných) kovariát  $X$ .

Tedy například  $\tau_{CATE} = E[Y(1) - Y(0) \mid X]$  je střední hodnota efektu ošetření v rámci subpopulace těch subjektů, jež mají konkrétní charakteristiky specifikované hodnotami kovariát (kovariáty)  $X$ ; analogicky pak  $\tau_{CATT}$ ,  $\tau_{CATC}$  pro subpopulace s kontrolním či aktivním ošetřením (a požadovanou úrovní kovariát  $X$ ).

V kapitole 4 je pak uveden i lokální efekt ošetření  $\tau_{LATE}$ . Kromě průměrných efektů lze v kontextu kauzální analýzy definovat také kvantilové a distribuční efekty.

## 2.5 Kauzalita před Rubinem

Modelování kauzality na základě konceptu Rubinových potenciálních odezev je moderní přístup, který do různých aplikací pronikl poměrně nedávno. Potřeba odhadovat kauzální efekty je ovšem dávná a před zavedením RCM se vztahy mezi proměnnými modelovaly pomocí strukturálních rovnicových modelů (Structural Equation Models, SEM). Ty jsou v aplikacích v oblasti managementu i marketingu poměrně známé a používané. Příklad použití SEM přístupu k modelování vlivu spokojenosti zákazníka na výkonnost podniku je v [Suchánek a Králová \(2016a\)](#) nebo [Suchánek a Králová \(2016b\)](#).

Bohužel SEM přístup má svá mnohá omezení, kvůli kterým je v jistém smyslu považován za již překonaný. Jednak má silné předpoklady pro rozložení a funkcionální formu vztahů, jednak neumí odlišit kauzální vztah od pouhé asociace. (Celý SEM model je založen výhradně na korelační matici.)

Zajímavý pokus o znovuoživení SEM přístupu v rámci kauzální inference je prezentován v [Pearl \(2000\)](#).





## Kapitola 3

# Metody odhadů v observačních studiích s regulárním mechanismem přiřazování

### 3.1 Základní předpoklady

V observačních studiích, kde mechanismus přiřazování není pod kontrolou výzkumníka, je potřeba věnovat velkou pozornost pečlivému a poctivému hledání všech kovariát (všech možných zavádějících proměnných) tak, aby mechanismus přiřazování ošetření byl jejich funkcí. Tyto kovariáty musejí být nejen známé, ale i pozorovatelné. Pomocí nich je pak možné odstranit vychýlení odhadu efektů ošetření způsobené *samovýběrem* (selfselection) – skutečností, že subjekty si „samy přiřazují“ aktivní/kontrolní ošetření závisle s potenciálními odezvami.

**Pokud lze (díky důkladné znalosti problematiky) obhájit, že mechanismus přiřazování je funkcí známých a pozorovatelných kovariát, pak lze předpoklad nezavádějícího mechanismu přiřazování považovat za splněný.** Potom je **podmíněně** na hodnotách kovariát  $X$  rozložení **potenciálních** odezev  $(Y(1), Y(0))$  pro každou úroveň ošetření  $W$  stejné.

(Tedy mezi subjekty se stejnými hodnotami kovariát je ošetření přiřazováno „jakoby“ náhodně a při modelování kauzální inference lze mechanismus přiřazování „ignorovat“.<sup>10</sup>) Alternativně k 2.3.1 lze tedy předpoklad nezavádějícího mechanismu přiřazování formulovat:

$$[Y(1), Y(0)] \perp\!\!\!\perp W \mid X \quad (3.1.1)$$

Bohužel tento předpoklad nelze testovat, lze jej pouze obhájit argumenty,

---

<sup>10</sup> Proto se jako synonymum pro „unconfoundedness“ (nezavádějící mechanismus přiřazování) používá i termín „ignorability“.

jež plynou ze znalosti problematiky, v níž chceme kauzální inferenci uplatnit. Na druhou stranu pokud ze znalosti problematiky plyne, že mimo pozorovatelné zavádějící proměnné existují i další nedostupné proměnné, které mají vliv jak na potenciální odezvy, tak na mechanismus přiřazování, pak podmínka nezavádějícího mechanismu přiřazování splněna není a metody z této kapitoly nelze použít. V některých situacích pak lze užít alespoň některé z metod v kapitole 4.

I když je předpoklad nezavádějícího mechanismu přiřazování splněn a při odhadování efektů ošetření kontrolujeme „správné“ kovariáty, stále může být obtížné odhadnout efekty ošetření bez vychýlení. Čím jsou si skupiny subjektů kategorizovaných podle jednotlivých úrovní ošetření v kovariátách podobnější, tím snadnější je vyhnout se při odhadu efektů vychýlení. Problém v tomto smyslu představují *nevyváženost* (imbalance) a *nedostatečný překryv*<sup>11</sup> (lack of overlap) jednotlivých kovariát.

- Nevyváženost znamená, že rozložení kovariát se napříč úrovněmi ošetření liší. Kdyby se vůbec nelišilo, situace by byla analogická randomizovanému experimentu. Čím více se naopak rozložení liší, tím více je nutné spoléhat se na správnou specifikaci modelu. (Ačkoliv se pojem nevyváženosti vztahuje k rozložení, prakticky se většinou pro každou úroveň ošetření uvádějí pouze průměry jednotlivých kovariát.)
- Překryv pro uvažovanou kovariátu reprezentuje tu oblast hodnot uvažované kovariáty, pro kterou lze pozorovat subjekty jak s aktivním, tak kontrolním ošetřením. (Analogicky pro  $k$  kovariát bude překryv reprezentovat  $k$ -rozměrný prostor.) Nedostatečný překryv znamená, že pro některé hodnoty uvažované kovariáty (resp.  $k$ -rozměrného vektoru kovariát) nemáme pozorování pro každou úroveň ošetření, a tak efekt ošetření nelze odhadnout s pomocí pozorovaných dat, jež jsou k dispozici (dá se případně odhadnout pouze na základě extrapolace modelu, pokud je dobře specifikován). Proto je pro odhad efektu ošetření **na základě dat** žádoucí, aby překryv platil pro všechny

<sup>11</sup> Požadavek překryvu je zajištěn splněním předpokladu probabilistického mechanismu přiřazování z podkapitoly 2.3.

hodnoty kovariát<sup>12</sup>  $x \in \mathcal{X}$ , kde symbol  $\mathcal{X}$  označuje prostor všech hodnot kovariát, které lze v kontextu studie uvažovat. Předpoklad překryvu lze formálně zapsat následně:

$$0 < P(W_i = 1 \mid X_i = x) < 1 \quad \text{pro libovolné } x \in \mathcal{X} \quad (3.1.2)$$

Předpoklad překryvu (overlap) 3.1.2 lze v literatuře najít i jako „positivity assumption“ (Petersen, Porter, Gruber, Wang a Laan, 2010) nebo „common support“ (Lechner, 2001).

V následující podkapitole budou uvedeny a zdůvodněny předpoklady, za kterých lze i při nerandomizovaném přiřazování ošetření v rámci observačních studií identifikovat na základě pozorovatelných dat efekt ošetření.

## 3.2 Předpoklady identifikace

Identifikovat efekt ošetření na odezvu pouze na základě observačních dat lze tehdy, když jsou splněny předpoklady **1**) nezavádějícího mechanismu přiřazování a **2**) překryvu pro všechna  $x \in \mathcal{X}$ . Současné platnosti těchto dvou předpokladů se někdy říká *silná ignorabilita* (strong ignorability<sup>13</sup>) (Rosenbaum a Rubin, 1983b). Za silné ignorability lze efekt ošetření  $\tau_{ATE}$  odhadnout pouze na základě pozorovaných dat jako  $E(\tau(X))$ , kde:

$$\tau(x) = E[Y(1) - Y(0) \mid X = x] \quad (3.2.1)$$

$$= E[Y(1) \mid X = x] - E[Y(0) \mid X = x] \quad (3.2.2)$$

$$= E[Y(1) \mid W = 1, X = x] - E[Y(0) \mid W = 0, X = x] \quad (3.2.3)$$

$$= E[Y \mid W = 1, X = x] - E[Y \mid W = 0, X = x]. \quad (3.2.4)$$

$\tau(x)$  je očekávaný rozdíl potenciálních odezev při aktivním a kontrolním ošetření podmíněně na hodnotách kovariát  $x$ . Rovnost vztahů 3.2.2 a 3.2.3 plyne z nezavádějícího mechanismu přiřazování v 3.1.1; rovnost vztahů 3.2.3

<sup>12</sup> V souladu s označením na začátku kapitoly 2 symbol  $X$  stále reprezentuje vektor kovariát  $X_1, \dots, X_k$ .

<sup>13</sup> Jak je uvedeno v 1.1, „unconfoundedness“ a „ignorability“ jsou synonyma. Tato publikace upřednostňuje první z nich a konzistentně používá termín „nezavádějící mechanismus přiřazování“. Ovšem termín „strong unconfoundedness“ se v anglické literatuře nepoužívá, proto je zde použit termín „silná ignorabilita“.

a 3.2.4 plyne z definice **pozorovatelné** odezvy v 2.2.1. Ve vztahu 3.2.4 již vystupují pozorovatelné  $Y$ .

Pokud platí předpoklad překryvu, pak  $\tau(x)$  lze určit pro libovolné hodnoty kovariát  $x \in \mathcal{X}$  s hustotou  $f_x$ , a efekt ošetření v celé populaci je tak

$$\tau_{ATE} = E(\tau(x)) = \int \tau(x) f_x(x) dx.$$

Analogicky lze odvodit pro další parametry, např. pro  $\tau_{ATT}$ :

$$\tau_1(x) = E[Y(1) - Y(0) \mid W = 1, X = x] \quad (3.2.5)$$

$$= E[Y(1) \mid W = 1, X = x] - E[Y(0) \mid W = 1, X = x] \quad (3.2.6)$$

$$= E[Y(1) \mid X = x] - E[Y(0) \mid X = x] \quad (3.2.7)$$

$$= E[Y(1) \mid W = 1, X = x] - E[Y(0) \mid W = 0, X = x] \quad (3.2.8)$$

$$= E[Y \mid W = 1, X = x] - E[Y \mid W = 0, X = x] \quad (3.2.9)$$

Rovnost vztahů 3.2.6 až 3.2.8 plyne opět z nezavádějícího mechanismu přiřazování, a v 3.2.9 je již vše pozorovatelné. Je dále vidět, že  $\tau_1(x) = \tau(x)$ , a tedy

$$\tau_{ATT} = E(\tau_1(x) \mid W = 1) = E(\tau(x) \mid W = 1) = \int \tau(x) f_{x|W=1}(x) dx.$$

Platí-li tedy předpoklady nezavádějícího mechanismu přiřazování a překryvu, pak je **nepozorovatelná odezva  $Y(0)$  u subjektu s aktivním ošetřením  $W = 1$  „v průměru“ stejná jako pozorovatelná odezva subjektu s kontrolním ošetřením** (a naopak)! Průměrný kauzální efekt ošetření tedy lze na základě dat identifikovat.

### 3.3 Odhady efektů pomocí klasického lineárního regresního modelu

Lineární regresní model se ke kauzální analýze historicky používal (předpoklad nezavádějícího mechanismu přiřazování je splněn, pokud je splněn předpoklad, že  $W_i$  jako prediktor je nekorelován s náhodnou chybou  $\varepsilon_i$ ), ale právě zmíněné problémy s častou nevyvážeností a nesplněným překryvem pro všechna  $x \in \mathcal{X}$  vedly v posledních desetiletích k odklonu od regresních modelů k metodám uvedeným níže.

Lineární regresní model lze tedy ke kauzální analýze použít tehdy, když je buď skutečně dobře specifikován,<sup>14</sup> nebo když jsou skupiny pro různé úrovně ošetření vyvážené vzhledem ke všem kovariátám. Pokud se místo extrapolace na základě specifikovaného modelu chceme v odhadech spoléhat na podporu dat, pak musí platit i překryv pro všechna  $x \in \mathcal{X}$ .

Nedostatečný překryv představuje pro získání nevychýleného odhadu efektu ošetření větší problém než nevyváženost. V podkapitolách níže uvedeme metody, které se vypořádávají s oběma problémy zároveň.

### 3.4 Subklasifikace

Jednoduchou metodou vhodnou pro situaci s jednou kvalitativní zavádějící proměnnou  $X$  (nebo „malým počtem“ kvalitativních zavádějících proměnných) je subklasifikace. Zavádějící vliv této kovariáty  $X$  se odstraní tak, že se populace rozdělí do strat (subpopulací) podle úrovní zavádějící proměnné  $X$  a následně se v rámci těchto strat získají střední hodnoty odezvy zvláště v subpopulaci s aktivním ošetřením a zvláště v subpopulaci s kontrolním ošetřením. Tyto střední hodnoty se zprůměrují napříč straty (s vahami odpovídajícími velikosti strat) a porovnají. Tedy formálně:

$$E(Y(w)) = \sum_x E(Y | W = w, X = x)P(X = x) \quad \text{pro } w = 0, 1 \quad (3.4.1)$$

$$\tau_{ATE} = E(Y(1)) - E(Y(0)) \quad (3.4.2)$$

Na výběrových datech se ve vzorci 3.4.1 nahradí střední hodnoty průměry a pravděpodobnosti relativními četnostmi. Odhady  $\tau_{ATT}$ , či  $\tau_{ATC}$  lze získat modifikací četnostních vah. Např. pro odhad  $\tau_{ATT}$  budou četnostní váhy odpovídat podílu subjektů s odpovídající úrovní kovariáty  $X = x$  mezi všemi subjekty s aktivním ošetřením.

Subklasifikace je nevhodná, pokud se vyskytne kvalitativních kovariát více. Nejen že je vytváření strat pro každou možnou kombinaci úrovní více kvalitativních kovariát nepřehledné a pracné. Může se i stát, že pro určitou

<sup>14</sup> Na základě empirického pravidla (Imbens a Rubin, 2009) jsou pro normalizované rozdíly  $\frac{\bar{X}_1 - \bar{X}_0}{\sqrt{S_0^2 + S_1^2}}$  s hodnotou nad 0,25 regresní modely na nesprávnou specifikaci velmi citlivé. Symboly  $\bar{X}_1$  a  $\bar{X}_0$  zde reprezentují výběrové průměry pro skupinu s aktivním a kontrolním ošetřením u jedné kovariáty,  $S_0^2$  a  $S_1^2$  jsou výběrové rozptyly pro obě skupiny a uvažovanou kovariátu.

kombinaci nejsou v souboru vůbec žádná pozorování, a tedy není splněn předpoklad překryvu.

### 3.5 Párování

Subklasifikace se k odhadnutí efektu ošetření také vůbec nehodí tehdy, pokud jsou jedna nebo více kovariát spojité. V tomto případě lze za předpokladu nezavádějícího mechanismu přiřazování použít některou z metod *párování* (matching), která mimo jiné zajistí i překryv, pokud tento předpoklad datový soubor nespĺňuje.

V metodách párování jde o to zredukovat výběrový soubor tak, aby si skupina subjektů s aktivním ošetřením a skupina subjektů s kontrolním ošetřením byly co nejpodobnější s ohledem na kovariáty  $X$ . V rámci široké škály metod pro párování je tím nejjednodušším přístupem některá z metod *párování* „jeden na jednoho“ (one-to-one matching); populární je např. *metoda nejbližšího souseda*. Při párovacích procedurách je chybějící nepozorovaná potenciální odezva  $Y_i(1 - W_i)$  k  $i$ -tému subjektu s pozorovanou odezvou  $Y_i$  při ošetření  $W_i$  nahrazena pozorovanou odezvou jiného subjektu s ošetřením  $1 - W_i$ , který se  $i$ -tému subjektu v hodnotách kovariát podobá co nejvíce (ideálně je stejný). (Hledáme tedy "dvojčata" pokud jde o hodnoty kovariát, která se ale mezi sebou liší úrovní ošetření.)

Je-li cílem výzkumníka odhadnout efekt ošetření  $\tau_{ATT}$ , pak je ke každému subjektu ze skupiny s aktivním ošetřením ve skupině s kontrolním ošetřením přiřazen subjekt s velmi podobnými hodnotami kovariát. Nejde zde primárně o to, aby takto vzniklý  $i$ -tý pár byl co nejpodobnější s ohledem na  $X$ , ale aby si v distribuci byly obě nově vzniklé skupiny (skupina subjektů s aktivním ošetřením a k ní napárovaná skupina vybraných vhodných subjektů s kontrolním ošetřením) co nejpodobnější s ohledem na  $X$ . Zjistit průměrný efekt ošetření pak lze srovnáním průměrů odezvy mezi nově vzniklými skupinami. Tedy například odhad očekávaného efektu ošetření na subpopulaci s aktivním ošetřením  $\tau_{ATT}$  je:

$$\hat{\tau}_{ATT} = \frac{1}{N_1} \sum_{i: W_i=1} (Y_i - \tilde{Y}_i) \simeq \frac{1}{N_1} \sum_{i: W_i=1} (Y_i(1) - Y_i(0)) \quad (3.5.1)$$

kde  $\tilde{Y}_i$  je pozorovaná odezva mezi subjekty s kontrolním ošetřením, která zastupuje nepozorovatelný kontrafaktuál  $Y_i(0)$  pozorovatelnému  $Y_i(1)$ .

Další možností, jak odhadnout efekt ošetření, je použít regresi, ovšem již ne na původních datech, ale na souboru dat redukováném párováním. (Tento soubor napárovaných dat je nyní již vyvážený a s dostatečným překryvem, tudíž senzitivita na špatnou specifikaci modelu či extrapolace na napárovaném souboru již nepředstavují problém.)

Přístup párování jeden na jednoho při odhadování  $\tau_{ATT}$  předpokládá, že ve skupině s kontrolním ošetřením bude dostatek subjektů podobných subjektům s aktivním ošetřením (s ohledem na kovariáty  $X$ ) a také že některé subjekty ze skupiny s kontrolním ošetřením zůstanou zřejmě nevyužité (obvykle v observačních studiích převažují subjekty s kontrolním ošetřením). Je také ale možné, že k některým subjektům ze skupiny s aktivním ošetřením nelze najít odpovídající subjekty ze skupiny s kontrolním ošetřením. Pak se kauzální inference bude vztahovat pouze k takovému podprostoru  $\mathcal{X}^* \subset \mathcal{X}$ , kde je zajištěn překryv.

Pozorované odezvě  $Y_i$   $i$ -tého subjektu ze skupiny s aktivním ošetřením lze jako náhradu jeho kontrafaktuálu přiřadit také průměr z  $M_i$  odezev subjektů ze skupiny s kontrolním ošetřením, jež jsou s ohledem na kovariáty podobné uvažovanému  $i$ -tému subjektu. Pak  $\tau_{ATT}$  bude odhadnuto následně:

$$\hat{\tau}_{ATT} = \frac{1}{N_1} \sum_{i: W_i=1} \left[ Y_i - \left( \frac{1}{M_i} \sum_{m=1}^{M_i} \tilde{Y}_{im} \right) \right] \quad (3.5.2)$$

kde  $\tilde{Y}_{im}$  je  $m$ -tá odezva ze všech  $M_i$  odezev subjektů z kontrolní skupiny, které jsou s ohledem na kovariáty  $X$  podobné subjektu  $i$  ze skupiny s aktivním ošetřením. Teoretické vlastnosti těchto a dalších odhadů efektů jsou popsány např. v [Abadie a Imbens \(2006\)](#).

Přiřazováním pozorovaných „náhrad“ kontrafaktuálů k pozorované odezvě  $Y_i$  je dosaženo jednoho z hlavních cílů párování a tím je korekce nedostatečného překryvu mezi skupinou s aktivním a skupinou s kontrolním ošetřením. Zároveň párování vylepšuje i vyváženost mezi skupinami.

Literatura k různým přístupům k párování a následně vlastnostem odhadů efektů je bohatá, významné práce jsou např. [Rubin \(1979\)](#), [Gu a Rosenbaum \(1993\)](#), [Rosenbaum \(1989\)](#), [Rosenbaum \(1995\)](#), [Rosenbaum \(2002\)](#), [Rubin a Thomas \(2000\)](#) či [Diamond a Sekhon \(2013\)](#).

### 3.5.1 Podobnost při párování

Je-li spojitá kovariáta pouze jedna, pak je párování analogické stratifikaci a „podobnost“ dvou subjektů s ohledem na tuto kovariátu lze definovat jednoduše jako vzdálenost těchto dvou subjektů v měřítku uvažované kovariáty. Ideální „nulovou“ vzdálenost (tedy dokonalou shodu obou subjektů v měřítku uvažované kovariáty) ovšem prakticky nelze u všech dvojic subjektů očekávat.

Pokud je kovariát více, tedy  $k \geq 2$ , pak vzdálenost  $d_{ij}$  dvou bodů  $x_i$  a  $x_j$  v  $k$  rozměrném prostoru lze definovat různými způsoby.

- Nejpopulárnější je *Mahalanobisova*<sup>15</sup> vzdálenost, která vzdálenost dvou bodů definuje:

$$d_{ij}^M = \sqrt{(x_i - x_j)' \Sigma^{-1} (x_i - x_j)} \quad (3.5.3)$$

kde  $\Sigma^{-1}$  je kovarianční matice vektoru  $X$ .

- Mezi další často užívané metriky patří metrika *genetického párování*, která vzdálenost dvou bodů definuje:

$$d_{ij}^{gen} = \sqrt{(x_i - x_j)' (\Sigma^{-1/2})' W \Sigma^{-1/2} (x_i - x_j)}$$

kde  $W$  je váhová matice získaná optimalizačním algoritmem popsaným např. v [Diamond a Sekhon \(2013\)](#).

- Další metrikou, která má v párování speciální postavení, je *propenzitní skór*, jemuž je věnována další podkapitola [3.5.2](#).

### Měřítko pro okolí (Caliper)

Někdy může dojít k situaci, že v datovém souboru k nějakému subjektu s aktivním ošetřením neexistuje do páru dostatečně podobný subjekt s kontrolním ošetřením. Ovšem některé párovací algoritmy vybírají do páru nejlepší (t.j. nejpodobnější, nejbližší) možný subjekt z těch, které jsou v párování k dispozici. Může se tak stát, že tím nejlepším subjektem

<sup>15</sup> Speciálním případem Mahalanobisovy vzdálenosti je Euklidovská vzdálenost, pokud je matice  $\Sigma$  jednotková (kovariáty jsou nekorelované a normované).



do páru je i subjekt dost nepodobný, takový, který již nelze považovat za dobré „dvojče“ a bylo by užitečné takovéto párování „zakázat“.

Pro definování skutečnosti, co je „dobře“ napárovaná dvojice, se používá *měřítka pro okolí* (caliper). Párování proběhne pouze u takových dvojic (subjektu s aktivním ošetřením se subjektem s kontrolním ošetřením), jejichž vzdálenost (ve zvolené metrice) je menší než nastavená hodnota měřítka pro okolí. Pokud se k subjektu s aktivním ošetřením nenajde s ohledem na nastavené měřítko okolí dostatečně podobné „dvojče“, pak je potřeba takovýto subjekt z datového souboru vyloučit.

Při platnosti předpokladu překryvu by měl ke každému subjektu s aktivním ošetřením (podmíněně na hodnotách kovariát  $X$ ) existovat subjekt s kontrolním ošetřením. Pokud uvnitř měřítka pro okolí nelze najít vhodný párovatelný subjekt z druhé skupiny, pak je to známka porušení předpokladu překryvu. Vyloučení takovýchto „nespárovatelných“ subjektů tedy řeší i problém nedostatečného překryvu. Nevýhodou je ovšem skutečnost, že je tímto krokem změněna populace pro inferenci.

### 3.5.2 Párování na základě propenzitních skóru

Propenzitní skór  $e(x)$  definovali [Rosenbaum a Rubin \(1983b\)](#) jako pravděpodobnost, že subjektu bude přiřazeno aktivní ošetření podmíněně na hodnotách všech jeho kovariát, tedy

$$e(x) = P(W = 1 \mid X = x). \quad (3.5.4)$$

Dále dokázali, že za předpokladu silné ignorability jsou ošetření  $W$  a vektor kovariát  $X$ , podmíněně na propenzitním skóru, vzájemně nezávislé.

Formálně

jestliže platí:

1.  $[Y(1), Y(0)] \perp\!\!\!\perp W \mid X$   
(přiřazování je nezavádějící)
2.  $0 < P(W = 1 \mid X = x) < 1$  pro lib.  $x \in \mathcal{X}$   
(je splněn předpoklad překryvu,)

potom:

$$W \perp\!\!\!\perp X \mid e(x) \quad (3.5.5)$$

Nebo alternativně k 3.5.5:  $P(W = 1 \mid e(X), X) = P(W = 1 \mid e(X))$ .

Jinými slovy u subjektů se stejným propenzitním skórem je rozložení kovariát  $X$  stejné. Pomocí propenzitních skóru tedy můžeme získat obě skupiny (s aktivním a s kontrolním ošetřením) vyvážené a říkáme, že propenzitní skór má *vyvažovací schopnost* (balancing property). Praktický význam je ten, že za účelem párování není již potřeba kontrolovat celý vektor kovariát (tedy všech  $k$  zavádějících proměnných), nýbrž pouze propenzitní skór (skalár).

Při randomizovaném experimentu je propenzitní skór známý (je definován designem experimentu), bohužel v observačních studiích známý není a je potřeba jej odhadnout. Nejčastěji se odhaduje pomocí logistické regrese, kde jsou vysvětlujícími proměnnými jednotlivé kovariáty vektoru  $X$  a modeluje se ošetření  $W$ . (Pokud má  $W$  více než dvě úrovně, pak lze použít např. multinomickou regresi. V případě „near-optimal generalised full matching“ jsou vyvinuty i další univerzálně použitelné algoritmy popsané v [Sävje, Higgins a Sekhon \(2017\)](#).)

Protože na základě propenzitních skóru bude provedeno párování a následně odhad efektů ošetření, je důležité tyto skóry správně odhadnout (t.j. najít správnou specifikaci logistické či jiné funkce). Kritériem pro správnou specifikaci logistické (či jiné použité) funkce je **co nejlepší vyváženost skupin** s aktivním a kontrolním ošetřením. Teoreticky je vyváženost (při správných propenzitních skórech) garantována, jelikož propenzitní skóry mají vyvažovací schopnost. Prakticky se nejdříve zkusí „nějaká rozumná“ logistická funkce, provede se párování a pak se porovná rozložení kovariát ve skupinách vytvořených párováním. Pokud nelze skupiny považovat za uspokojivě vyvážené, přistoupí se k úpravě specifikace logistického modelu tak, aby se vyváženost skupin zlepšila. Před samotným párováním je ještě užitečné graficky posoudit překryv pomocí histogramů propenzitních skóru skupiny s kontrolním a skupiny s aktivním ošetřením. Tvary histogramů nemusejí být stejné či podobné, stačí, když jsou definovány přes stejné intervaly propenzitních skóru, a je tedy splněn předpoklad překryvu.

Tvar použité logistické funkce je interpretačně nezajímavý, jediným kritériem, které je zajímavé s ohledem na tvar logistické funkce, je konečná vyváženost skupin. Vyváženost posuzujeme například na základě popisných statistik jednotlivých kovariát, tedy typicky průměrů kovariát v každé

skupině, případně rozptylů, kvantilů apod. Dále lze testem porovnávat distribuční funkce mezi skupinou s aktivním a kontrolním ošetřením pro každou kovariátu zvlášť (například pomocí *Kolmogorov-Smirnovova* testu), ovšem při interpretaci  $p$ -hodnot je potřeba brát v úvahu i konečný počet subjektů ve výběrovém souboru zredukovaném po párování.

Proces vylepšování odhadů propenzitních skóre probíhá zcela bez ohledu na hodnoty odezev  $Y$ ! Máme-li již konečné odhady propenzitních skóre, použijeme je pak jako metriku pro párování analogicky např. k Mahalanobiově metrice. Praktickou výhodou propenzitních skóre je, že nyní párování neprobíhá na základě  $k$ -rozměrného vektoru kovariát  $X$ , nýbrž na základě skalárního propenzitního skóre  $e(X)$ . Dále již lze odhadnout efekt ošetření na základě průměrů odezev či regrese na vyvážených skupinách.

Párování odstraňuje vychýlení odhadu efektu ošetření způsobené kovariátami  $X$  (zavádějícími proměnnými) a efekt ošetření lze pomocí párování odhadnout jednoduše standardními statistickými postupy. Zvolený algoritmus párování ale zároveň může odhad rozptylu odhadu efektu komplikovat. V případě one-to-one párování „jeden na jednoho“ je odhad rozptylu odhadu efektu standardní. Jestliže se ale musejí některé náhrady kontrafaktuálů použít opakovaně (with replacement) – zejména při malém rozsahu skupiny s kontrolním ošetřením nebo při malém překryvu obou skupin – pak je nutné rozptyl odhadu efektu odhadovat pokročilejšími statistickými metodami, nejčastěji se používají sandwichové odhady.

## 3.6 Další využití propenzitních skóre

Propenzitní skóre lze pro odhad efektu ošetření užít i jinak než jako metriku pro párování. Další možnosti jejich užití jsou uvedeny níže:

### 3.6.1 Stratifikace pomocí propenzitních skóre

Tato metoda navržená v [Rosenbaum a Rubin \(1983b\)](#) rozdělí datový soubor do strata na základě hodnot propenzitního skóre. Subjekty v rámci jednoho strata (s velmi podobnými hodnotami propenzitního skóre) jsou považovány za subjekty z „jakoby“ téměř randomizovaného experimentu, kde se úroveň ošetření přiřazuje subjektům náhodně. Potom lze efekt ošetření uvnitř tohoto strata získat porovnáním průměrných odezev obou

skupin a následně se efekty jednotlivých strat zprůměrují pomocí vah odpovídajících počtu subjektů v jednotlivých stratech. Formálně:

Hodnoty propenzitního skóru (jedná se o pravděpodobnosti) rozdělíme nejdříve do  $J$  strat následně:  $0 = c_0 < c_1 < c_2 < \dots < c_J = 1$ .

Dále zavedeme indikátor  $B_{ij}$  skutečnosti, zda se  $i$ -tý subjekt nachází v  $j$ -tém stratu následným způsobem: pro  $i = 1, \dots, N$  a  $j = 1, \dots, J$  je

$$B_{ij} = \begin{cases} 1 & \text{jestliže } c_{j-1} \leq e(X_i) < c_j \\ 0 & \text{jinak.} \end{cases}$$

Střední hodnota efektu ošetření (ATE) v  $j$ -tém stratu je:

$$\tau_j = E[Y_i(1) - Y_i(0) \mid B_{ij} = 1] \quad \text{pro } j = 1, \dots, J$$

a její odhad v  $j$ -tém stratu je:

$$\hat{\tau}_j = \bar{Y}_{j1} - \bar{Y}_{j0},$$

kde  $\bar{Y}_{j1}, \bar{Y}_{j0}$  jsou průměrné odezvy v  $j$ -tém stratu pro obě skupiny:

$$\bar{Y}_{j1} = \frac{1}{N_{j1}} \sum_{i: W_i=1}^N B_{ij} \cdot Y_i; \quad \bar{Y}_{j0} = \frac{1}{N_{j0}} \sum_{i: W_i=0}^N B_{ij} \cdot Y_i$$

přičemž  $N_{j1}$  a  $N_{j0}$  jsou postupně počty objektů s aktivním ošetřením v  $j$ -tém stratu a počty objektů s kontrolním ošetřením v  $j$ -tém stratu.

Při velkém počtu  $J$  a malé délce intervalu  $c_j - c_{j+1}$  lze hodnoty propenzitního skóru subjektů uvnitř  $j$ -tého intervalu považovat za téměř totožné a odhad efektu ošetření v  $j$ -tém intervalu za téměř nevychýlený. Celkový odhad efektu ošetření v celé populaci bude potom vážený průměr efektů získaných ve všech stratech:

$$\hat{\tau} = \sum_{j=1}^J \hat{\tau}_j \cdot \left( \frac{N_{j0} + N_{j1}}{N} \right)$$

### 3.6.2 Vážení propenzitními skóry

Cílem metody známé jako IPTW (Inverse Probability Treatment Weighting) je vytvoření dvou skupin pseudopopulace, kde jedna skupina pseudopopulace reprezentuje původní populaci, kdyby všem subjektům bylo přiřazeno

aktivní ošetření, a druhá skupina pseudopopulace reprezentuje původní populaci, kdyby všem subjektům bylo přiřazeno kontrolní ošetření. Potom lze efekt ošetření  $W$  na odezvu  $Y$  získat jednoduše porovnáním středních hodnot odezvy  $Y$  v obou skupinách pseudopopulace. Tento přístup nahrazuje randomizovaný experiment, ač má k dispozici pouze observační data.

Princip metody spočívá v tom, že jednotlivým pozorováním (v rámci observační studie) je přiřazena váha tak, aby finální skupiny pseudopopulace měly stejné rozložení zavádějících proměnných. Připomeňme, že propenzitní skór  $e(x)$  je pravděpodobnost přiřazení aktivního ošetření podmíněná hodnotami všech zavádějících proměnných (3.5.4). Subjektům, jimž bylo ve skutečnosti přiřazeno aktivní ošetření, se přiřadí váha  $\frac{1}{e(x)}$ , a subjektům, jimž bylo ve skutečnosti přiřazeno kontrolní ošetření, se přiřadí váha  $\frac{1}{1-e(x)}$ . Pro pevně zafixované hodnoty vektoru kovariát je pak počet nasimulovaných subjektů v obou skupinách pseudopopulace stejný a mechanismus přiřazení ošetření již není ovlivněn zavádějícími kovariáty  $X$ . (V pseudopopulaci má každý subjekt stejnou pravděpodobnost dostat aktivní nebo kontrolní ošetření. Simuluje se tedy randomizovaný experiment.) Názorně je idea metody IPTW popsána v komentáři k obrázku 3.1.

Střední hodnoty potenciálních odezev jsou pak

$$E[Y_i(1)] = E\left[\frac{W_i \cdot Y_i}{e(X_i)}\right]; \quad E[Y_i(0)] = E\left[\frac{(1 - W_i) \cdot Y_i}{(1 - e(X_i))}\right]$$

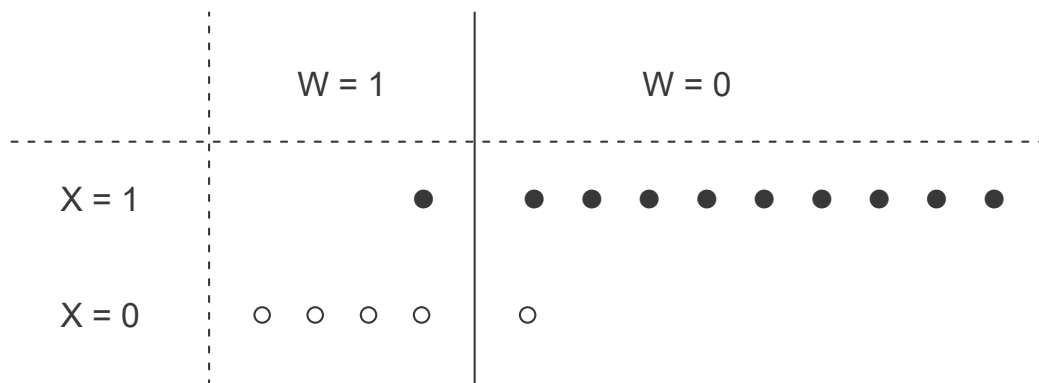
a efekt ošetření  $\tau_{ATE}$  je potom

$$\tau_{ATE} = E\left[\frac{W_i \cdot Y_i}{e(X_i)} - \frac{(1 - W_i) \cdot Y_i}{(1 - e(X_i))}\right]. \quad (3.6.1)$$

Jako odhad efektu ošetření se nabízí

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^N \left[ \frac{W_i \cdot Y_i}{e(X_i)} - \frac{(1 - W_i) \cdot Y_i}{(1 - e(X_i))} \right]. \quad (3.6.2)$$

Odhad  $\hat{\tau}$  3.6.2 ovšem vyžaduje znalost hodnot propenzitních skóre  $e(x)$ , které se ale v praxi většinou pouze odhadují (obvykle pomocí logistické regrese). Odhad efektu ošetření  $\tau_{ATE}$  pomocí vzorce 3.6.3 již využívá



**Obrázek 3.1:**  $X \in \{0, 1\}$  je pro jednoduchost skalární zavádějící proměnná nabývající pouze dvou hodnot a tečky na obrázku reprezentují původní populaci. Aktivní ošetření  $W = 1$  bylo přiřazeno pouze 10 % takových subjektů, pro něž  $X = 1$ . To znamená, že jejich propenzitní skór je  $e(1) = P(W = 1 | X = 1) = 0,1$ . Obdobně aktivní ošetření  $W = 1$  bylo přiřazeno až 80 % takových subjektů, pro něž  $X = 0$ , tedy jejich propenzitní skór je  $e(0) = P(W = 1 | X = 0) = 0,8$ . (Zavádějící vliv proměnné  $X$  je zřejmý.) Při párování „jeden na jednoho“ by se např. k jedinému subjektu – takovému, že  $X = 1$  a  $W = 1$  – napároveň jediný subjekt s kontrolním ošetřením a hodnotou kovariáty  $X = 1$ . Zbylých osm subjektů ze skupiny s kontrolním ošetřením a hodnotou kovariáty  $X = 1$  by se nevyužilo. Nyní ale budou pomocí IPTW vytvořeny dvě skupiny pseudopopulace, jedna s aktivním a druhá s kontrolním ošetřením tak, že budou simulovat randomizovaný experiment a žádný subjekt z původní populace nebude „zahozen“. Subjektům, pro něž  $X = 1$  a dostaly aktivní ošetření, se přiřadí váha  $\frac{1}{e(1)} = \frac{1}{0,1} = 10$ . Subjektům, pro něž  $X = 1$  a dostaly kontrolní ošetření, se přiřadí váha  $\frac{1}{1-e(1)} = \frac{1}{0,9} = \frac{10}{9}$ . Tedy jeden subjekt s aktivním ošetřením dostane váhu 10 a celkově je tedy nasimulováno  $10 \cdot 1 = 10$  „pseudosubjektů“ s aktivním ošetřením pro  $X = 1$ . Podobně devět subjektů s kontrolním ošetřením dostane váhu  $10/9$  a celkově je tedy nasimulováno  $10/9 \cdot 9 = 10$  „pseudosubjektů“ s kontrolním ošetřením pro  $X = 1$ . Analogicky pro  $X = 0$  bude v nasimulované pseudopopulaci  $4 \cdot \frac{1}{0,2} = 5$  subjektů s aktivním ošetřením a  $1 \cdot \frac{1}{0,8} = 5$  subjektů s kontrolním ošetřením. Zavádějící vliv proměnné  $X$  je tak odstraněn.

odhady propenzitních skóru  $\hat{e}(x)$  a je dokonce vydatnější než  $\hat{\tau}$  (Hirano, Imbens a Rider, 2003).

$$\hat{\tau}_{ipw} = \frac{\sum_{i=1}^N \frac{W_i \cdot Y_i}{\hat{e}(X_i)}}{\sum_{i=1}^N \frac{W_i}{\hat{e}(X_i)}} - \frac{\sum_{i=1}^N \frac{(1 - W_i) \cdot Y_i}{1 - \hat{e}(X_i)}}{\sum_{i=1}^N \frac{1 - W_i}{1 - \hat{e}(X_i)}} \quad (3.6.3)$$

Výraz  $\frac{\sum_{i=1}^N \frac{W_i \cdot Y_i}{\hat{e}(X_i)}}{\sum_{i=1}^N \frac{W_i}{\hat{e}(X_i)}}$  můžeme chápat jako průměrnou hodnotu  $Y$  v pseudo-populaci s aktivním ošetřením, jelikož jmenovatel odpovídá počtu prvků v pseudopopulaci s aktivním ošetřením a čitatel odpovídá součtu hodnot  $Y$  pro všechny subjekty v pseudopopulaci s aktivním ošetřením. (A analogicky pro pseudopopulaci s kontrolním ošetřením.)

### 3.6.3 Marginální strukturální modely

Porovnání středních hodnot potenciálních odezev, jak je uvedeno výše, neumožňuje odhadnout efekt ošetření v případě, kdy je tento efekt modifikován dalšími vysvětlujícími proměnnými  $V$ . (Ty nepatří mezi zavádějící proměnné  $X$ !) Za této situace lze použít *Marginální Strukturální modely* (MSM), které připomínají klasickou regresi, ovšem modelují střední hodnotu **potenciálních** odezev, nikoliv **pozorovaných** odezev, a nejsou podmíněné na zavádějících proměnných, jelikož využívají subjekty z výše zavedené pseudopopulace simulované pomocí metody IPTW. S ohledem na charakter odezvy  $Y$  (spojitá nebo binární či četnostní apod.) lze analogicky klasickému regresnímu přístupu použít lineární MSM, logistický MSM či zobecněný MSM.

#### Lineární MSM

Pro lineární MSM (zatím bez další proměnné, jež by mohla efekt odezvy modifikovat) modelujeme střední hodnotu potenciální odezvy  $E(Y(w))$  následovně:

$$E(Y(w)) = \psi_0 + \psi_1 w, \quad w = 0, 1 \quad (3.6.4)$$

Potom

$$E(Y(0)) = \psi_0 \quad \text{a} \quad E(Y(1)) = \psi_0 + \psi_1$$

Tedy  $\psi_1$  je průměrný efekt ošetření;  $\psi_1 = E(Y(1)) - E(Y(0))$ .

Je-li efekt ošetření modifikován další proměnnou  $V$ , pak lze lineární MSM upravit následovně:

$$E(Y(w) | V) = \psi_0 + \psi_1 w + \psi_2 V + \psi_3 wV, \quad w = 0, 1. \quad (3.6.5)$$

Tedy  $E(Y(1) | V) - E(Y(0) | V) = \psi_1 + \psi_3 V$

### Zobecněný MSM

Analogicky zobecněným lineárním modelům (GLzM), kde  $g$  je linkovací funkce, bude v zobecněném MSM modelována střední hodnota potenciální odezvy následovně:

$$E(Y(w)) = g^{-1}(\psi_0 + \psi_1 w), \quad w = 0, 1 \quad (3.6.6)$$

Jak již bylo uvedeno v úvodu této sekce, MSM modelují potenciální odezvy  $Y$  pro pevně nastavenou hodnotu ošetření  $W$  (symbolicky vyjádřeno malým písmenem  $w$ ), zatímco klasické regresní modely modelují odezvu  $Y$  podmíněně na hodnotě ošetření  $W$  ovlivněné možnými zavádějícími proměnnými, tedy

$$E(Y | W) = g^{-1}(\psi_0 + \psi_1 W) \quad (3.6.7)$$

Kdyby datový soubor pocházel z randomizovaného experimentu, pak by i klasický regresní model dobře odhadl efekt ošetření. V případě dat z observačních studií však (kvůli vlivům zavádějících proměnných) klasický regresní model použít nelze.

Uvážíme-li, že skupiny pseudopopulace vytvořené pomocí IPTW již mají stejné rozložení zavádějících proměnných (jak skupina pseudopopulace s aktivním, tak skupina pseudopopulace s kontrolním ošetřením), můžeme parametry MSM získat i na observačních datech. Technicky se odhady parametrů MSM odhadují stejně jako v GLzM, jen je potřeba u každého pozorování  $i$  zohlednit váhy  $\frac{1}{W_i e(X_i) + (1 - W_i)(1 - e(X_i))}$ . Odhad rozptylu odhadů pak lze získat například pomocí bootstrapu nebo robustním sandwichovým odhadem rozptylu.<sup>16</sup> Podrobněji například v [Robins, Hernán a Brumback \(2000\)](#).

Problém mohou představovat pozorování, která mají extrémně velkou váhu, jelikož silně přispívají k variabilitě odhadu kauzálního efektu (a naznačují problém s předpokladem překryvu). Proto je vhodná inspekce rozložení propenzitních skóre a v případě extrémně velkých či extrémně malých hodnot propenzitních skóre je možné zvážit použití metody *ořezávání* (trimming) nebo *osekávání* (truncation).

<sup>16</sup> Tato řešení zohledňují váhy v pseudopopulacích, nezohledňují ovšem variabilitu v odhadu propenzitních skóre.



### **Ořezávání (trimming)**

Ořezávání spočívá v odstranění těch subjektů z analýzy, jejichž propenzitní skóre jsou extrémní – jsou na chvostech rozložení propenzitních skóre. (Tím je ovšem populace pro inferenci změněna!) Obvykle se postupuje tak, že se odstraní ty subjekty s aktivním ošetřením, jejichž skóre jsou větší než 98. percentil z rozložení propenzitních skóre ve skupině s kontrolním ošetřením. Podobně se odstraní ty subjekty s kontrolním ošetřením, jejichž skóre jsou menší než druhý percentil z rozložení propenzitních skóre ve skupině s aktivním ošetřením. (Postup odpovídá situaci, kdy je rozložení propenzitních skóre ve skupině s aktivním ošetřením posunutá „doprava“ od rozložení propenzitních skóre s kontrolním ošetřením.)

### **Osekávání (truncation)**

Osekávání nemění původní populaci, všechny subjekty zůstanou v souboru pro následnou analýzu zachovány. Ovšem extrémní hodnoty vah spočtených na základě propenzitních skóre se nahrazují nejvyšší (resp. nejnižší) výzkumníkem nastavenou akceptovatelnou hodnotou. Tato hodnota může být stanovena v absolutních číslech nebo na základě percentilu. Je-li tedy nejvyšší akceptovatelná hodnota pro váhu nastavena na hodnotu 80, pak se všem subjektům, jimž je pomocí IPTW přiřazena váha větší než 80, váha sníží na nastavenou hodnotu 80. Tento postup vede k vychýlení odhadu efektu, ale snižuje rozptyl odhadu efektu a vede k menší střední kvadratické chybě.



## Kapitola 4

# Metody odhadů v observačních studiích s neregulárním mechanismem přiřazování

Metody kauzální analýzy popsané v kapitole 3 předpokládaly mimo jiné nezavádějící mechanismus přiřazování, což prakticky znamená, že hodnoty všech zavádějících proměnných jsou pozorovatelné a dostupné. Pro takové observační studie, kde nejsou hodnoty všech zavádějících proměnných dostupné, již obecný přístup k zjištění nevychýleného odhadu efektu ošetření k dispozici není. Ovšem pro některé speciální situace řešení existuje.

Mezi známé metody, jež nevyžadují dostupnost všech zavádějících proměnných (ale mají jiné speciální předpoklady), patří zejména metoda instrumentálních proměnných upravená pro kauzální analýzu, metoda regresní diskontinuity a metoda rozdílu v rozdílech. Dále je v posledních letech rozvíjena i mediační analýza s ohledem na kauzální interpretaci (Imai, Keele a Yamamoto, 2010) či (Huber, Lechner a Mellace, 2016) a strojové učení opět upravené pro kauzální interpretaci, vhodné zejména pro vysokodimenzionální úlohy (Belloni, Chernozhukov a Hansen, 2014).

V této kapitole bude uvedena metoda instrumentálních proměnných (bude použita v kapitole 6) a její speciální případ regresní diskontinuity.

### 4.1 Instrumentální proměnné

Problém vychýlení odhadu efektu z důvodu samovýběru (selfselection) v situacích, kdy nelze vychýlení korigovat pomocí zavádějících proměnných, se tradičně ve statistice i ekonometrii řeší prostřednictvím instrumentálních proměnných. Tradičně užívané lineární regresní modely s instrumentálními

proměnnými ovšem předpokládají homogenní efekt ošetření (t.j. ošetření působí na každý subjekt populace stejně) a s tímto předpokladem pak odhadují průměrný efekt ošetření  $ATE$  pro celou populaci. Tento předpoklad je ovšem ve většině aplikací nereálný, a tradičním přístupem interpretované odhady získané metodou instrumentálních proměnných jsou tedy zavádějící.

Současná literatura k instrumentálním proměnným nabízí **novou kauzální interpretaci** těchto tradičních odhadů pro situace, kdy jsou kauzální efekty heterogenní: základní model pro diskrétní instrument (bude uveden níže) popsali [Imbens a Angrist \(1994\)](#) a [Angrist, Imbens a Rubin \(1996\)](#); modely s rozšířením na další kovariáty pak např. [Frölich \(2007\)](#) a modely se spjitými instrumenty jsou popsány např. ve [Vytlačil \(2002\)](#).

#### 4.1.1 Model s binárním instrumentem a binárním ošetřením

Při zkoumání vlivu ošetření  $W$  na odezvu  $Y$  bez možnosti kontrolovat vliv zavádějících proměnných  $X$  je v rámci metod instrumentálních proměnných potřebná nová proměnná, kterou nazýváme *instrument* (či *instrumentální proměnná*) a bude značena symbolem  $Z$ . Jedná se o takovou proměnnou, pomocí které lze predikovat hodnoty odezvy  $Y$ , přičemž na tuto odezvu působí  $Z$  **výlučně** skrze ošetření  $W$ . A dále, instrumentální proměnná nesmí souviset se zavádějícími proměnnými.

S ohledem na tuto podmínku je ideálním instrumentem taková proměnná, jejíž hodnoty se subjektům přiřazují náhodně, protože náhodné přiřazení zajišťuje nezávislost instrumentu na zavádějících proměnných. S takovou randomizovanou proměnnou se setkáváme například u experimentů s *designem s povzbuzením* (encouragement design).

Příkladem experimentu s designem s povzbuzením je následující situace. Společnost chce u dealerů svých produktů zjistit, jaký efekt na zisky z prodeje dealerů (proměnná  $Y$ ) má účast/neúčast dealerů na firemním školení (proměnná  $W$ ). Nemá ale přístup k zavádějícím proměnným  $X$  (například motivaci dealerů apod.) a přitom ony mohou ovlivnit jak ochotu dealerů účastnit se školení, tak jejich prodejní schopnosti ovlivňující zisky z prodeje.

V rámci experimentu s designem s povzbuzením tedy společnost rozešle **náhodně vybraným** dealerům motivační dopis, v němž je „povzbuzuje“

k účasti na školení. Obdržení dopisu nemá přímý vliv na zisky, ale ovlivňuje je prostřednictvím účasti na školení. Ovšem ne každý, kdo dopis dostal, se školení nakonec zúčastní a i mezi těmi, kdo dopis nedostali, mohou být tací, kteří se školení zúčastní. Indikátor skutečnosti, zda dealer motivační dopis dostal, nebo ne, je binární proměnná  $Z$ ,  $Z \in \{1, 0\}$ .

Proměnná  $Z$  vzhledem k randomizaci zcela jistě nesouvisí se zavádějícími proměnnými  $X$ , a tudíž jsou subpopulace těch, kteří dopis dostali, a těch, kteří dopis nedostali, s ohledem na hodnoty nepozorovatelných zavádějících proměnných stejné. Instrumentální proměnnou lze v kontextu observačních studií chápat jako analogii „povzbuzující“ proměnné v experimentech s designem s povzbuzením.

Místo efektu ošetření na odezvu se následně posuzuje efekt instrumentu na odezvu a v případě binárního instrumentu jde o porovnání odezvy mezi subpopulací s hodnotou instrumentu  $Z = 0$  a subpopulací s hodnotou instrumentu  $Z = 1$ . Tento efekt se nazývá „*intent-to-treat*“ efekt (*ITT*) a formálně bude zaveden v 4.1.3. V příkladu s dealery *ITT* efekt odpovídá rozdílu v ziscích prodeje mezi těmi, kdo motivační dopis dostali, a těmi, kdo jej nedostali.

Kdyby se všichni ti, kdo „povzbuzující“ dopis dostali, školení nakonec zúčastnili a naopak školení se nezúčastnil nikdo z těch, kdo dopis nedostali, potom by *ITT* efekt odpovídal průměrnému efektu ošetření v celé populaci. V tomto případě by všichni účastníci experimentu vykázali dokonalou shodu mezi hodnotou instrumentu  $Z$  a ošetřením  $W$ . Tedy pro  $Z_i = 1$  je i  $W_i = 1$  a analogicky pro  $Z_i = 0$  je i  $W_i = 0$ . Takovéto subjekty se v kontextu kauzální analýzy nazývají „*compliers*“.

Pokud ovšem některé subjekty tuto shodu nevykazují (podrobně bude tato skupina subjektů popsána níže), pak hodnotu *ITT* efektu nelze interpretovat jako průměrný efekt ošetření na celou populaci. Tato skutečnost byla ve statistické i ekonometrické literatuře dlouhou dobu ignorována. [Imbens a Angrist \(1994\)](#) a následně [Angrist, Imbens a Rubin \(1996\)](#) se s tímto problémem vypořádali tak, že koncept potenciálních odezev Rubinova kauzálního modelu upravili na modely s instrumentálními proměnnými.

### 4.1.2 Koncept potenciálních odezev v modelech s binární instrumentální proměnnou

Na rozdíl od kapitoly 2, kde bylo jednotlivým subjektům přiřazováno samotné ošetření  $W_i$ , bude nyní přiřazováno pouze „povzbuzení“ k ošetření  $Z_i$ . Toto povzbuzení k ošetření je v roli instrumentální proměnné, kde hodnota  $Z = 1$  reprezentuje přiřazení povzbuzení a hodnota  $Z = 0$  reprezentuje nepřičazení povzbuzení.

Ošetření  $W_i$  je nyní na instrumentu  $Z_i$  závislé, tedy  $W_i = W_i(Z_i)$ , a může nabývat dvou hodnot – jedničku, když  $i$ -tý subjekt obdrží aktivní ošetření, – a nulu, když  $i$ -tý subjekt obdrží kontrolní ošetření. Například platí-li pro subjekt  $i$ :  $W_i(1) = 1$ , pak subjekt  $i$  patří mezi „compliers“ – byl povzbuzen k ošetření a také dostal aktivní ošetření. Ovšem subjekt  $j$ , pro který  $W_j(1) = 0$ , nepatří mezi „compliers“, jelikož byl sice povzbuzen k ošetření, ale nakonec aktivní ošetření nedostal.

Vzhledem k tomu, že ošetření je závislé na instrumentu, dvojice  $W_i(0)$ ,  $W_i(1)$  představuje potenciální odezvy ošetření  $W$  na instrument  $Z$ .

Pokud jde o odezvu  $Y$ , její pozorovaná hodnota bude nyní ovlivněna jak přiřazenou hodnotou instrumentální proměnné  $Z$ , tak hodnotou ošetření  $W(Z)$ , tedy  $Y_i = Y_i(Z_i, W_i(Z_i))$ . Proto dvojice  $Y_i(0, W_i(0))$ ,  $Y_i(1, W_i(1))$  představuje potenciální odezvy proměnné  $Y$  na instrument  $Z$ .

**Potenciální odezvy jsou tedy na základě instrumentu  $Z$  definovány jak pro ošetření  $W$ , tak pro odezvu  $Y$ .**

Uvedené proměnné a jejich značení pro  $i$ -tý subjekt se formálně zavádí následně:

Randomizovaný instrument

$$Z_i \in \{0, 1\}$$

Pozorované ošetření

$$W_i = W_i(Z_i) = \begin{cases} W_i(0) & \text{pro } Z_i = 0 \\ W_i(1) & \text{pro } Z_i = 1 \end{cases}$$

Pozorovaná odezva

$$Y_i = Y_i(Z_i, W_i(Z_i)) = \begin{cases} Y_i(0, 0) & \text{pro } Z_i = 0, W_i(Z_i) = 0 \\ Y_i(0, 1) & \text{pro } Z_i = 0, W_i(Z_i) = 1 \\ Y_i(1, 0) & \text{pro } Z_i = 1, W_i(Z_i) = 0 \\ Y_i(1, 1) & \text{pro } Z_i = 1, W_i(Z_i) = 1 \end{cases}$$

Exogenita instrumentu<sup>17</sup> je v přístupu s potenciálními odezvami vyjádřena následně:

$$[Y_i(0), Y_i(1), W_i(0), W_i(1)] \perp\!\!\!\perp Z_i \quad (4.1.1)$$

### 4.1.3 Předpoklady pro kauzální inferenci v modelech s instrumentální proměnnou

V příkladu z úvodu sekce 4.1.1 byly zmíněny některé typy subjektů v závislosti na tom, jak reagovaly na povzbuzení ke školení. Typ „*compliers*“ byli ti, kteří na motivační dopis ( $Z = 1$ ) reagovali účastí na školení ( $W(1) = 1$ ) nebo kteří na absenci motivačního dopisu ( $Z = 0$ ) reagovali neúčastí na školení ( $W(0) = 0$ ).

Ve skupině subjektů mohou být ale i další *typy souladu* (compliance types)<sup>18</sup> mezi odezvou ošetření a instrumentální proměnnou. Takové subjekty, které bez ohledu na hodnotu instrumentu volí vždy aktivní ošetření, se nazývají typ „*always-taker*“ a subjekty, které bez ohledu na hodnotu instrumentu vždy volí kontrolní ošetření, se nazývají typ „*never-taker*“. Dalším typem jsou „*rebelové*“, u kterých je hodnota ošetření vždy jiná než hodnota instrumentu – ti se označují jako typ „*defier*“.

<sup>17</sup> Na rozdíl od konvenčně definované exogenity instrumentu v rámci lineárních modelů s instrumentálními proměnnými (předpoklad  $E(Z_i \cdot \varepsilon_i) = 0$ ) v modelu s potenciálními odezvami není potřeba specifikovat regresní funkci a tak definovat rezidua  $\varepsilon$ .

<sup>18</sup> Označení *complier*, *defier*, *always-taker*, *never-taker* je převzato z Angrist, Imbens a Rubin (1996).

Označuje-li symbol  $T_i$  typ  $i$ -tého subjektu s ohledem na shodu instrumentu a ošetření, pak lze zmíněné typy explicitně definovat následně:

$$T_i = \begin{cases} \text{complier} & \text{pro } W_i(0) = 0 \wedge W_i(1) = 1 \Leftrightarrow W_i(1) > W_i(0) \\ \text{defier} & \text{pro } W_i(0) = 1 \wedge W_i(1) = 0 \Leftrightarrow W_i(1) < W_i(0) \\ \text{always-taker} & \text{pro } W_i(0) = W_i(1) = 1 \\ \text{never-taker} & \text{pro } W_i(0) = W_i(1) = 0 \end{cases}$$

Místo „complier“, „defier“, „always-taker“ a „never-taker“ budou také používána písmena  $c, d, a, n$ , tedy  $T \in \{c, d, a, n\}$ .

Bohužel z pozorovaných dat nelze určit, zda konkrétní subjekt je typu „complier“, nebo ne. Např. pro  $i$ -tý subjekt, pro nějž je  $(Z_i, W_i(Z_i), Y_i) = (1, 1, y_i)$ , připadá v úvahu jak typ „complier“, tak typ „always-taker“. U tohoto subjektu nelze zjistit, jaká by byla realizace  $(1 - Z_i, W_i(1 - Z_i), Y_i)$ . Přesto bude pro určení odhadu kauzálního efektu ošetření  $W$  na odezvu  $Y$  nutné určit podíl „compliers“ v celé populaci, jak bude dále upřesněno.

Aby bylo možné prostřednictvím instrumentu  $Z$  odhadnout kauzální efekt ošetření  $W$  na odezvu  $Y$ , musí instrument splňovat předpoklady shrnuté na konci podkapitoly 4.1.3. Kauzální efekt instrumentu  $Z$  na odezvu  $Y$  u  $i$ -tého subjektu je potom definován jako rozdíl potenciálních odezev

$$Y_i(1, W_i(1)) - Y_i(0, W_i(0)), \quad (4.1.2)$$

přičemž z dvojice potenciálních odezev  $Y_i(1, W_i(1)), Y_i(0, W_i(0))$  je pouze jedna pozorovatelná. Ovšem za předpokladu, že platí SUTVA (mezi jednotlivými subjekty není žádná interference) a přiřazení hodnoty instrumentu  $Z$  jednotlivým subjektům je náhodné (díky randomizaci instrumentu  $Z$  je rozložení potenciálních odezev  $Y$  v obou skupinách pro  $Z = 1$  a pro  $Z = 0$  stejné), lze průměrný kauzální efekt instrumentu  $Z$  na odezvu  $Y$  identifikovat a vyjádřit následně:

$$\begin{aligned} ITT &= E[Y_i(1, W_i(1)) - Y_i(0, W_i(0))] = \\ &= E(Y \mid Z = 1) - E(Y \mid Z = 0) \end{aligned} \quad (4.1.3)$$

Efekt  $ITT$  ale není primárně zajímavý. Instrumentální proměnné slouží k tomu, aby byl jejich prostřednictvím zjištěn kauzální efekt **ošetření** na



odezvu. Subjekty typu „always-taker“ a „never-taker“ nenesou žádnou informaci o vlivu ošetření na odezvu. V tomto smyslu jsou zajímavé pouze subjekty typu „complier“, pro ty je koncept kontrafaktualů relevantní a informačně s ohledem na efekt ošetření na odezvu zajímavý.

Problém pro odhad efektu ošetření na odezvu naopak představují subjekty typu „defier“. Proto [Imbens a Angrist \(1994\)](#) zavedli pro každý subjekt  $i = 1, \dots, N$  předpoklad  $W_i(1) \geq W_i(0)$ . Tento předpoklad je známý jako předpoklad *monotonie*<sup>19</sup> a zajišťuje, že typ „defier“ nemůže nastat. (V praktických situacích může být obhajoba tohoto předpokladu obtížná.)

Má-li instrumentální proměnná  $Z$  přispět k odhadu efektu ošetření  $W$  na odezvu  $Y$ , musí být dále zajištěno, že  $Z$  na  $Y$  působí výlučně skrze ošetření  $W$ . Tedy že u subjektů např. typu „always-taker“ nebude mít hodnota instrumentu  $Z$  absolutně žádný vliv na odezvu  $Y$ . (V kontextu příkladu z úvodu kapitoly u takového dealera, jenž by se účastnil školení v každém případě bez ohledu na to, zda dostal, či nedostal motivační dopis, nesmí být hodnota jeho zisků z prodeje nikterak ovlivněna skutečností, zda motivační dopis dostal, nebo ne.) Tento předpoklad je známý jako předpoklad *restrikce na výlučnost* (*exclusion restriction*) a formálně jej lze vyjádřit následně:  $Y(1, W) = Y(0, W)$  pro  $W \in \{0, 1\}$ . Instrument je tak „vyloučen“ z možnosti ovlivnit hodnotu odezvy  $Y$  u těch, kdo jsou „non-compliers“. Proto stačí potenciální odezvy  $Y$  zapisovat pouze s ohledem na úroveň ošetření  $W$ , tedy  $(Y(0), Y(1))$ .

Poslední požadavek kladený na instrument  $Z$  garantuje jeho relevanci v tom smyslu, že instrument musí silně souviset s ošetřením  $W$ . V kontextu příkladu s dealery – pokud by byl podíl účastníků školení ve skupině s motivačním dopisem a ve skupině bez motivačního dopisu velmi podobný, pak motivační dopis jako instrument neposloužil. Nevygeneroval skupinu subjektů typu „complier“, kteří jediní jsou zajímaví s ohledem na informaci o efektu ošetření na odezvu. Tento předpoklad je známý jako předpoklad *relevance* (nebo také *nenulový kauzální efekt  $Z$  na  $W$* ) a formálně jej lze vyjádřit následně:  $E[W_i(1) - W_i(0)] \neq 0$ . Tento předpoklad je empiricky ověřitelný.

<sup>19</sup> Předpoklad monotonie je ekvivalentní s předpokladem „no-defiance“ ([Pearl, 2000](#)).

**Proměnná  $Z$  je tedy instrumentální proměnnou pro kauzální efekt ošetření  $W$  na odezvu  $Y$  tehdy, když platí:**

- 1) SUTVA pro  $W_i(z)$  i  $Y_i(z, w)$
- 2) randomizace instrumentu  $Z$  (nebo nezavádějící mechanismus přiřazování hodnoty instrumentu<sup>20</sup>  $Z$ )
- 3) monotonie
- 4) restrikce na výlučnost<sup>21</sup>
- 5) relevance

#### 4.1.4 Odhad efektu $W$ na $Y$ v modelu s binární instrumentální proměnnou

V úvodu kapitoly byl zmíněn problém s interpretací odhadů efektů ošetření v tradičních modelech s instrumentálními proměnnými za předpokladu heterogenity ošetření. [Imbens a Angrist \(1994\)](#) se s tímto problémem vypořádali tak, že místo odhadu  $ATE$  (průměrný efekt  $W$  na  $Y$  v rámci celé populace) se soustředili na odhad efektu ošetření  $W$  na odezvu  $Y$  pouze v subpopulaci subjektů typu „complier“.

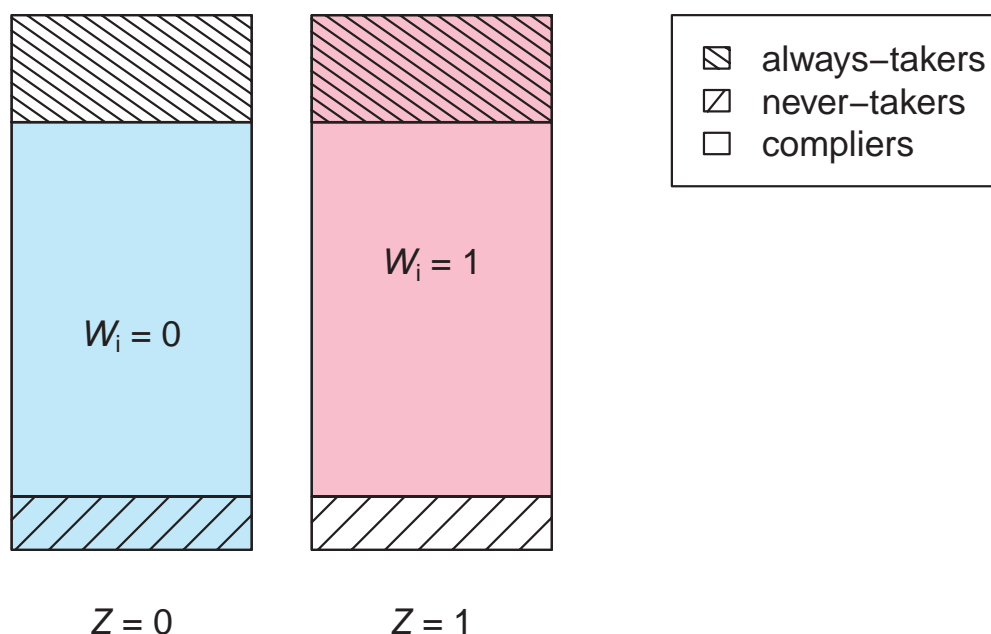
Efekt  $ITT$  (4.1.3) je efekt vyjádřený pro celou populaci, která se (za předpokladu monotonie) rozpadá do tří subpopulací – subjektů typu „never-taker“, „always-taker“ a „complier“. Například subpopulace subjektů typu „never-taker“ přirozeně nenese žádnou informaci o vlivu ošetření  $W$  na odezvu  $Y$ . Jedinou subpopulací, která je pro identifikaci efektu ošetření  $W$  zajímavá, je subpopulace subjektů typu „complier“. Tyto subjekty, jsou-li ve skupině s hodnotou instrumentu  $Z = 0$ , zůstanou bez ošetření, tedy pro ně je  $W = 0$ . A naopak jsou-li ve skupině s hodnotou instrumentu  $Z = 1$  pak jejich hodnota ošetření je  $W = 1$ . Průměrný efekt ošetření  $W$  na

<sup>20</sup> Pokud není k dispozici randomizovaný instrument (nebo i instrument z přirozeného experimentu), pak lze v observačních studiích za instrument uvažovat takovou proměnnou  $Z$ , jejíž hodnota je přiřazena nezavádějícím mechanismem s ohledem jak na potenciální odezvy ošetření  $W$  na  $Z$ , tak potenciální odezvy  $Y$  na  $Z$ . Obvykle tuto podmínku proměnná  $Z$  splní jen podmíněně na hodnotách nějakých vhodných kovariát  $X$ . Nalezení vhodného instrumentu je tedy mimo randomizaci obtížné.

<sup>21</sup> Předpoklad 1) a zároveň 4) zajišťují splnění předpokladu exogenity instrumentu, jak je známý z tradičních modelů s instrumentálními proměnnými.

odezvu  $Y$  v rámci této subpopulace lze tedy získat porovnáním průměrných odezev  $Y$  mezi oběma skupinami lišícími se úrovní instrumentu.

Problém ovšem představuje skutečnost, že z pozorovaných dat nelze přímo určit, který subjekt je typu „complier“. Například pro pozorované  $Y(Z, W(Z)) = Y(1, 1)$  nelze zjistit, zda se jedná o typ „complier“, nebo „always-taker“. Analogicky u  $Y(0, 0)$  není zřejmé, zda je subjekt typu „complier“, nebo „never-taker“. Na druhou stranu  $Y(0, 1)$  je za platných předpokladů určitě „always-taker“ a  $Y(1, 0)$  je určitě „never-taker“. Co určit lze, je podíl jednotlivých subpopulací v celé populaci. (Názorně na obrázku 4.1.)



**Obrázek 4.1:** Podíl „always-takers“, „never-takers“ a „compliers“ je díky randomizaci instrumentu  $Z$  stejný jak ve skupině s hodnotou instrumentu  $Z = 0$ , tak ve skupině s hodnotou instrumentu  $Z = 1$ . V rámci skupiny subjektů s hodnotou instrumentu  $Z = 0$  je modrou barvou vyjádřena skupina subjektů s kontrolním ošetřením  $W_i = 0$ , mezi nimiž jsou subjekty typu „never-taker“ a subjekty typu „complier“, jež nelze navzájem odlišit. Bílou barvou jsou ale jednoznačně určeny subjekty typu „always-taker“. Jejich podíl lze tedy určit jako podíl subjektů s aktivním ošetřením mezi všemi subjekty s hodnotou instrumentu  $Z = 0$ . Analogicky ve skupině subjektů s hodnotou instrumentu  $Z = 1$  je růžovou barvou vyjádřena skupina subjektů s aktivním ošetřením  $W_i = 1$ , mezi nimiž jsou subjekty typu „always-taker“ a subjekty typu „complier“, jež nelze od sebe odlišit. Bílou barvou jsou ale jednoznačně určeny subjekty typu „never-taker“. Jejich podíl lze tedy určit jako podíl subjektů s kontrolním ošetřením mezi všemi subjekty s hodnotou instrumentu  $Z = 1$ .

Pravděpodobnost příslušnosti libovolného subjektu  $i$  k jednotlivým subpopulacím bude označena následně:  $P_t = P(T_i = t)$  pro  $t \in \{c, a, n\}$ . („Defiers“ není potřeba za předpokladu monotonie uvažovat.) Potom lze podíl jednotlivých subpopulací „always-takers“, „never-takers“ a „compliers“ v rámci celé populace identifikovat ve třech krocích:

- Nejdříve bude určen podíl „always-takers“. V rámci skupiny těch, pro které je instrument  $Z = 0$ , lze pozorovat aktivní ošetření  $W = 1$  pouze u „always-takers“. Podíl „always-takers“ v celé populaci lze tedy vyjádřit jako pravděpodobnost  $P_a = P(W_i = 1 \mid Z = 0)$ . (Vzhledem k randomizaci instrumentu  $Z$  je podíl všech subpopulací ve skupině s instrumentem  $Z = 1$  i ve skupině s instrumentem  $Z = 0$  stejný.)
- Podíl „never-takers“ bude určen analogicky. V rámci skupiny těch, pro které je instrument  $Z = 1$ , lze pozorovat kontrolní ošetření  $W = 0$  pouze u „never-takers“. Podíl „never-takers“ v celé populaci lze tedy vyjádřit jako pravděpodobnost  $P_n = P(W_i = 0 \mid Z = 1)$ .
- Nakonec bude určen podíl „compliers“<sup>22</sup> pomocí předchozích pravděpodobností:  

$$P_c = 1 - P_a - P_n.$$

Efekt  $ITT$  (4.1.3) lze nyní rozložit:

$$ITT = ITT_c \cdot P_c + ITT_a \cdot P_a + ITT_n \cdot P_n \quad (4.1.4)$$

kde

$$ITT_c = E[Y_i(1, W_i(1)) - Y_i(0, W_i(0)) \mid W_i(1) = 1, W_i(0) = 0]$$

$$ITT_a = E[Y_i(1, W_i(1)) - Y_i(0, W_i(0)) \mid W_i(1) = 1, W_i(0) = 1]$$

$$ITT_n = E[Y_i(1, W_i(1)) - Y_i(0, W_i(0)) \mid W_i(1) = 0, W_i(0) = 0]$$

Za předpokladu restrikce na výlučnost (exclusion restriction) je vyloučen vliv  $Z$  na  $Y$  jinak než výhradně skrze  $W$ . Proto  $ITT_a = 0 = ITT_n$

<sup>22</sup> Podíl „compliers“ lze vyjádřit jako  $P_c = P[W_i(1) > W_i(0)] = E[W_i(1) - W_i(0)] = E[W_i \mid Z_i = 1] - E[W_i \mid Z_i = 0]$ . Poslední rovnost plyne z randomizace instrumentu  $Z$ . Podrobněji v příloze A.2.

a tedy 4.1.4 lze upravit na

$$ITT = ITT_c \cdot P_c \quad (4.1.5)$$

Efekt  $ITT_c$  Imbens a Angrist (1994) nazvali *střední hodnota lokálního efektu ošetření* (Local Average Treatment Effect) a pro tento efekt se používá symbol  $\tau_{LATE}$ . Tedy

$$\begin{aligned} \tau_{LATE} = ITT_c &= E[Y_i(1) - Y_i(0) \mid W_i(1) = 1, W_i(0) = 0] \quad (4.1.6) \\ &= E[Y_i(1) - Y_i(0) \mid T_i = complier] \end{aligned}$$

Efekt  $\tau_{LATE}$  je tedy efekt ošetření  $W$  na odezvu  $Y$  v rámci subpopulace „compliers“ a lze jej z 4.1.5 vyjádřit jako podíl ITT a podílu „compliers“ v rámci celé populace.

$$\tau_{LATE} = \frac{ITT}{P_c} \quad (4.1.7)$$

Z praktického pohledu ovšem není nutné odhadovat  $\tau_{LATE}$  prostřednictvím rozkladu  $ITT$ , jelikož Imbens a Angrist (1994) dokázali (podrobně v příloze A.2), že

$$\begin{aligned} \tau_{LATE} &= \frac{E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]}{E[W_i \mid Z_i = 1] - E[W_i \mid Z_i = 0]} = \\ &= \frac{E[Y_i(Z_i - E(Z_i))]}{E[W_i(Z_i - E(Z_i))]} = \end{aligned} \quad (4.1.8)$$

$$= \frac{\text{cov}(Y_i, Z_i)}{\text{cov}(W_i, Z_i)}. \quad (4.1.9)$$

Efekt  $\tau_{LATE}$  pak lze u uvažovaného binárního instrumentu (a bez dalších vysvětlujících proměnných) odhadnout pomocí Waldovy statistiky nebo obecně pomocí metody dvojstupňového modelu známého jako 2SLS (two-stage-least-squares), která se tradičně v metodách s instrumentálními proměnnými používá. Metoda 2SLS je podrobně popsána v příloze A.1.

Waldův odhad efektu  $\tau_{LATE}$  má tvar:

$$\hat{\tau}_{LATE} = \frac{\frac{1}{N_1} \sum_{i=1}^N y_i z_i - \frac{1}{N_0} \sum_{i=1}^N y_i (1 - z_i)}{\frac{1}{N_1} \sum_{i=1}^N w_i z_i - \frac{1}{N_0} \sum_{i=1}^N w_i (1 - z_i)}, \quad (4.1.10)$$

kde  $N_1$  je počet subjektů pro které je instrument  $Z = 1$ , tedy  $N_1 = \sum_{i=1}^N z_i$ , a  $N_0$  je počet subjektů, pro které je instrument  $Z = 0$ , tedy  $N_0 = N - N_1$ .

Waldův odhad  $\hat{\tau}_{LATE}$  je konzistentní, pro konečné soubory je však vychýlený (zejména jde o rozsah podsouboru „compliers“). Čím slabší je instrument  $Z$  (t.j.  $\text{cov}(W, Z) \rightarrow 0$ ), tím je vychýlení větší. Důsledky slabých instrumentů jsou popsány např. v [Bound, Jaeger a Baker \(1995\)](#).

#### 4.1.5 Interpretace efektu $\tau_{LATE}$

Nyní je zřejmé, že za předpokladu heterogenity odhady efektu ošetření získané pomocí instrumentální proměnné odhadují ve skutečnosti kauzální vliv ošetření na odezvu **v rámci subpopulace „compliers“**. Přesto může být interpretace odhadů obtížná.

Subpopulace „compliers“ závisí na volbě instrumentu, tedy pro různé instrumenty může být různá. Proto se odhady efektu ošetření získané prostřednictvím různých instrumentů mohou navzájem lišit. Navíc nemusí být jasné, kdo vlastně „compliers“ jsou, a pro některé situace nemusí být efekt  $\tau_{LATE}$  zajímavý a  $\tau_{ATE}$  či  $\tau_{ATT}$  nelze bez dalších omezujících předpokladů identifikovat.

V některých aplikačních úlohách ale lze například z povahy problému předpokládat jednostrannou „non-compliance“, kde pro všechny subjekty  $i$  platí  $W_i(0) = 0$  (tento předpoklad tedy vylučuje „always-takers“ a „defiers“). Potom  $\tau_{LATE} = \tau_{ATT}$ .

Pro binární instrumentální proměnnou  $Z$  a binární ošetření  $W$  lze  $\tau_{LATE}$  odhadnout neparametricky pomocí Waldova odhadu. Metody *2SLS* jsou univerzálnější v tom smyslu, že umožňují i modely s víceúrovňovým či spojitým ošetřením, s víceúrovňovým či spojitým instrumentem a s možností začlenit do modelu další pozorovatelné vysvětlující proměnné (ne zavádějící!). V tradičním ekonometrickém přístupu ovšem umožňují pouze odhady za předpokladu homogenity ošetření, což je v praxi často nerealistický předpoklad.

#### 4.1.6 Podmíněný instrument

V aplikačních úlohách je často nemožné získat plně randomizovaný instrument. Někdy je ale k dispozici taková proměnná, která může plnit funkci validního instrumentu, pokud je podmíněna na nějakých pozorovatelných zavádějících kovariátách. *LATE* odhady za těchto okolností popsal [Frölich \(2007\)](#) a zobecnili [Hong a Nekipelov \(2010\)](#).

Odhady efektů pomocí instrumentální proměnné podmíněné na zavádějících kovariátách lze v rámci prostředí **R** získat pomocí funkce `estlate` v balíku `ri` od autorů Peter M. Aronow a Cyrus Samii.<sup>23</sup>

## 4.2 Regresní diskontinuita

### 4.2.1 Úvod k modelům regresní diskontinuity

Metody regresní diskontinuity (RD) jsou užitečné v situacích, kdy se subjektům přiřazuje ošetření  $W$  jednoznačně (nebo téměř jednoznačně) na základě hodnoty skalární **spojité** kovariáty  $X$ . Tato kovariáta ovšem souvisí buď přímo, nebo prostřednictvím dalších nepozorovatelných zavádějících proměnných také s potenciálními odezvami  $Y(0)$  a  $Y(1)$ .

Subjektům, u kterých je hodnota kovariáty  $X$  menší než hraniční hodnota  $c$ , je přiřazeno kontrolní ošetření, a subjektům, u kterých je hodnota kovariáty  $X$  větší než hraniční hodnota  $c$ , je přiřazeno aktivní ošetření. Odezva  $Y$  tedy v okolí hranice  $c$  „poskočí“, má-li ošetření na odezvu skutečný efekt. Takto vnesenou „diskontinuitu“ do jinak spojité odezvy  $Y$  v bodě  $c$  je pak možné interpretovat jako efekt ošetření na odezvu. Subjekty s hodnotami  $X$  velmi blízko hranice  $c$  jsou si velmi podobné, a tak na ně lze nahlížet jako na subjekty v rámci „randomizovaného experimentu“. Proměnná  $X$ , na základě které se rozhoduje o přiřazení aktivního, či kontrolního ošetření, se nazývá *účinná proměnná* (forcing variable nebo také running variable).

Původní idea regresní diskontinuity je již dlouho známá ([Thistlethwaite a Campbell, 1960](#)), ovšem po zavedení Rubinova kauzálního modelu a jeho implementaci do modelů RD zažívá metoda svoji renesanci a je extrémně populární zejména v oblasti evaluace programů. V roce 2008 jí například bylo věnováno speciální číslo v *Journal of Econometrics*. Historické využití RD ve statistice, psychologii a ekonomii je dobře popsáno v [Cook \(2008\)](#), přehledové články popisující moderní modely s RD, jejich varianty a aplikace jsou například v [Imbens a Lemieux \(2008\)](#), [Lee a Lemieux \(2010\)](#) či [Van der Klaauw \(2008\)](#).

<sup>23</sup> Funkce využívá úpravu 2SLS tak, jak je popsána v kapitole 4.5.2 v [Angrist a Pischke \(2009\)](#). Funkce dále umožňuje použití vah odpovídající podílu compliers, jak je popsáno v [Aronow a Carnegie \(2013\)](#).

Základní RD modely budou nejdříve ilustrovány na následující situaci.

Určitý benefit kreditní karty je nabízen pouze klientům se solidní platební historií měřenou například pomocí *FICO* skóru. Aktivnímu ošetření odpovídá kreditní karta s benefitem, kontrolnímu karta bez benefitu. *FICO* skór odpovídá spojité proměnné  $X$ , na základě které je uskutečněna nabídka benefitu ke kartě. Těm klientům, u nichž je *FICO* skór větší než stanovená hraniční hodnota  $c$ , je bankou učiněna nabídka benefitu ke kartě. V případě *FICO* skóru se tedy jedná o účinnou proměnnou. Odezvou  $Y$  je například objem půjček držitele karty a cílem marketérů banky je zjistit, zda má nabízený benefit skutečný efekt na odezvu.

Pokud by banka v nabízení benefitu nedělala žádné výjimky a všichni klienti, kterým byla učiněna nabídka benefitu, by tento benefit skutečně využili, pak by se jednalo o model s *ostrou* regresní diskontinuitou (sharp regression discontinuity). Realističtější je ale situace, že ne všichni klienti osloveni nabídkou ji skutečně využijí, a naopak, některým klientům těsně pod hranicí  $c$  může banka benefit také mimořádně nabídnout. V tomto případě mluvíme o *neostré* regresní diskontinuitě (fuzzy regression discontinuity).

#### 4.2.2 Model ostré regresní diskontinuity

Každému subjektu  $i$  je přiřazeno binární ošetření  $W_i \in \{0, 1\}$  na základě hodnoty účinné proměnné  $X_i$ . Ta je spojitá a jednoznačně determinuje hodnotu ošetření podle hraničního bodu  $c$  následovně:

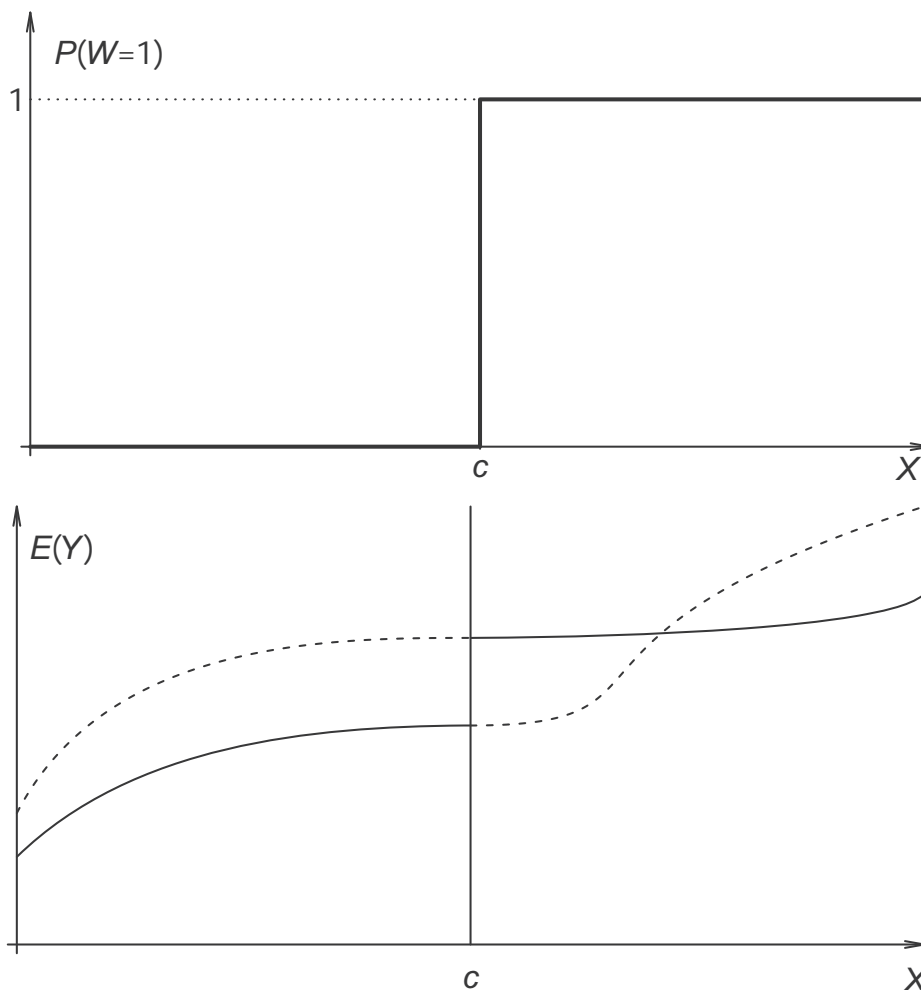
$$W_i = \begin{cases} 1 & \text{pro } X_i \geq c \\ 0 & \text{pro } X_i < c, \end{cases}$$

či alternativně pomocí indikátorové funkce  $\underline{1}$ ;  $W_i = \underline{1}(X_i \geq c)$ , kde

$$\underline{1}(\cdot) = \begin{cases} 1 & \text{když výrok v závorce je pravdivý} \\ 0 & \text{když výrok v závorce je nepravdivý.} \end{cases}$$

Pravděpodobnost přiřazení aktivního ošetření v závislosti na hodnotě účinné proměnné  $X$ ,  $P[W = 1 \mid X = x]$ , je ilustrována na obrázku 4.2 nahoře. Funkce  $P[W = 1 \mid X = x]$  je všude konstantní s výjimkou bodu  $c$ , kde „ostře“ poskočí z hodnoty 0 na hodnotu 1 (s tím souvisí termín





**Obrázek 4.2:** Ostrá RD

Graf nahore: na svislé ose je pravděpodobnost přiřazení ošetření v závislosti na  $x$ ,  $P(W = 1 | X = x)$ ; na vodorovné ose je účinná proměnná  $X$ .

Grafy dole: na svislé ose je  $E(Y)$ . Grafy představují střední hodnoty potenciálních odezev v závislosti na účinné proměnné  $X$ ,  $E[Y(0) | X = x]$  a  $E[Y(1) | X = x]$ . Čárkovaná čára odpovídá nepozorovatelným středním hodnotám, plná čára odpovídá pozorovatelným; rozdíl funkčních hodnot v bodě  $c$  odpovídá efektu  $\tau_{SRD}$  (4.2.2).

„sharp“ RD). V hraničním bodě  $X_i = c$  lze považovat přiřazení ošetření za „náhodné“.

Účinná proměnná  $X_i$  **souvisí** s hodnotami odezvy  $Y_i(0)$  i  $Y_i(1)$  (ať už přímo, nebo prostřednictvím dalších kovariát), ale vliv ošetření na odezvu s kontrolou této proměnné není možný, jelikož neplatí předpoklad překryvu (3.1.2). Pro hodnoty účinné proměnné  $X_i < c$  je z obou potenciálních odezev možné pozorovat výlučně odezvu  $Y_i(0)$  a pro hodnoty  $X_i \geq c$  je z obou potenciálních odezev možné pozorovat výlučně odezvu  $Y_i(1)$ .

Na obrázku 4.2 dole jsou dva grafy obou středních hodnot **potenciálních** odezev v závislosti na  $x$ ,  $E[Y(0) | X = x]$  a  $E[Y(1) | X = x]$ . Obě podmíněné střední hodnoty potenciálních odezev jsou spojité funkce. Třetí graf (překrývá po částech předchozí dva grafy) s plnou čarou reprezentuje střední hodnoty **pozorovaných** odezev v závislosti na  $x$ ,  $E[Y | X = x]$ .

$$E[Y | X = x] = E[Y | W = 0, X = x] \cdot P[W = 0 | X = x] + E[Y | W = 1, X = x] \cdot P[W = 1 | X = x] \quad (4.2.1)$$

Střední hodnota pozorovaných odezev je funkce v bodě  $X = c$  nespojitá. Velikost „skoku“ v bodě  $c$  odpovídá střední hodnotě lokálního efektu ošetření (*LATE*) na subjekty v blízkém okolí bodu  $c$  a v kontextu ostré regresní diskontinuity se značí  $\tau_{SRD}$ , tedy

$$\tau_{SRD} = E[Y_i(1) - Y_i(0) | X_i = c]. \quad (4.2.2)$$

Protože neplatí předpoklad překryvu, nelze se při identifikaci efektu ošetření  $\tau_{SRD}$  vyhnout extrapolaci (ta je ovšem dostačující v blízkém okolí bodu  $c$ ). Pro odhad efektu  $\tau_{SRD}$  se pak nabízí rozdíl limit v okolí bodu  $c$ ,

$$\lim_{x \rightarrow c^+} E[Y_i | X_i = x] - \lim_{x \rightarrow c^-} E[Y_i | X_i = x].$$

Aby bylo možné pomocí tohoto rozdílu  $\tau_{SRD}$  skutečně identifikovat, musí platit již zmíněná spojitost středních hodnot<sup>24</sup> potenciálních odezev, tedy musí platit předpoklad, že

funkce  $E[Y(0) | X = x]$  a funkce  $E[Y(1) | X = x]$  jsou spojité pro všechna<sup>25</sup>  $x$ .

<sup>24</sup> Alternativně lze použít i tvrdší předpoklad spojitosti podmíněných distribučních funkcí  $F_{Y(0)|X}(y | x)$  a  $F_{Y(1)|X}(y | x)$ .

<sup>25</sup> Zejména musí být spojité v bodě  $c$ . Požadavek spojitosti na všechna  $x$  plyne spíše z praktického pohledu v tom smyslu, že by bylo neobvyklé očekávat spojitost speciálně jen v bodě  $c$ .

Potom

$$\begin{aligned}
E[Y(0) | X = c] &= \lim_{x \rightarrow c^-} E[Y(0) | X = x] = \\
&= \lim_{x \rightarrow c^-} E[Y(0) | W = 0, X = x] = \\
&= \lim_{x \rightarrow c^-} E[Y | X = x] \\
E[Y(1) | X = c] &= \lim_{x \rightarrow c^+} E[Y(1) | X = x] = \\
&= \lim_{x \rightarrow c^+} E[Y(1) | W = 1, X = x] = \\
&= \lim_{x \rightarrow c^+} E[Y | X = x]
\end{aligned} \tag{4.2.3}$$

Z 4.2.2 a 4.2.3 tedy plyne:

$$\tau_{SRD} = \lim_{x \rightarrow c^+} E[Y | X = x] - \lim_{x \rightarrow c^-} E[Y | X = x] \tag{4.2.4}$$

Díky předpokladu spojitosti středních hodnot potenciálních odezev tedy lze nesplněný předpoklad překryvu kompenzovat extrapolací těchto středních hodnot v blízkém okolí hranice  $c$ . Pro odhad  $\tau_{SRD}$  tedy stačí odhadnout například obě regresní funkce  $Y$  na  $X$  pro  $x$  vpravo i vlevo od hranice  $c$ , určit jejich hodnoty v bodě  $c$  a spočítat rozdíl těchto hodnot. (Podmíněné střední hodnoty potenciálních odezev lze v levém i pravém okolí odhadovat i neparametricky.)

### 4.2.3 Model neostré regresní diskontinuity

V ostrém modelu RD subjekty s hodnotami účinné proměnné  $X$  vlevo od hranice  $c$  vždy dostaly kontrolní ošetření, tedy  $P[W = 1 | X < c] = 0$ , a subjekty s hodnotami účinné proměnné  $X$  vpravo od hranice  $c$  vždy dostaly aktivní ošetření, tedy  $P[W = 1 | X \geq c] = 1$ . Pravděpodobnost přiřazení aktivního ošetření byla tedy výhradně buď nula, nebo jedna,  $P[W = 1 | X = x] \in \{0, 1\}$ . K nárůstu této pravděpodobnosti z hodnoty nula do hodnoty jedna docházelo v hraničním bodě  $X = c$ .

V neostrém modelu RD se u pravděpodobnosti aktivního ošetření v hraničním bodě  $c$  „skok z nuly na jedničku“ neočekává, očekává se pouze diskontinuita. Tedy

$$\lim_{x \rightarrow c^-} P[W_i = 1 | X_i = x] \neq \lim_{x \rightarrow c^+} P[W_i = 1 | X_i = x] \tag{4.2.5}$$

V příkladu s kreditními kartami klienti s hodnotami *FICO* skóru jak vlevo, tak i vpravo od hraniční hodnoty  $c$  mohou být držiteli benefitu ke kartě, ovšem pravděpodobnost vlastnění karty s benefitem prudce vzroste pro  $FICO = c$ . Klientům s *FICO* skórem s hodnotou alespoň  $c$  banka totiž poslala dopis s nabídkou benefitu ke kartě, klientům s *FICO* skórem menším než  $c$  banka dopis neposlala. Přesto i takovíto klienti mohou o benefit zažádat a případně jej získat a naopak ne všichni ti, kdo dopis s nabídkou dostali, skutečně o kartu s benefitem zažádali. Proto pravděpodobnostní funkce přiřazení aktivního ošetření (držení karty s benefitem) v závislosti na hodnotách  $x$  v hraničním bodě  $c$  „prudce vzroste“, ovšem již ne z nuly na jedničku, jak tomu bylo u ostré RD.

Hranice  $c$  se tak ve vztahu k hodnotě účinné proměnné dostává do role instrumentální proměnné. Pro subjekty v blízkém okolí hranice  $c$  tak lze definovat kategorie „always-taker“, „never-taker“ a „complier“. Subjekt je

$$\left\{ \begin{array}{l} \text{complier v bodě } x = c, \text{ když} \\ \lim_{x \rightarrow c^-} P[W_i = 1 \mid X_i = x] = 0 \wedge \lim_{x \rightarrow c^+} P[W_i = 1 \mid X_i = x] = 1 \\ \text{never-taker v bodě } x = c, \text{ když} \\ \lim_{x \rightarrow c^-} P[W_i = 1 \mid X_i = x] = 0 \wedge \lim_{x \rightarrow c^+} P[W_i = 1 \mid X_i = x] = 0 \\ \text{always-taker v bodě } x = c, \text{ když} \\ \lim_{x \rightarrow c^-} P[W_i = 1 \mid X_i = x] = 1 \wedge \lim_{x \rightarrow c^+} P[W_i = 1 \mid X_i = x] = 1 \end{array} \right.$$

„Compliers“ jsou tedy subjekty, kterým by bylo přiřazeno aktivní ošetření, kdyby hranice  $c$  byla menší nebo rovna  $X_i$ , a kontrolní ošetření, kdyby hranice  $c$  byla větší než  $X_i$ . Hodnota jejich ošetření je tedy jednoznačně určena jejich  $x$ -ovou polohou ve vztahu k  $c$ .

Vzhledem k existenci „compliers“ a „noncompliers“ v rámci neostré RD nelze předpokládat homogenitu efektu ošetření pro jednotlivé subjekty, a tak (analogicky k parametrům v rámci modelů s instrumentální proměnnou) lze identifikovat střední hodnotu lokálního efektu ošetření (*LATE*) pouze pro „compliers“ a pro  $X = c$  (prakticky subjekty v blízkém okolí bodu  $c$ ).

Střední hodnota lokálního efektu ošetření se v kontextu neostré regresní diskontinuity značí  $\tau_{FRD}$  a [Hahn, Todd a Van der Klaauw \(2001\)](#) dokázali,

že pro ni platí:

$$\begin{aligned}\tau_{FRD} &= E[Y_i(1) - Y_i(0) \mid \text{subjekt } i \text{ je complier} \wedge X_i = c] = \\ &= \frac{\lim_{x \rightarrow c^+} E[Y_i \mid X_i = x] - \lim_{x \rightarrow c^-} E[Y_i \mid X_i = x]}{\lim_{x \rightarrow c^+} E[W_i \mid X_i = x] - \lim_{x \rightarrow c^-} E[W_i \mid X_i = x]}\end{aligned}\quad (4.2.6)$$

Pro identifikaci střední hodnoty efektu ošetření na subjekty v blízkém okolí bodu  $c$  (lokální efekt *LATE*) je kromě předpokladu spojitých středních hodnot podmíněných potenciálních odezev (analogicky k ostrým RD modelům) potřeba zajistit i splnění předpokladu monotonie, tedy že  $W_i(x)$  je neklesající funkce vzhledem k  $x$  pro  $x = c$ . (Restrikce na výlučnost je splněna implicitně.)

Pravděpodobnost přiřazení aktivního ošetření v závislosti na hodnotě účinné proměnné  $X$ ,  $P[W = 1 \mid X = x]$  (alternativně zapsáno  $E[W \mid X = x]$ ), je ilustrována na obrázku 4.3 nahoře. Pro  $x = c$  lze podíl „always-takers“, „never-takers“ a „compliers“ (obrázek 4.3 nahoře) vyjádřit za předpokladu monotonie následovně (význam značení  $P_t$  je analogický značení v sekci 4.1.4):

$$P_a = \lim_{x \rightarrow c^-} P[W = 1 \mid X = x]$$

$$P_n = 1 - \lim_{x \rightarrow c^+} P[W = 1 \mid X = x]$$

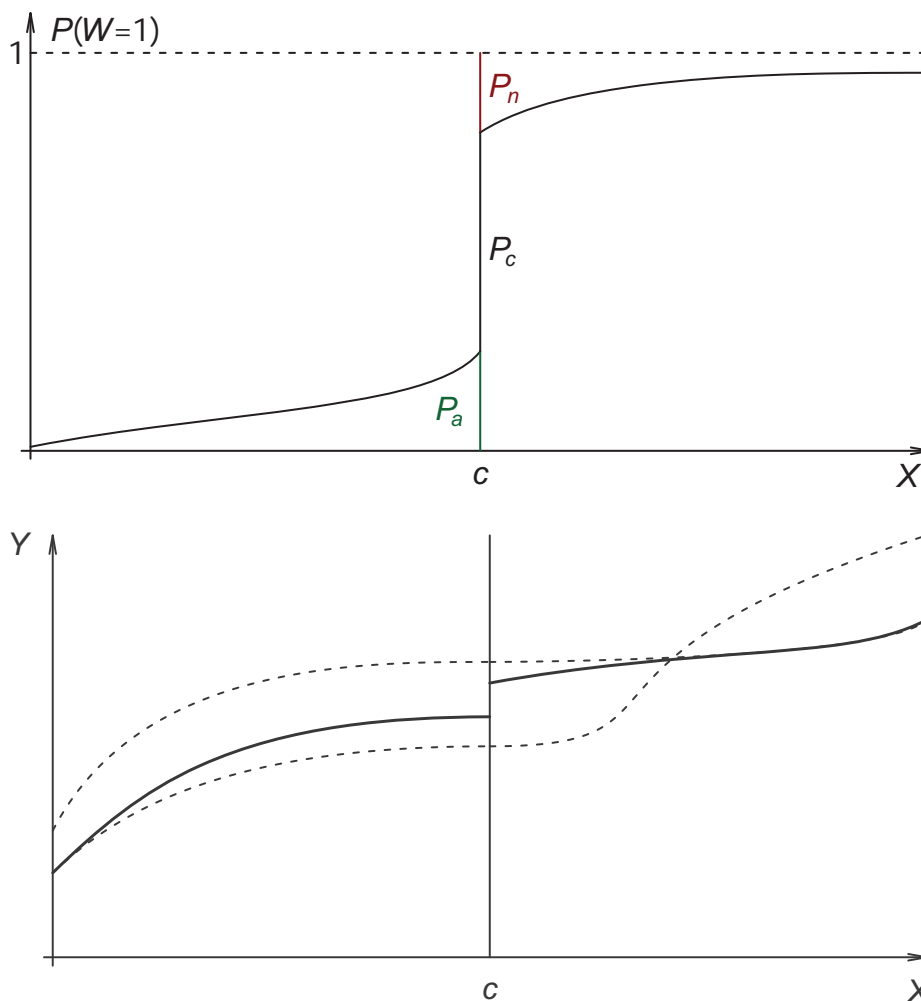
$$P_c = 1 - P_a - P_n = \lim_{x \rightarrow c^+} P[W = 1 \mid X = x] - \lim_{x \rightarrow c^-} P[W = 1 \mid X = x]$$

Podíl „compliers“,  $P_c$ , odpovídá jmenovateli ve výrazu 4.2.6.

Na obrázku 4.3 dole jsou grafy dvou nepozorovatelných středních hodnot **potenciálních** odezev v závislosti na  $x$ ,  $E[Y(0) \mid W = 0, X = x]$  a  $E[Y(1) \mid W = 1, X = x]$ . Třetí graf s plnou čarou reprezentuje střední hodnoty **pozorovaných** odezev v závislosti na  $x$ :

$$\begin{aligned}E[Y \mid X = x] &= E[Y(0) \mid W = 0, X = x] \cdot P[W = 0 \mid X = x] + \\ &+ E[Y(1) \mid W = 1, X = x] \cdot P[W = 1 \mid X = x]\end{aligned}$$

„Skok“ v bodě  $x = c$  v tomto třetím grafu odpovídá čitateli ve výrazu 4.2.6.



**Obrázek 4.3:** Neostrá RD

Graf nahoře: na svislé ose je pravděpodobnost přiřazení aktivního ošetření v závislosti na  $x$ ,  $P(W = 1 | X = x)$ ; na vodorovné ose je účinná proměnná  $X$ .

Grafy dole: na svislé ose je  $E(Y)$ . Čárkovaný graf odpovídá nepozorovatelným středním hodnotám potenciálních odezev  $E[Y(0) | W = 0, X = x]$  a  $E[Y(1) | W = 1, X = x]$  a plný graf odpovídá pozorovaným středním hodnotám  $E(Y | X = x)$ ; rozdíl funkčních hodnot pro plné čáry v bodě  $c$  **není** efekt  $\tau_{FRD}$ , pouze odpovídá čitateli ve výrazu 4.2.6.

#### 4.2.4 Interní a externí validita

$\tau_{SRD}$  je validní pouze pro subjekty v blízkém okolí  $c$  a  $\tau_{FRD}$  je validní pouze pro subjekty v blízkém okolí  $c$ , které jsou navíc „compliers“.

Odhady tedy nemají externí validitu. Ovšem subpopulace „compliers“ může být zajímavá, pokud reprezentuje tu část populace, která by byla silně zasažena změnou nastavení hodnoty  $c$ .

### 4.2.5 Odhad efektu ošetření v modelech s ostrou i neostrou regresní diskontinuitou

Pro identifikaci  $\tau_{SRD}$  i  $\tau_{FRD}$  nebylo nutné znát funkcionální formu vztahu mezi  $Y$  a  $X$ . Odhad všech potřebných limit lze proto získat i neparametricky.

Historicky se používaly parametrické odhady na základě regresních (včetně nelineárních) modelů, jež využívaly všechna pozorování v datovém souboru. V současnosti je populární lokální lineární regrese, jak ji popsali např. [Imbens a Lemieux \(2008\)](#).



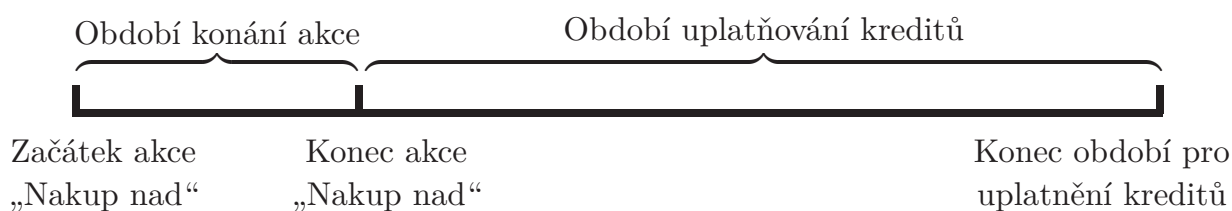


## Kapitola 5

# Aplikace metod odhadu kauzálního efektu na observačních datech s regulárním mechanismem přiřazování

### 5.1 Vliv akce „Nakup nad“ na zisky

V této kapitole bude zjišťován vliv účasti v akci na podporu prodeje „Nakup nad“ na průměrný zisk připadající na uživatele v různě definovaných obdobích. Uživatelé, kteří v období od 16. 11. 2017 do 19. 11. 2017 uskutečnili nákupy v součtu za aspoň 500 Kč, získali kredit. Ten mohli ve formě slevy uplatnit na další nákupy v období od 20. 11. 2017 do 30. 11. 2017 (obrázek 5.1). V tomto období bylo možné uplatnit i slevy za kredity z jiných akcí na podporu prodeje.



**Obrázek 5.1:** Časový průběh akce „Nakup nad“ a následného období uplatňování kreditů

Je rozumné očekávat, že v období konání akce by měly být prodeje a tím i zisky v důsledku akce „Nakup nad“ vyšší než obvykle. Na druhou stranu uplatněné kredity představují pro společnost  $F$  náklady, jež zisky společnosti v období, kdy je možné uplatnit kredity, výrazně snižují, a to dokonce ve skupině účastníků akce až do záporných hodnot.

Vzniká tedy otázka, jak moc a jestli vůbec se akce „Nakup nad“ společnosti  $F$  vyplácí. V období konání akce část uživatelů nakupuje více a při splnění podmínky celkové útraty nad 500 Kč získává kredity. Nicméně není jasné, zda by právě tato část uživatelů nenakupovala více i bez uvažované akce na podporu prodeje z důvodů jejich individuálních vlastností souvisejících s ochotou nakupovat. Pokud ano, pak by těmto uživatelům přiřazené a později uplatněné kredity představovaly pro společnost  $F$  zbytečné náklady. Proto společnost potřebuje zjistit, jestli se jí tato akce na podporu prodeje vyplácí, což v kontextu kauzální analýzy znamená zjistit nevychýlený efekt účasti v akci „Nakup nad“ na zisky společnosti.

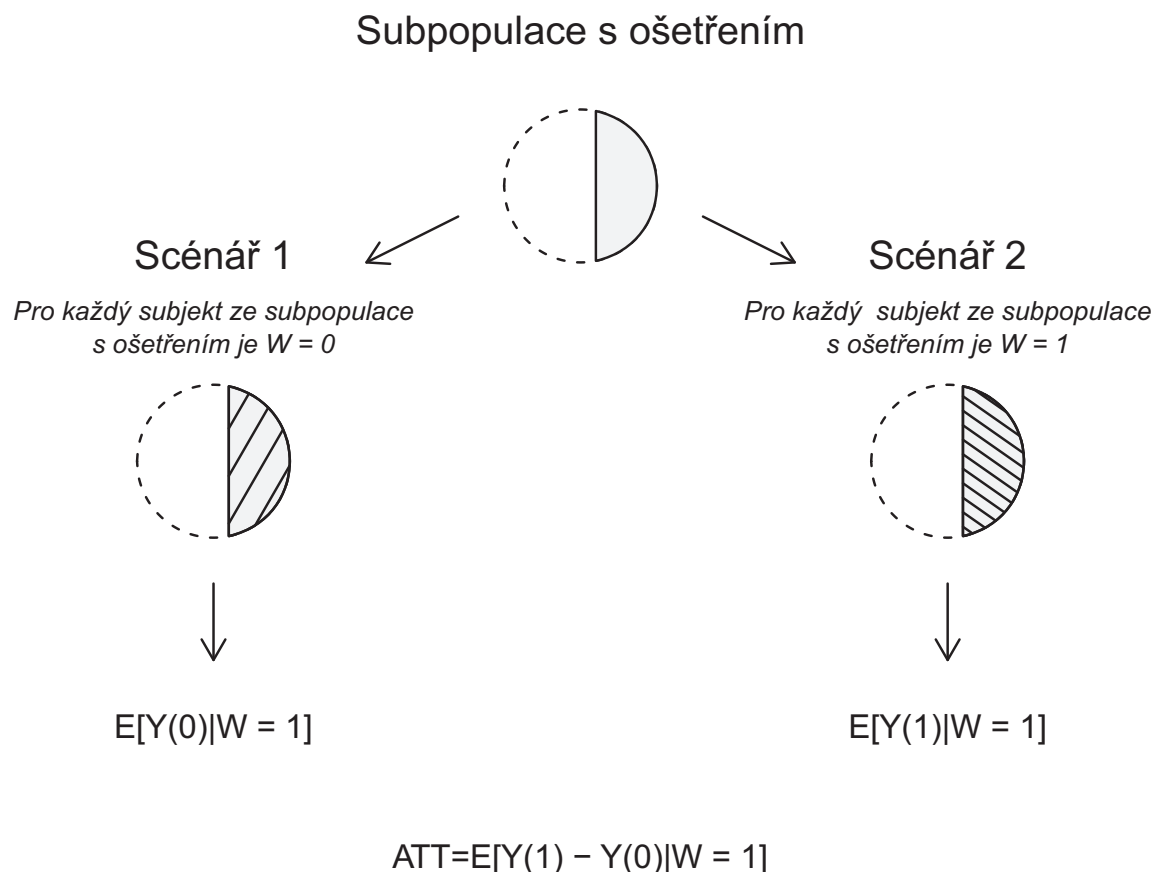
Efekt účasti v akci „Nakup nad“ na hrubé či čisté zisky společnosti ovšem nelze získat prostým porovnáním průměrných zisků vytvořených skupinou uživatelů, kteří se akce zúčastnili, se skupinou uživatelů, jež se akce nezúčastnili. Takové srovnání by bylo zkreslující z důvodu nekontrolovaného vlivu zavádějících proměnných, jež ovlivňují jak rozhodnutí uživatelů účastnit se akce „Nakup nad“, tak velikost jejich útrat a tedy následně zisků společnosti.

Například průměrný čistý zisk společnosti v období uplatnění kreditů je ve skupině účastníků akce  $-22$  Kč, ve skupině těch, kdo se akce neúčastnili, je to  $+259$  Kč (podrobně v tabulce 5.1). Je tedy celý rozdíl  $-281$  Kč v neprospěch skupiny účastníků se akce nevychýleným efektem akce „Nakup nad“ v tomto období? A na jakou populaci by měl být tento rozdíl vztažen? Na populaci všech uživatelů, nebo pouze na populaci těch, kdo se účastnili akce?

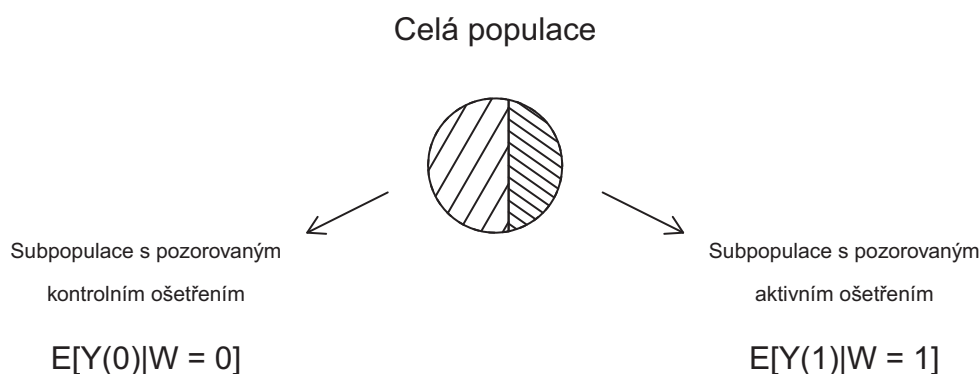
**Tabulka 5.1:** Pozorované hodnoty průměrných zisků (odezev  $Y_1$ ,  $Y_2$  a  $Y_3$ ) na původních datech listopadové akce „Nakup nad“. Podrobná interpretace pozorovaných odezev  $Y_1$ ,  $Y_2$  a  $Y_3$  je v tabulce 5.3.

Pozorované zisky				
	celkový průměr	průměr pro skup. s aktivním ošetřením	průměr pro skup. s kontrolním ošetřením	rozdíl průměrů
$Y_1$	254,254 2	142,095 8	273,340 9	$-131,245 2$
$Y_2$	218,463 9	$-21,933 6$	259,373 8	$-281,307 4$
$Y_3$	288,973 3	417,200 2	267,152 1	150,048 2

Pro posouzení efektivity podpory prodeje prostřednictvím akce „Nakup nad“ je s ohledem na zisky společnosti zajímavá zejména subpopulace těch uživatelů, kteří se akce účastnili, protože pouze u nich vznikají společnosti náklady ve formě získaných a uplatněných kreditů. Zejména je tedy z praktického pohledu zajímavá otázka, jaké by byly průměrné zisky společnosti **na subpopulaci účastníků akce**, kdyby se tito uživatelé akce ve skutečnosti akce neúčastnili (efekt  $\tau_{ATT}$ ). Toto zúžení celé populace na subpopulaci účastníků akce v kontextu efektu akce na zisky společnosti je graficky ilustrováno na obrázku 5.2, zatímco obě pozorované subpopulace a jejich pozorované průměrné zisky jsou graficky ilustrovány na obrázku 5.3.



**Obrázek 5.2:** Subpopulace s ošetřením odpovídá těm uživatelům společnosti  $F$ , kteří se účastnili listopadové akce „Nakup nad“. Scénář 1 reprezentuje tu stejnou subpopulaci, kdyby dostala kontrolní ošetření, tedy kdyby se neúčastnila akce „Nakup nad“. Scénář 2 reprezentuje stále tu stejnou subpopulaci, kdyby dostala aktivní ošetření, tedy kdyby se akce „Nakup nad“ naopak účastnila. Zisky (tedy odezva  $Y$ ) při scénáři 1 reprezentují nepozorovatelné kontrafaktuály pozorovatelných zisků při scénáři 2 na uvažované subpopulaci. Rozdíl průměrných zisků za scénářů 1 a 2 odpovídá efektu akce  $\tau_{ATT}$  na subpopulaci s ošetřením.

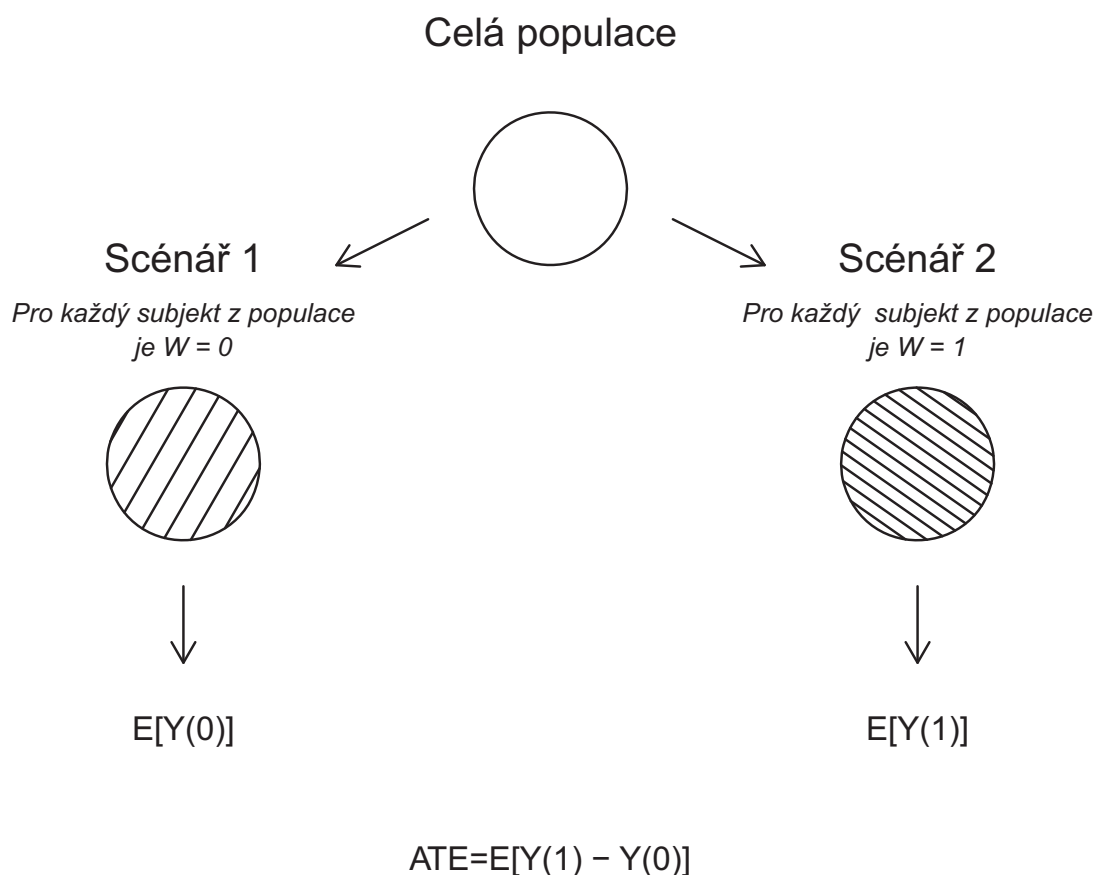


**Obrázek 5.3:** Celá populace všech uživatelů společnosti  $F$  je rozdělená na subpopulaci těch uživatelů, kteří se účastnili listopadové akce „Nakup nad“, a subpopulaci těch, kteří se listopadové akce „Nakup nad“ neúčastnili. Tito uživatelé reprezentují observační datový soubor s pozorovanými zisky v rámci každé subpopulace. Na těchto subpopulacích jsou také počítány pozorované popisné charakteristiky sledovaných zavádějících proměnných, jež jsou uvedeny v tabulce 5.2.

Vynořuje se také spíše akademická otázka, jaký by byl efekt účasti v akci „Nakup nad“ na zisky společnosti, kdyby se akce účastnili všichni uživatelé. V tomto případě by se jednalo o efekt  $\tau_{ATE}$ , tedy efekt akce na celou populaci uživatelů. Vzhledem k tomu, že nelze nějakým umělým zásahem donutit celou populaci uživatelů, aby se účastnila akce, je tento efekt čistě teoretický. Může ale posloužit marketérům společnosti  $F$  jako inspirace, jaké zisky by bylo možné očekávat, kdyby se podařilo zapojit do akce celou populaci uživatelů. Kontext posuzování vlivu akce na celou populaci je graficky ilustrován na obrázku 5.4.

Prakticky velmi využitelná je ovšem související otázka, totiž jaké by byly zisky společnosti, kdyby se akce neúčastnil nikdo. (Takto postavená otázka není ekvivalentní s tím, že by se akce vůbec nekonala, nicméně je významově blízko a pro marketéry může být její zodpovězení užitečné.)

Zisky společnosti  $F$  budou měřeny za různě definovaná období. Jako ztrátové období v rámci subpopulace účastníků akce se na základě popisných statistik v tabulce 5.1 jeví období uplatňování kreditů. V tomto období účastníci akce nakupují méně, jelikož intenzivně nakupovali během předchozích čtyř dnů akce „Nakup nad“ a během tohoto období navíc zisky společnosti snižují uplatňováním slevy formou získaných kreditů.



**Obrázek 5.4:** Scénář 1 reprezentuje celou populaci, kdyby všechny subjekty populace dostaly kontrolní ošetření, tedy kdyby se nikdo neúčastnil akce „Nakup nad“. Scénář 2 reprezentuje celou populaci, kdyby všechny subjekty populace dostaly aktivní ošetření, tedy kdyby se naopak každý účastnil akce „Nakup nad“. Rozdíl průměrných zisků ze scénářů 1 a 2 odpovídá efektu akce  $\tau_{ATE}$  v rámci celé populace všech uživatelů společnosti  $F$ .

Hrubé zisky<sup>26</sup> za období uplatňování kreditů budou označeny  $Y_1$  a čisté zisky za stejné období budou označeny  $Y_2$ . Celkové čisté zisky za období od začátku konání akce až do konce období uplatňování kreditů budou označeny  $Y_3$  a promítají se do nich jak intenzivní nákupy v době konání akce, tak uplatněné kredity, jež zisk v následném období snižují. Tato proměnná je s ohledem na měření efektu akce nejzajímavější.

Celkově tedy lze formulovat následující výzkumné otázky a ty lze rozdělit do dvou skupin podle toho, zda se vztahují k subpopulaci účastníku akce, nebo zda se vztahují k celé populaci uživatelů společnosti  $F$ .

<sup>26</sup> Podrobný popis proměnných  $Y_1$ ,  $Y_2$  a  $Y_3$  je v tabulce 5.3.

### **Výzkumné otázky v rámci subpopulace účastníků akce:**

- VO1.** Jaký by byl průměrný hrubý zisk  $Y_1$  v rámci subpopulace skutečných účastníků akce v období od ukončení akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se tito uživatelé akce na podporu prodeje „Nakup nad“ neúčastnili? Jaký je skutečný efekt účasti v akci „Nakup nad“ na hrubý zisk  $Y_1$  v rámci subpopulace účastníků akce?
- VO2.** Jaký by byl průměrný čistý zisk  $Y_2$  v rámci subpopulace skutečných účastníků akce v období od ukončení akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se tito uživatelé akce na podporu prodeje „Nakup nad“ neúčastnili? Jaký je skutečný efekt účasti v akci „Nakup nad“ na čistý zisk  $Y_2$  v rámci subpopulace účastníků akce?
- VO3.** Jaký by byl průměrný čistý zisk  $Y_3$  v rámci subpopulace skutečných účastníků akce v období od začátku akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se tito uživatelé akce na podporu prodeje „Nakup nad“ neúčastnili? Jaký je skutečný efekt účasti v akci „Nakup nad“ na čistý zisk  $Y_3$  v rámci subpopulace účastníků akce?

Uvedené otázky mají ryze praktický charakter, protože odpovědi na ně umožňují společnosti  $F$  odhadovat skutečný dopad účasti v akci na podporu prodeje „Nakup nad“ (a zprostředkovaně akce samotné) na její zisky.

### **Výzkumné otázky v rámci celé populace uživatelů:**

- VO4.** Jaký by byl průměrný hrubý zisk  $Y_1$  v rámci populace všech uživatelů společnosti  $F$  v období od ukončení akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se nikdo z populace všech uživatelů akce na podporu prodeje „Nakup nad“ neúčastnil? Jaký je skutečný efekt účasti v akci „Nakup nad“ na hrubý zisk  $Y_1$  v rámci populace všech uživatelů společnosti  $F$ ?
- VO5.** Jaký by byl průměrný čistý zisk  $Y_2$  v rámci populace všech uživatelů společnosti  $F$  v období od ukončení akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se nikdo z populace všech uživatelů akce na podporu prodeje „Nakup nad“ neúčastnil? Jaký je skutečný

efekt účasti v akci „Nakup nad“ na čistý zisk  $Y_2$  v rámci populace všech uživatelů společnosti  $F$ ?

**VO6.** Jaký by byl průměrný čistý zisk  $Y_3$  v rámci populace všech uživatelů společnosti  $F$  v období od začátku akce „Nakup nad“ do konce období uplatňování kreditů, kdyby se nikdo z populace všech uživatelů akce na podporu prodeje „Nakup nad“ neúčastnil? Jaký je skutečný efekt účasti v akci „Nakup nad“ na čistý zisk  $Y_3$  v rámci populace všech uživatelů společnosti  $F$ ?

Uvedené otázky, jež se týkají celé populace, mají teoretický charakter, jelikož prakticky není k dispozici „násilný“ mechanismus,<sup>27</sup> který by donutil subpopulaci těch, kteří se akce neúčastnili, aby se jí účastnili a naopak. Prakticky využitelná je odpověď na otázku, jaký by byl průměrný zisk  $Y_i$ , kdyby se akce nikdo neúčastnil, protože je významově blízko otázce, jaké by byly zisky na celé populaci, kdyby se akce nekonala.

Průměrné zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  jsou chápány jako průměrné zisky připadající na uživatele v rámci buď celé populace všech uživatelů, nebo vymezené subpopulace.

Hledání odpovědi na otázku „Jaké by byly zisky, kdyby se akce na podporu prodeje „Nakup nad“ nikdo neúčastnil?“ je analogické odhadování „základních tržeb“ upravenému na kontext jednotlivých výzkumných otázek.

### 5.1.1 Popis datového souboru

Populace uživatelů, ke které se vztahují výzkumné otázky **VO1.–VO6.**, je definovaná jako soubor 102 525 takových uživatelů společnosti  $F$ , kteří jsou aktivní (tedy během října 2017, měsíce předcházejícího listopadové akci „Nakup nad“, byli aspoň jednou na stránkách společnosti  $F$ ) a odebírají newsletter společnosti.

Účast uživatele v akci na podporu prodeje „Nakup nad“ je v roli binárního ošetření  $W \in \{0, 1\}$ , přičemž 1 znamená aktivní ošetření, tedy skutečnost, že uživatel nakoupil po dobu konání akce za alespoň 500 Kč a získal kredit; 0 znamená, že po dobu konání akce nakoupil méně než

<sup>27</sup> S takovýmto mechanismem se ale lze setkat v aplikačních úlohách, kde roli „násilného“ přiřazení ošetření plní např. nějaké státní plošné nařízení či veřejná politika.

za 500 Kč a nezískal kredit. Uživatelů s aktivním ošetřením je 14 910, uživatelů s kontrolním ošetřením je 87 615.

V roli odezvy jsou postupně tři proměnné  $Y_1$ ,  $Y_2$  a  $Y_3$ .  $Y_1$  udává hrubý zisk ze všech nákupů zákazníka (tedy ne utracenou částku, jak bude uvedeno v kapitole 6) během období uplatňování kreditů a  $Y_2$  udává čistý zisk ze všech nákupů zákazníka během období uplatňování kreditů. Odezva  $Y_3$  udává čistý zisk ze všech nákupů zákazníka během období od začátku akce do konce období uplatňování kreditů.

V roli kovariát, jež reprezentují zavádějící proměnné s vlivem jak na ošetření, tak i odezvy, jsou tyto proměnné: objem dřívějších útrat (GS), počet dřívějších nákupů (NPurchases), počet dřívějších návštěv na stránkách společnosti (NVisits), objem dřívějších uplatněných kreditů (CreditsVol), podíl dřívějších otevřených e-mailů (OpenRate), získání výhod věrnostního programu (Platinum) a velikost bydliště (City2). U všech těchto proměnných lze pozorovat vyšší hodnoty (potenciálně asociované s vyššími zisky společnosti) ve skupině účastníků akce než ve skupině těch, kdo nákupy nad 500 Kč neuskutečnili, jak je vidět z tabulky 5.2.

**Tabulka 5.2:** Průměr a směrodatná odchylka pro každou z uvažovaných kovariát – zvláště pro pozorované uživatele, kteří se akce neúčastnili (skupina 0), a pro pozorované uživatele, kteří se akce účastnili (skupina 1). Dále tabulka uvádí standardizované rozdíly *SMD* (podrobně v 1.2 a 5.2.1), přičemž čím větší je absolutní hodnota *SMD*, tím více jsou skupiny nevyvážené.

	0	1	SMD
<i>n</i>	87 615	14 910	
GS (mean (sd))	9 691,71 (1 8454,16)	17 070,47 (31 547,31)	0,286
NPurchases (mean (sd))	15,65 (30,77)	28,02 (54,93)	0,278
NVisits (mean (sd))	143,96 (220,07)	216,61 (297,27)	0,278
CreditsVol (mean (sd))	348,51 (1 226,13)	979,32 (4 634,41)	0,186
OpenRate (mean (sd))	0,17 (0,27)	0,21 (0,29)	0,153
Platinum (mean (sd))	0,16 (0,36)	0,28 (0,45)	0,299
City2 (mean (sd))	0,68 (0,47)	0,68 (0,47)	0,001

Tato skutečnost naznačuje, že uvedené proměnné mohou zkreslovat odhad efektu akce na zisky, pokud by byl efekt odhadován pouze porovnáním zisků mezi oběma skupinami (subpopulacemi) uživatelů.



Průměry a směrodatné odchylky pro jednotlivé kovariáty jsou spočteny na pozorovaných původních subpopulacích, jež jsou symbolicky reprezentovány obrázkem 5.3. V tabulce 5.2 je zaznamenán kromě průměrů a směrodatných odchylek pro jednotlivé kovariáty také standardizovaný rozdíl mezi skupinovými průměry pro jednotlivé kovariáty (*SMD*) a jeho vzorec je uveden v 5.2.1.

Přehledný popis všech potřebných proměnných obsahuje tabulka 5.3.

Vzhledem k tomu, že veškeré následné analýzy odhadují efekt akce na celé populaci (výběrová data jsou tedy totožná s celou populací), není potřeba uvádět směrodatnou chybu odhadů. Pokud bychom ale pozorovanou populaci brali jako jen jednu z mnoha možností realizace superpopulace, potom znalost směrodatných chyb odhadů umožňuje i zobecnění na superpopulaci. Navíc např. v případě optimálního párování je v níže uváděné směrodatné chybě podle [Abadie a Imbens \(2006\)](#) zohledněna i náhodnost samotného párovacího algoritmu. V případě optimálního párování na základě odhadnutých propenzitních skóru ovšem variabilita samotných odhadů propenzitních skóru podchycena není. V následujících výstupech analýz jsou tedy uváděny i směrodatné chyby odhadů efektů.

### 5.1.2 Základní předpoklady identifikace efektu ošetření

V pozorovaném datovém souboru je u každého uživatele u sledované odezvy vždy pouze jedna pozorovaná potenciální odezva (buď zisk při aktivním ošetření, tedy při účasti v akci „Nakup nad“, nebo zisk při kontrolním ošetření, tedy zisk bez účasti v akci „Nakup nad“). Pro identifikaci kauzálního efektu účasti v akci „Nakup nad“ na zisk je ale potřeba mít u každého uživatele obě potenciální odezvy. Tedy pokud pozorujeme zisk u *i*-tého účastníka, jenž se akce účastnil, je potřeba vědět, jaký by byl zisk u téhož *i*-tého účastníka, kdyby se akce neúčastnil; jinými slovy je potřeba odhadnout jeho kontrafaktuál.

Aby tedy mohl být efekt akce „Nakup nad“ na zisky **na základě pozorovaných dat** korektně identifikován, je potřeba ověřit splnění následujících podmínek identifikace efektu. (V této kapitole je kauzální efekt odhadován za předpokladu regulárního mechanismu přiřazování ošetření, tedy podmínky se vztahují i k zavádějícím proměnným.)

**Tabulka 5.3:** Přehled proměnných v analýzách provedených za účelem vyhodnocení efektivity akce na podporu prodeje „Nakup nad“.

Název proměnné	Popis proměnné
<b>Ošetření:</b>	
$W$	Binární proměnná nabývající hodnot 0, 1; 1 reprezentuje aktivní ošetření a znamená, že zákazník po dobu konání akce „Nakup nad“ uskutečnil nákup za aspoň 500 Kč a získal kredit; 0 reprezentuje kontrolní ošetření a znamená, že zákazník uskutečnil nákup za méně než 500 Kč a nezískal kredit; stručně bude v textu účast/neúčast na akci „Nakup nad“
<b>Seznam odezev:</b>	
$Y_1$	Hrubý celkový zisk (marže) připadající na uživatele v období uplatňování kreditů (od ukončení akce do konce období pro uplatnění kreditů): 20. 11. 2017–30. 11. 2017 včetně
$Y_2$	Čistý zisk ( $Y_1$ po odečtení všech kreditů) na uživatele v období uplatňování kreditů (od ukončení akce do konce období pro uplatnění kreditů): 20. 11. 2017–30. 11. 2017 včetně
$Y_3$	Čistý zisk na uživatele v období od začátku akce do konce období pro uplatnění kreditů: 16. 11. 2017–30. 11. 2017 včetně
<b>Seznam zavádějících kovariát:</b>	
GS	Spojité proměnná nabývající nezáporných hodnot; udává objem útrat v korunách za celou historii zákazníka až do začátku konání akce „Nakup nad“
NPurchases	Celočíselná proměnná nabývající nezáporných hodnot; udává počet nákupů za celou historii uživatele až do začátku konání akce „Nakup nad“
NVisits	Celočíselná proměnná nabývající nezáporných hodnot; udává počet dnů návštěv za celou historii uživatele až do začátku konání akce „Nakup nad“
CreditsVol	Spojité proměnná nabývající nezáporných hodnot; udává celkový objem uplatněných kreditů v korunách za celou historii uživatele až do začátku konání akce „Nakup nad“
OpenRate	Spojité proměnná z intervalu (0, 1); udává podíl otevřených e-mailů za celou historii uživatele až do začátku konání akce „Nakup nad“; pokud uživatel žádný newsletter nikdy neodebíral, má hodnotu 0
Platinum	Binární proměnná nabývající hodnot 0, 1; 1 znamená, že zákazník získal výhody věrnostního programu
City2	Binární proměnná nabývající hodnot 0, 1; 1 znamená, že zákazník udává adresu ve větším městě (tedy s bohatou nabídkou služeb a zboží nabízeného společností $F$ ), 0 znamená menší sídlo (s chudší nabídkou)

## 1. Předpoklad nezavádějícího mechanismu přiřazování

Tento velmi důležitý předpoklad má zajistit, že při kontrole všech relevantních zavádějících proměnných je již mechanismus přiřazování ošetření (volba uživatele, zda nakoupí za aspoň 500 Kč, či méně) „jakoby“ náhodný. Bohužel tento předpoklad nelze statisticky testovat a je potřeba jej alespoň nepřímo obhájit na základě znalosti kontextu problému.

Na začátku seznamu zavádějících proměnných v tabulce 5.3 jsou  $GS$  a  $NPurchases$ . Obě proměnné nesou informaci o nákupní historii uživatele a popisují jeho tendenci utrácet (obě proměnné spolu souvisejí, ale stejného objemu útrat lze dosáhnout větší četností malých nákupů či menší četností větších nákupů). Lze usoudit, že ti uživatelé, u nichž objem útrat či četnost nákupů dosahují velkých hodnot, budou mít i v budoucnu tendenci hodně nakupovat a tím i snadněji dosáhnou na splnění podmínky nákupu nad 500 Kč. Ovlivňují tedy hodnotu ošetření  $W$ . Zároveň ale lze očekávat, že ve sledovaném období (buď od začátku akce do konce období uplatňování kuponů, nebo během období uplatňování kuponů) budou generovat společnosti větší zisk, tedy ovlivňují odezvu. Tento fenomén v kontextu podobného business modelu jako u společnosti  $F$  popsali např. [Kimes a Dolakia \(2011\)](#) ve své studii chování spotřebitele. Proměnné  $GS$  a  $NPurchases$  tedy splňují definici zavádějících proměnných.

Další proměnnou ze seznamu zavádějících proměnných je  $NVisits$ , tedy počet dno-návštěv (opakované návštěvy během jednoho dne se počítají jako jedna návštěva). Časté návštěvy na stránkách společnosti  $F$  naznačují uživatelovo vyhledávání konkrétních produktů a služeb z nabídky společnosti či průběžné sledování nabídky a slev, motivované např. newsletterem. Obojí povede k nákupu u společnosti  $F$  spíše, než by k němu došlo v případě zákazníka, který stránky společnosti často nenavštěvuje. Tedy uživatel s častějšími návštěvami spíše splní podmínku akce (vliv na  $W$ ), ale také jeho nákupy mohou vést k většímu zisku společnosti než v případě zákazníka, který stránky často nenavštěvuje (vliv na  $Y$ ). Jedná se tedy o zavádějící proměnnou.

Proměnná *CreditsVol* udává počet dosud uplatněných kreditů z různých dřívějších akcí na podporu prodeje, které společnost  $F$  pořádá. Vyšší hodnoty *CreditsVol* tedy nepřímo udávají vyšší míru zapojení do jiných dřívějších akcí (a splnění podmínek akcí pro získání kreditů). Proto lze očekávat, že uživatel s vyšší hodnotou proměnné *CreditsVol* spíše zareaguje na akci „Nakup nad“ než uživatel s nižší hodnotou této proměnné (vliv na  $W$ ). Zákazníci, u kterých pozorujeme hodně uplatněných kreditů, jsou ti, kteří nutně hodně nakupují. Jednak kredity získávají většinou za nákupy (s výjimkou například narozeninových kreditů), jednak následně, aby je mohli uplatnit, musejí uskutečnit další nákupy, tedy generují společnosti větší zisk než uživatelé s menším počtem dosud uplatněných kreditů (vliv na  $Y$ ). Jedná se tedy opět o zavádějící proměnnou.

Proměnná *OpenRate* udává dosavadní podíl otevřených emailů, tedy vypovídá o tom, do jaké míry je uživatel vystaven reklamě na produkty a služby společnosti v emailech a odkazovaných newsletterech. Také částečně vypovídá i o míře loajality zákazníka ke společnosti  $F$ , bez ní by totiž častěji mailly mazal bez otevření. U takovéhoho zákazníka lze spíše očekávat splnění podmínky akce „Nakup nad“ než u zákazníka, jenž mailly otevírá málo nebo vůbec (vliv na  $W$ ). Od zákazníků vystavených intenzivnější reklamě v součinnosti s možnou vyšší loajalitou vůči společnosti  $F$  lze očekávat i větší zisky (vliv na  $Y$ ). Proměnná *OpenRate* je tedy zavádějící proměnná.

Binární stavová proměnná *Platinum* vyjadřuje svojí jedničkovou hodnotou vysokou míru loajality zákazníka ke společnosti  $F$  – jedná se o věrného zákazníka. Věrný zákazník je ten, který často a hodně nakupuje, a tedy lze předpokládat i jeho vyšší ochotu zapojit se do akce „Nakup nad“ než u zákazníka s nulovou hodnotou této proměnné. Zároveň lze u tohoto zákazníka očekávat i větší zisky, jak je v kontextu podobného business modelu jako u společnosti  $F$  popsáno na webových stránkách Forbes v příspěvku Wang (2011) nebo v dizertační práci Silva (2012) (vliv na  $Y$ ). Jedná se tedy opět o zavádějící proměnnou.

Poslední proměnná *City2* je opět stavová binární a její jedničková hodnota udává, že registrovaná adresa uživatele je větší město. To

v kontextu záznamů dat společnosti  $F$  znamená, že v tomto větším městě, kde se uživatel pravděpodobně často vyskytuje (uživatel může udávat i jinou adresu, než kde tráví většinu svého času), je bohatá nabídka zboží a služeb. Tedy na podnět utratit aspoň 500 Kč může snadněji reagovat proto, že může pohodlně vybírat z bohaté nabídky (vliv na  $W$ ). Zároveň se ve větších městech koncentrují vzdělanější lidé s vyššími příjmy, tedy uživatelé z větších měst mohou více utrácet (vliv na  $Y$ ). Proměnná  $City2$  má tedy také charakter zavádějící proměnné.

U všech sedmi uvedených proměnných lze očekávat vliv jak na odezvu, tak na ošetření, vykazují tedy charakter zavádějících proměnných. Jejich vliv je tudíž potřeba kontrolovat, aby nedošlo k vychýlení odhadu efektu akce na podporu prodeje „Nakup nad“ na zisky. Nedá se ovšem dokázat, že výčet zavádějících proměnných je úplný.

## 2. Předpoklad SUTVA

Tento předpoklad požaduje, aby potenciální odezvy každého uvažovaného subjektu nebyly ovlivněny ošetřením jakéhokoliv jiného subjektu. V kontextu akce „Nakup nad“ by tento předpoklad mohl být porušen, pokud by kategorie produktů, z níž by chtěl uživatel nakupovat v době od začátku akce do konce měsíce, byla již vyčerpána dřívějšími účastníky akce. Potenciální zisk (odezvy  $Y_1, Y_2, Y_3$ ) na uvažovaného uživatele by tedy byl ovlivněn účastí v akci „Nakup nad“ (ošetření  $W$ ) jiných uživatelů. Dle zkušeností společnosti  $F$  k tomu nedochází, jelikož nabídka produktů je natolik bohatá, že její vyčerpání během sledovaného období je téměř nemožné.

SUTVA dále předpokládá, že neexistuje více verzí pro jednotlivé úrovně ošetření. Toto je zaručeno charakterem akce – buď účastník splní podmínku nákupu za aspoň 500 Kč a kredit získá, nebo ne. Předpoklad stabilní hodnoty ošetření subjektu (SUTVA) je tedy splněn.

## 3. Exogenita zavádějících proměnných

Zavádějící proměnné nesmějí být ovlivněny ošetřením. Pro splnění tohoto předpokladu stačí, aby hodnota každé kovariáty u každého

subjektu  $i$  byla stejná bez ohledu na to, zda je  $i$ -tému subjektu přiřazeno aktivní, či kontrolní ošetření (tedy  $X(w) = X(1 - w)$  pro  $w = 0, 1$ ).

Vzhledem k tomu, že všechny zavádějící kovariáty uvedené v tabulce 5.3 mají buď charakter proměnných před ošetřením (např. *NPurchases*), nebo mají trvalý charakter (např. proměnná *City2*), jejich hodnota nemůže být úrovní ošetření (účasti/neúčasti na akci „Nakup nad“) ovlivněna. Tento předpoklad je tedy splněn.

#### 4. Překryv

Překryv pro uvažovanou kovariátu reprezentuje tu oblast hodnot uvažované kovariáty, pro kterou je pravděpodobnost přiřazení jak aktivního, tak kontrolního ošetření ostře kladná.<sup>28</sup> Vzhledem k praktickému charakteru výzkumných otázek VO1 až VO3 by bylo interpretačně výhodné, kdyby tato oblast pokrývala celý pozorovaný rozsah hodnot všech sledovaných zavádějících kovariát, protože pak by bylo zřejmé, k jaké populaci vztáhnout výsledky analýz.

Předpoklad dostatečného překryvu bude v následujících podkapitolách posuzován pro každou použitou metodu zvlášť.

## 5.2 Párování na základě Mahalanobisovy metriky

V této i následující podkapitole 5.3 budou použity takové metody, které umožní odpovědět na výzkumné otázky VO1 až VO3, jež zjišťují, jaké by byly zisky  $Y_1$ ,  $Y_2$  a  $Y_3$ , kdyby se ti uživatelé, kteří se účastnili akce, hypoteticky akce neúčastnili. Tedy ke každému uživateli s pozorovaným aktivním ošetřením je potřeba najít jeho kontrafaktuál (z principu nepozorovatelný). Následný odhad efektu akce na zisky je tedy vztažen na subpopulaci s aktivním ošetřením, a jedná se proto o odhady typu  $\tau_{ATE}$  (podkapitola 2.4 a obr. 5.2).

Pro tyto odhady je výhodné, když je subjektů s kontrolním ošetřením podstatně více než subjektů s aktivním ošetřením. Toto je i případ posuzovaného datového souboru k akci „Nakup nad“, jelikož uživatelů, kteří se

---

<sup>28</sup> Přesněji – kladná by měla být pro vícerozměrný vektor kovariát.

účastnili akce, je pouze 14 910, zatímco uživatelů, kteří se akce neúčastnili, je 87 615.

Párování na základě Mahalanobisovy metriky hledá k pozorovanému subjektu ze skupiny s aktivním ošetřením jeho „nejpodobnější dvojče“ (s ohledem na zavádějící kovariáty) ze skupiny s kontrolním ošetřením, a to slouží jako náhrada nepozorovatelného kontrafaktuálu. Tato „podobnost“ je reprezentována vzdáleností na základě Mahalanobisovy metriky (3.5.3).

Byly použity dva párovací algoritmy, párování „jeden na jednoho“ metodou nejbližšího souseda a optimální párování. Při párování pomocí metody nejbližšího souseda záleží na pořadí, v jakém jsou subjekty s pozorovaným aktivním ošetřením v datovém souboru seřazeny, v případě optimálního párování pořadí subjektů v datovém souboru vliv na výsledek párování nemá.<sup>29</sup>

Pro nalezení párů a následně odhadů efektů byly použity balíky `Matching` a `MatchIt` jazyka `R`.

Po provedení párování jednotlivými algoritmy je potřeba posoudit kvalitu použitých párovacích algoritmů co do vyváženosti a zároveň vyhodnotit s tím související překryv. Pokud lze napárované skupiny považovat vzhledem ke všem kovariátám za vyvážené, pak se skupiny liší pouze ošetřením a rozdíly v odezvě mezi skupinami lze přisuzovat již výlučně vlivu ošetření (tedy situace připomíná randomizovaný experiment).

### 5.2.1 Vyváženost párování

Pro posouzení vyváženosti vzhledem ke všem zavádějícím kovariátám je obvyklé uvádět především průměr pro pozorovanou skupinu s aktivním ošetřením a průměr pro skupinu napárovaných „dvojčat“ ze skupiny s kontrolním ošetřením.

---

<sup>29</sup> Při párování pomocí metody nejbližšího souseda prochází párovací algoritmus postupně skupinou subjektů s aktivním ošetřením a ke každému najde nejbližší možný, ovšem dosud nepoužitý subjekt ze skupiny s kontrolním ošetřením. Takovýto napárovaný subjekt již nelze opakovaně použít. Tak se může stát, že k jednotlivému subjektu s aktivním ošetřením je sice nalezen vždy ten nejlepší zbylý kontrolní subjekt, ovšem v průměru tento způsob nemusí vést k největší možné podobnosti párů. Optimální algoritmy naopak hledají ze všech možných kombinací dvojic tu sadu párů, která vede v průměru k největší podobnosti párů. Toto je ovšem výpočetně zejména na velkých datových souborech náročné.

Pokud jsou napárované skupiny s ohledem na všechny kovariáty vyvážené, měly by mít stejné průměry. Posoudit tuto skutečnost na základě  $p$ -hodnot např.  $t$ -testu ovšem není vzhledem k velikosti souborů vhodné, jelikož signifikance rozdílu s velikostí souboru roste. Proto se k posouzení vyváženosti používají standardizované rozdíly mezi průměry  $SMD$  definované vztahem:<sup>30</sup>

$$SMD = \frac{M_1 - M_0}{\sqrt{\frac{S_1^2 + S_0^2}{2}}}, \quad (5.2.1)$$

kde  $M_1$ ,  $M_0$  jsou po řadě výběrové průměry skupiny s aktivním a skupiny s kontrolním ošetřením a  $S_1^2$ ,  $S_0^2$  jsou výběrové rozptyly obou skupin. Je-li absolutní hodnota  $|SMD| \in \langle 0; 0,1 \rangle$ , pak se napárované skupiny považují za vyvážené, pro  $|SMD| \in \langle 0,1; 0,2 \rangle$  lze vyváženost akceptovat, jinak nelze mluvit o vyváženém párování.

Vyváženost napárovaných skupin pro jednotlivé párovací algoritmy porovnáváme s vyvážeností nenapárovaných původních pozorovaných dat, a tu posoudíme na základě statistik v již uvedené tabulce 5.2.

V nenapárovaném souboru překračuje  $SMD$  u všech zavádějících proměnných (s výjimkou *City2*) hodnotu 0,1. V případě proměnných *GS*, *NPurchase*, *NVisits* a *Platinum* se  $SMD$  dokonce blíží hodnotě 0,3. Původní datový soubor je tedy silně nevyvážený a bylo by chybou interpretovat rozdíl mezi odezvami (zisky společnosti  $F$ ) ve skupinách s kontrolním a aktivním ošetřením jako výlučný vliv ošetření (účasti na akci „Nakup nad“).

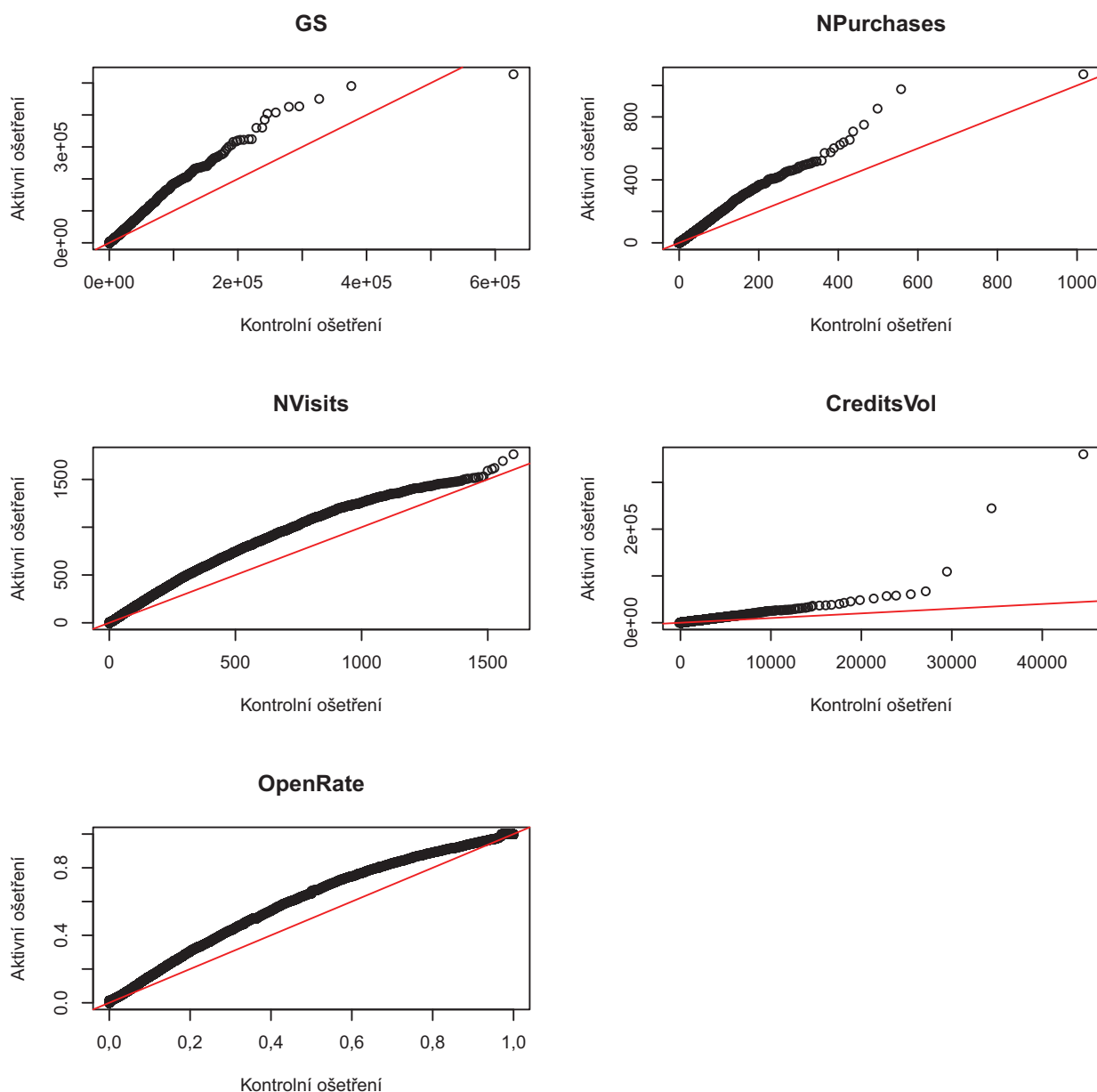
Ačkoliv je obvyklé posuzovat vyváženost na základě  $SMD$ , zejména u spojitých kovariát je užitečné posoudit shodu rozložení mezi skupinami také graficky (dvě skupiny se stejným  $SMD$  nutně nemusejí mít stejné rozložení), například pomocí kvantil-kvantilového grafu ke každé zavádějící proměnné.

U binárních proměnných lze průměry interpretovat jako podíl úspěchů (jednotkových hodnot) ve skupině s kontrolním či aktivním ošetřením, pro tyto proměnné tedy informace o průměrech postačuje k představě o jejich rozložení. Graficky lze tuto představu doplnit například sloupcovými grafy.

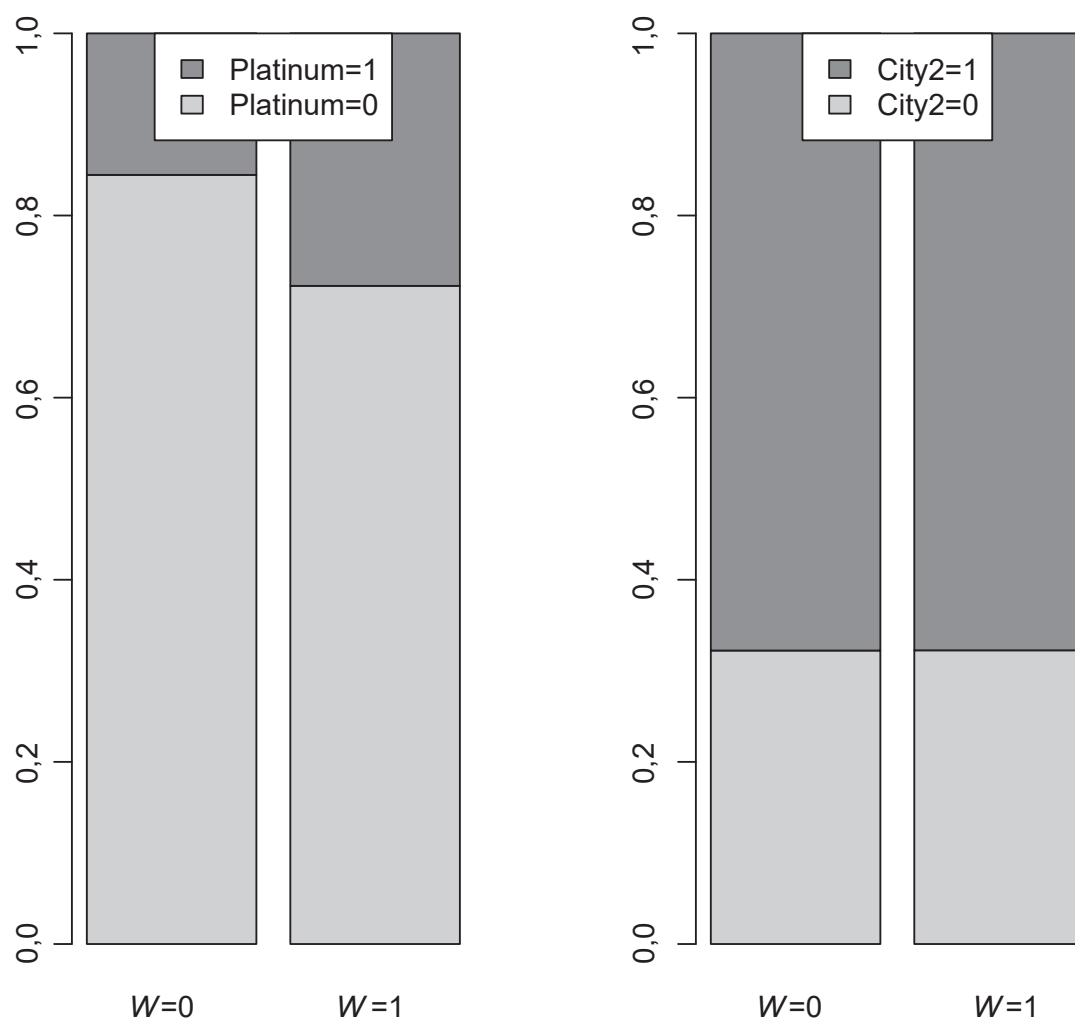
<sup>30</sup> Vážený průměr výběrových rozptylů ve jmenovateli  $SMD$  zavedli [Rosenbaum a Rubin \(1985\)](#).



Nevyváženost původních pozorovaných dat je zřejmá z obrázku 5.5, kde jsou kvantil-kvantilové grafy pro proměnné *GS*, *NPurchases*, *NVisits*, *CreditsVol* a *OpenRate*, a z obrázku 5.6, kde jsou sloupcové grafy pro proměnné *Platinum* a *City2*.



**Obrázek 5.5:** Kvantil-kvantilové grafy pro původní nenapárovaný datový soubor: graf vynáší pro každou kovariátu na vodorovnou osu kvantily rozložení pro skupinu s pozorovaným kontrolním ošetřením a na svislou osu kvantily rozložení pro skupinu s pozorovaným aktivním ošetřením. Čím více se body grafu vzdalují od proložené červené přímky, tím více je vyváženost porušena. U všech uvedených zavádějících proměnných je na původních observačních datech vidět silná nevyváženost.



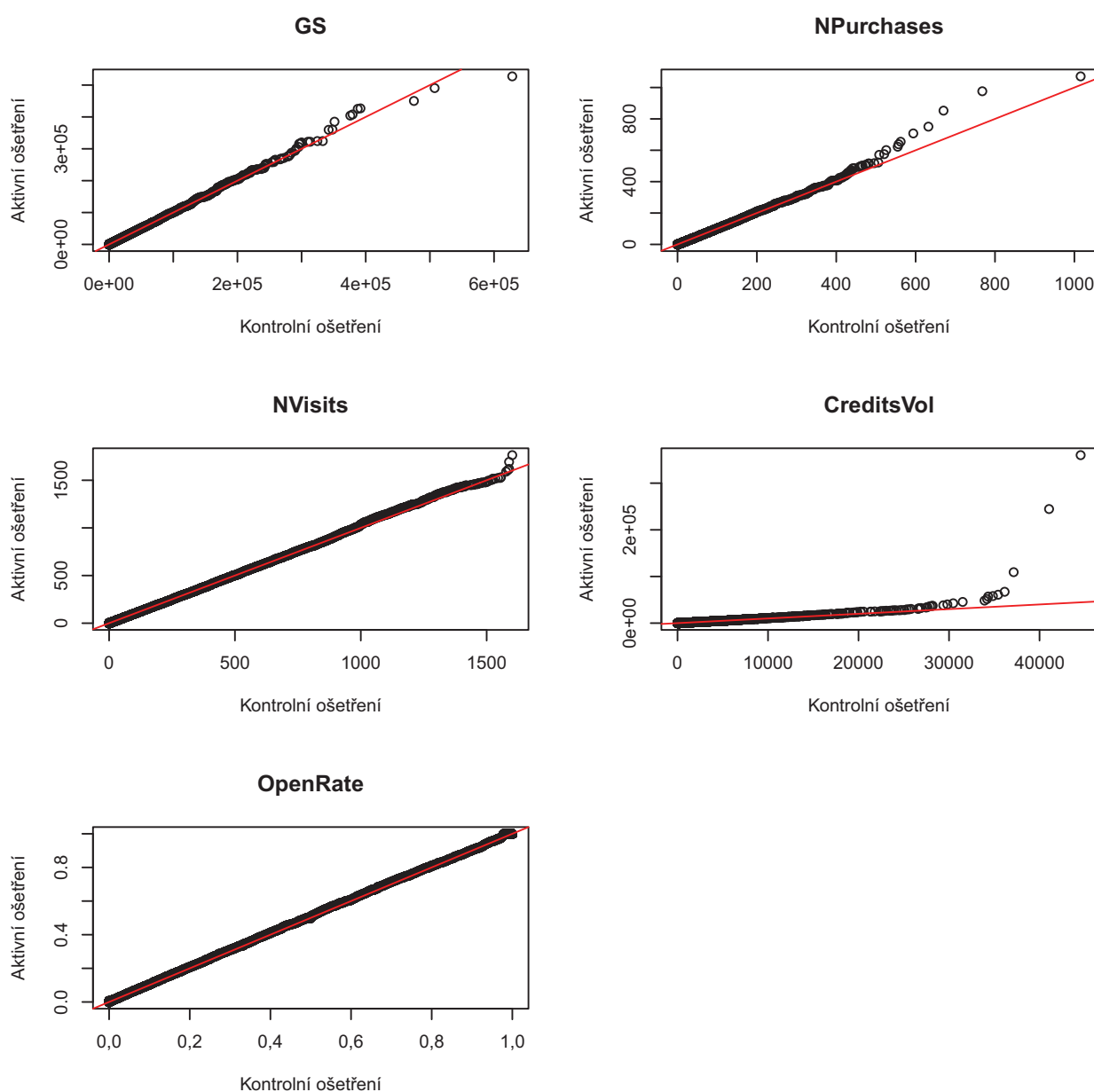
**Obrázek 5.6:** Sloupcové grafy pro původní nenapárovaný datový soubor: graf pro každou kovariátu zobrazuje výši sloupce relativní četnosti pro skupinu s kontrolním ošetřením ( $W = 0$ ) a pro skupinu s aktivním ošetřením ( $W = 1$ ). U proměnné *Platinum* je opět vidět silná nevyváženost. Malý rozdíl v relativních četnostech proměnné *City2* (dle tabulky 5.2 je  $SMD > 0$ ) je graficky nezaznamatelný

### Soubor napárovaný metodou nejbližšího souseda

Výsledkem párování pomocí metody nejbližšího souseda je soubor, v němž bylo nalezeno ke **každému** subjektu z původní pozorované skupiny s aktivním ošetřením právě jedno „dvojče“ bez opakování ze skupiny s kontrolním ošetřením. V případě shod (více vhodných dvojčat do páru k uvažovanému ošetřenému subjektu) bylo právě jedno „dvojče“ vybráno náhodně. Tím je triviálně splněn předpoklad překryvu. Párování proběhlo výlučně na základě zavádějících kovariát, odezvy  $Y_1$ ,  $Y_2$  a  $Y_3$  tedy výsledek párovací procedury nikterak neovlivnily.

Vyváženost kovariát mezi skupinou s aktivním ošetřením a k ní napárouvanou skupinou s kontrolním ošetřením (v roli kontrafaktuálů) lze posoudit na základě *SMD* hodnot v tabulce 5.4 nebo vizuálně na základě grafů 5.7 pro spojité kovariáty či 5.8 pro binární kovariáty.

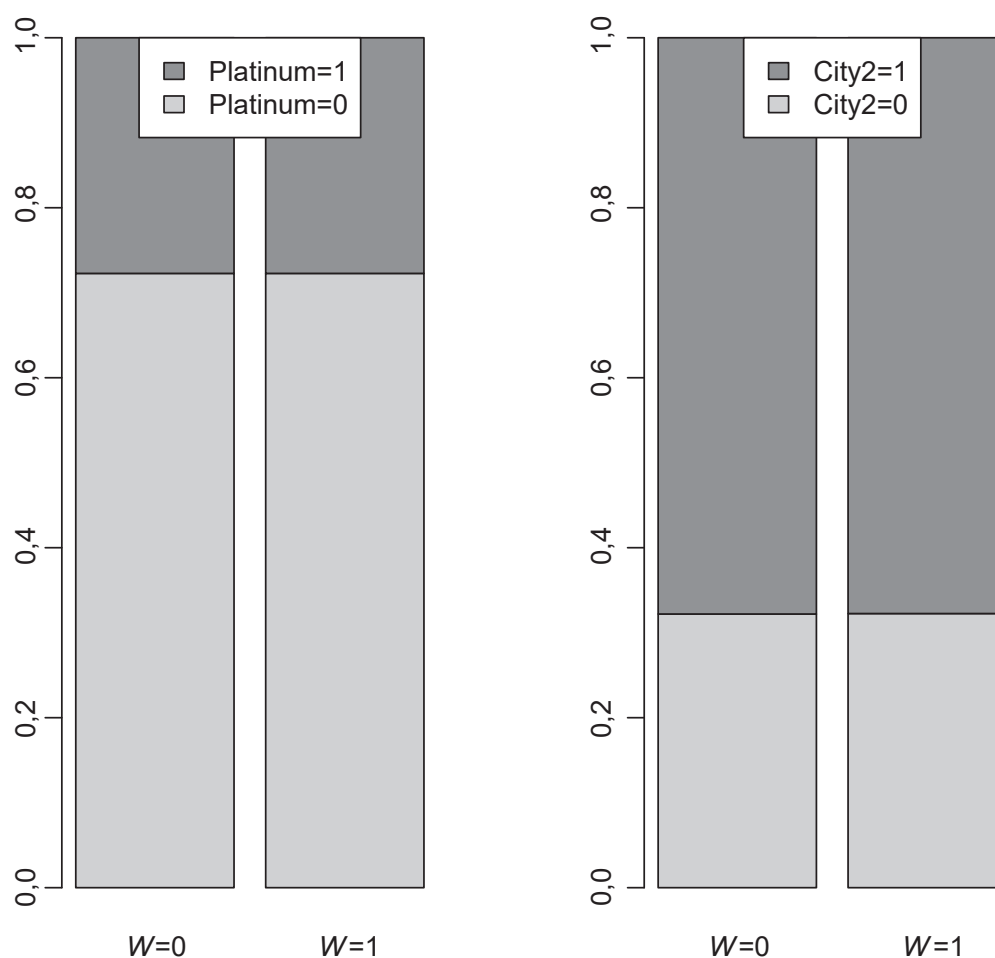
Z tabulky 5.4 je vidět, že výsledek párování pomocí metody nejbližšího souseda je velmi uspokojivý. Všechny kovariáty lze na základě standardizovaných rozdílů průměrů považovat za vyvážené, jelikož hodnoty *SMD*  $\ll 0,1$



**Obrázek 5.7:** Kvantil-kvantilové grafy pro soubor napárouvaný pomocí metody nejbližšího souseda: vyváženost se v porovnání s původním datovým souborem výrazně zlepšila, neuspokojivě napárouvané jsou pouze subjekty s velkými hodnotami proměnných *NPurchases* a *CreditsVol*.

**Tabulka 5.4:** Číselné charakteristiky kovariát v souboru napárovaném metodou nejbližšího souseda. Ve sloupci označeném hodnotou „1“ je všech 14 910 subjektů s pozorovaným ošetřením. Ve sloupci označeném hodnotou „0“ je 14 910 napárovaných kontrafaktuálů, kteří zastupují pozorované subjekty s aktivním ošetřením, kdyby dostaly kontrolní ošetření. Rozdíl v rozložení všech kovariát mezi oběma napárovanými skupinami je posouzen na základě hodnot *SMD*.

	0	1	SMD
	<i>n</i> 14 910	14 910	
GS (mean (sd))	16 845,32 (31 017,66)	17 070,47 (31 547,31)	0,007 20
NPurchases (mean (sd))	27,79 (52,81)	28,02 (54,93)	0,004 30
NVisits (mean (sd))	214,36 (291,42)	216,61 (297,27)	0,007 64
CreditsVol (mean (sd))	861,71 (2 462,65)	979,32 (4 634,41)	0,031 69
OpenRate (mean (sd))	0,21 (0,29)	0,21 (0,29)	0,013 21
Platinum (mean (sd))	0,28 (0,45)	0,28 (0,45)	< 0,000 01
City2 (mean (sd))	0,68 (0,47)	0,68 (0,47)	0,001 15



**Obrázek 5.8:** Sloupcové grafy pro soubor napárovaný pomocí metody nejbližšího souseda: obě binární proměnné jsou vyvážené.

pro každou uvažovanou kovariátu. Také vizuální posouzení vyváženosti pomocí kvantil-kvantilových grafů na obrázku 5.7 či sloupcových grafů na obrázku 5.8 ukazuje významné zlepšení proti obdobným grafům před párováním (5.5 a 5.6).

### Soubor napárovaný optimální metodou párování

Algoritmus optimálního párování byl nastaven tak, že se při hledání kontrolních subjektů vhodných pro párování k subjektům ze skupiny s aktivním ošetřením mohly subjekty v obou skupinách opakovat. Přitom byla při opakování párům přiřazena taková váha, aby celkový převážený počet párů zůstal stejný jako počet subjektů v původní pozorované skupině s aktivním ošetřením a žádný subjekt ze skupiny s aktivním ošetřením nebyl z párování vyloučen.

Dále se v případě shod do páru k uvažovanému subjektu s aktivním ošetřením využily všechny vhodné subjekty ze skupiny s kontrolním ošetřením, ale převáženě – s váhami reflektujícími jejich četnost (postup je tedy zcela deterministický). Celkově je součet vah nově vytvořených párů roven 14 910 (odpovídá počtu původně pozorovaných subjektů s aktivním ošetřením), přestože v tabulce 5.5 jsou četnosti subjektů v napárovaných skupinách větší.

**Tabulka 5.5:** Číselné charakteristiky kovariát v souboru napárovaném pomocí metody optimálního párování. V obou sloupcích pro aktivní i kontrolní ošetření je 913 963 subjektů, ovšem každému z nich je přiřazena taková váha, aby po převážení byl počet párů stejný jako počet pozorovaných subjektů s aktivním ošetřením, tedy 14 910. Uvedené průměry i směrodatné odchylky jsou počítány již na váženém souboru. Z nich jsou také spočteny hodnoty *SMD*.

	0		1		SMD
	<i>n</i> 913 963		913 963		
GS (mean(sd))	16 937,04	(31 478,11)	17 070,47	(31 546,25)	0,004 2
NPurchases (mean(sd))	27,98	(53,8)	28,02	(54,93)	0,000 9
NVisits (mean(sd))	215,4	(293,73)	216,61	(297,26)	0,004 1
CreditsVol (mean(sd))	897,89	(2 720,79)	979,32	(4 634,25)	0,021 4
OpenRate (mean(sd))	0,21	(0,29)	0,21	(0,29)	0,009 2
Platinum (mean(sd))	0,28	(0,45)	0,28	(0,45)	0,000 0
City2 (mean(sd))	0,68	(0,47)	0,68	(0,47)	0,002 6

Výsledný napárovaný soubor by byl stejný i při jiném pořadí subjektů v původním datovém souboru a s ohledem na požadavky nastavené v odstavcích výše se jedná o „nejlepší možný“ napárovaný soubor. Vyuváženost kovariát mezi skupinou s aktivním ošetřením a k ní napárovanou skupinou s kontrolním ošetřením (v roli kontrafaktuálů) byla opět posouzena na základě  $SMD$  hodnot v tabulce 5.5, přičemž potřebné průměry a rozptyly byly vážené.

Z tabulky 5.5 je vidět, že výsledek párování pomocí optimálního párování je také velmi uspokojivý. Všechny kovariáty lze na základě standardizovaných rozdílů průměrů považovat za vyvážené, jelikož hodnoty  $SMD \ll 0,1$  pro každou uvažovanou kovariátu. Ve srovnání se standardizovanými rozdíly u metody nejbližšího souseda (5.4) jsou některé hodnoty  $|SMD|$  ještě menší.

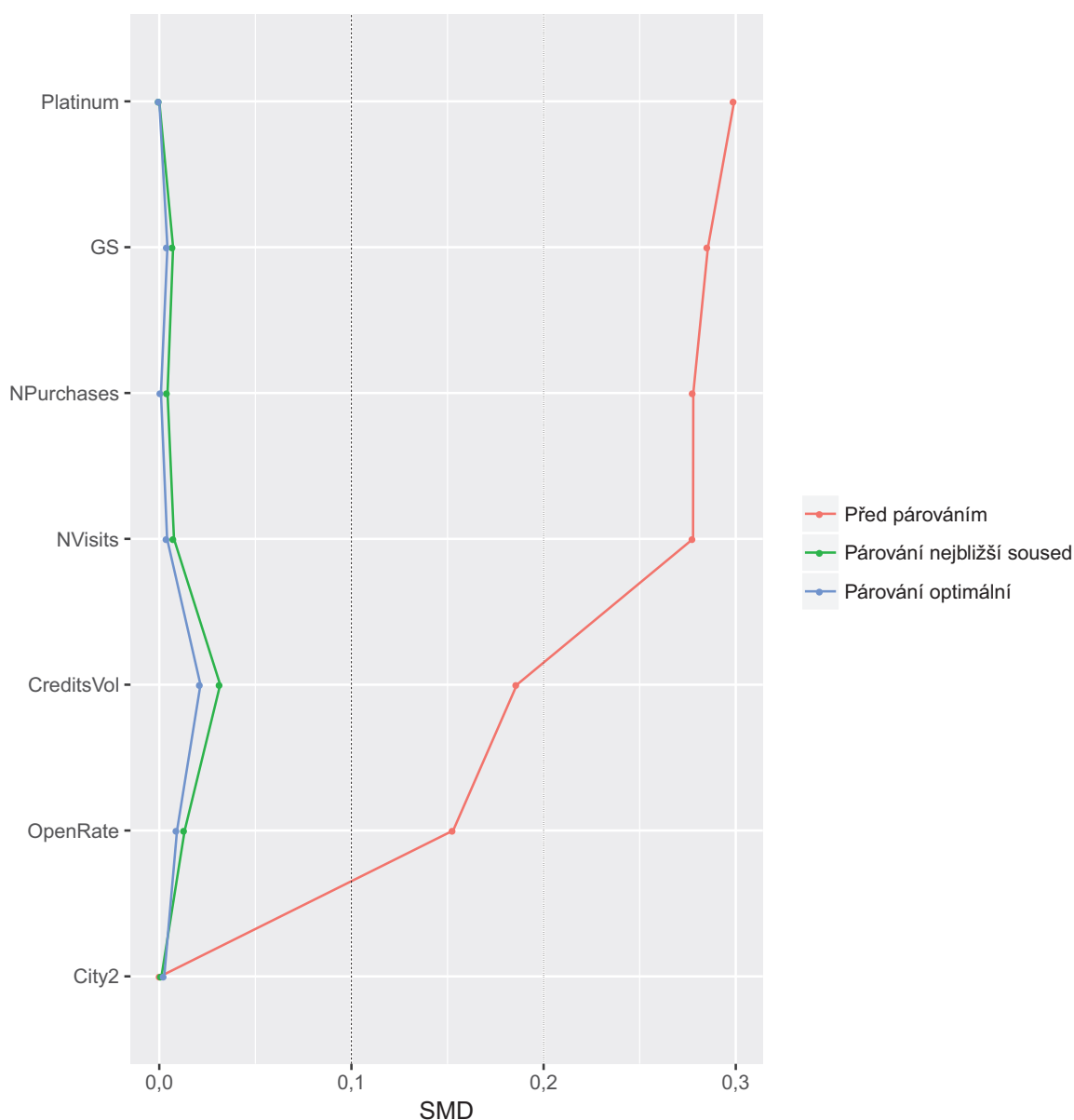
Porovnávat vyváženost mezi různými metodami párování na základě tabulek je málo přehledné, proto je pro pohodlné porovnání vyváženosti mezi různými párovacími procedurami výhodná grafická vizualizace. V grafu na obrázku 5.9 jsou uvedeny hodnoty  $SMD$  všech kovariát pro původní pozorovaný soubor, pro soubor napárovaný metodou nejbližšího souseda a soubor napárovaný optimálně. Kovariáty jsou seřazeny od nejhůře vyvážených k nejlépe vyváženým na základě původního pozorovaného souboru.

Obě použité párovací procedury vedou k dobře vyváženým kovariátům (všechny hodnoty  $SMD$  napárovaných souborů jsou menší než 0,1) a takto napárované soubory nahrazují randomizovaný experiment. Procedura optimálního párování se na použitých datech jeví jako mírně lepší.

### 5.2.2 Odhady efektů

Pro posouzení kauzálního vlivu účasti v akci na podporu prodeje „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  společnosti  $F$  v rámci subpopulace těch, kdo se akce účastnili, lze použít napárované datové soubory získané pomocí obou použitých párovacích procedur. Oba napárované soubory jsou vyvážené a v žádném z nich nebyl z párování vyloučen žádný subjekt ze subpopulace s aktivním ošetřením.

Porovnáme-li tedy průměry všech kovariát skupiny s aktivním ošetřením ve všech uvedených tabulkách číselných charakteristik (tabulky 5.2, 5.4, 5.5), pak je vidět, že se nijak neliší. Odhadnuté výsledné efekty z napárovaných



**Obrázek 5.9:** Graf pro porovnání vyváženosti dosažené různými párovacími procedurami na základě Mahalanobisovy vzdálenosti. Hodnota 0,1 na ose *SMD* udává hraniční hodnotu, pod kterou je kovariáta považována za vyváženou.

souborů lze tedy zobecnit na celou subpopulaci<sup>31</sup> původně pozorovaných

<sup>31</sup>Někdy je potřeba v zájmu vyváženosti některé subjekty ze skupiny s aktivním ošetřením (v uvažovaném kontextu odhadu  $\tau_{ATT}$ ) vypustit a vyváženosti je následně dosaženo na souboru, který nereprezentuje celou původní subpopulaci. Potom lze i odhadnuté kauzální efekty zobecnit jen na část původní subpopulace s aktivním ošetřením.

subjektů s aktivním ošetřením.

Například v případě souboru napárovaného pomocí metody nejbližšího souseda se každý uživatel, jenž se ve skutečnosti účastnil akce „Nakup nad“, dostal do napárované skupiny s aktivním ošetřením právě jednou. Proto je populace uživatelů, kteří se v listopadu účastnili akce, totožná s uživateli, kteří reprezentují případy s aktivním ošetřením v párech. Odhady efektů  $\tau_{ATT}$  z takto napárovaného souboru lze tedy považovat za odhady  $\tau_{ATT}$  v populaci skutečných účastníků akce „Nakup nad“.

Směrodatné chyby odhadu pro výpočet intervalových odhadů efektů lze v případě metody nejbližšího souseda získat standardními statistickými metodami (Budíková, Králová a Maroš, 2010), jelikož každý uživatel, jenž se účastnil akce „Kup nad“, se dostal do napárované skupiny s aktivním ošetřením právě jednou s váhou jedna.

V případě optimálního párování se mohou některé subjekty dostat do výběru opakovaně, tudíž není možné odhadovat směrodatnou chybu odhadu  $\tau_{ATT}$  standardními statistickými postupy. Pak lze použít např. Abadie-Imbensův odhad směrodatné chyby odhadu efektu  $\tau_{ATT}$  pro optimálně napárované soubory (s opakováním), jehož asymptotické vlastnosti jsou popsány v Abadie a Imbens (2006). Tento odhad je také implementován v balíku `Matching` jazyka `R`.

Výsledné odhady kauzálních efektů akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  na základě obou párovacích procedur jsou zaznamenány v tabulkách 5.6 a 5.7, jejich interpretace v kontextu měření efektivity akce na podporu prodeje je v uvedena podkapitole 5.5.

### 5.3 Párování na základě propenzitních skóreů

Metody párování na základě Mahalanobisovy vzdálenosti hledají dvojice do párů „vícerozměrně“ na základě vektoru všech kovariát. Naproti tomu metody párování pomocí propenzitních skóreů hledají dvojice do páru na základě skalární hodnoty propenzitního skóreu. Ten vyjadřuje pravděpodobnost, že uvažovaný subjekt dostane (s ohledem na jemu příslušné hodnoty kovariát zavádějících proměnných) aktivní ošetření. Párování proběhne tak, že k subjektu, který dostal aktivní ošetření, je napárovaný takový subjekt s pozorovaným kontrolním ošetřením, který má stejnou hodnotu



**Tabulka 5.6:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci subpopulace účastníků akce získané na základě párování pomocí metody nejbližšího souseda s Mahalanobisovou metrikou

Výsledky pro metodu nejbližšího souseda					
	průměr pro skupinu s aktivním ošetřením	průměr pro kontrafaktuály s kontrolním ošetřením	rozdíl průměrů	SE	interval spolehlivosti pro efekt
$Y_1$	142,096	274,790	-132,694	3,942	(-140,42; -124,97)
$Y_2$	-21,933	255,365	-277,298	4,807	(-286,72; -267,88)
$Y_3$	417,200	263,665	153,535	6,186	(141,41; 165,66)

**Tabulka 5.7:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci subpopulace účastníků akce získané na základě optimálního párování (s opakováním a deterministicky ošetřenými shodami) s Mahalanobisovou metrikou

Výsledky pro metodu optimálního párování					
	průměr pro skupinu s aktivním ošetřením	průměr pro kontrafaktuály s kontrolním ošetřením	rozdíl průměrů	SE	interval spolehlivosti pro efekt
$Y_1$	142,096	274,488	-132,392	4,024	(-140,28; -124,51)
$Y_2$	-21,934	255,543	-277,477	4,941	(-287,16; -267,79)
$Y_3$	417,200	264,072	153,129	6,458	(140,47; 165,79)

propenzitního skóru jako uvažovaný aktivně ošetřený subjekt. Párování na základě propenzitních skóre zaručuje vyváženost napárovaných skupin pro všechny kovariáty, které ovlivňují propenzitní skór (3.5.5).

Pro párování na základě propenzitních skóre je nezbytná jejich znalost pro každé pozorování. V praxi jsou však tyto skóre obvykle neznámé, a proto se musejí nejdříve na základě hodnot kovariát odhadnout. Propenzitní skóre se obvykle odhadují pomocí logistické (či v případě víceúrovňového ošetření multinomické) regrese. Alternativou mohou být různé metody strojového učení, např. neuronové sítě, gradient boosting či náhodné lesy (Brown, Merrigan a Royer, 2018).

### 5.3.1 Odhady propenzitních skóru

Propenzitní skór je v kontextu řešeného problému pravděpodobnost, že se uživatel rozhodne splnit podmínky akce „Nakup nad“, a získá kredit (tedy „je mu přiřazeno“ aktivní ošetření, účastní se akce). Toto jeho rozhodnutí je podmíněno jeho individuálním nákupním charakterem popsaným hodnotami vektoru kovariát z tabulky 5.3.

Pravděpodobnost, že se uživatel „rozhodne“ účastnit se akce, tedy pravděpodobnost přiřazení aktivního ošetření, bude odhadnuta pomocí logistické regrese. Odhady parametrů logistického modelu jsou ovšem interpretačně zcela nezajímavé (s ohledem na párování), zajímavé jsou výlučně odhady propenzitních skóru. Pokud vyváženost skupin napárovaných na základě odhadnutých propenzitních skóru není vyhovující, je možné se k modelu logistické regrese vrátit a vylepšit jej („vylepšit“ logistický model v této souvislosti znamená vylepšit vyváženost napárovaných skupin pomocí lépe odhadnutých propenzitních skóru).

Odhad modelu logistické regrese je v tabulce 5.8. (Propenzitní skóry lze v jazyce **R** odhadnout tradičně pomocí funkce `glm`, `family=binomial`, následné párování pak pomocí funkcí z balíku `Matching`, příloha B.)

**Tabulka 5.8:** Výstup pro model logistické regrese. Kvalita modelu je v kontextu řešeného problému posuzována vyvážeností napárovaných skupin.

	Odhad	SE	$z$ hodnota	$\Pr(>  z )$
(Intercept)	-1,98E+00	0,02	-111,31	0,00E+00
GS	3,61E-06	0,00	4,83	1,37E-06
NPurchases	-2,53E-03	0,00	-4,57	4,81E-06
NVisits	1,19E-04	0,00	2,05	4,07E-02
CreditsVol	1,13E-04	0,00	13,12	2,65E-39
OpenRate	3,29E-01	0,03	10,38	3,02E-25
Platinum	4,40E-01	0,03	14,25	4,60E-46
City2	-3,94E-02	0,02	-2,05	4,05E-02

Před tím než bude na základě odhadnutých propenzitních skóru provedeno párování, je potřeba posoudit předpoklad překryvu. Dokonalý překryv zajistí, že ke každému subjektu ze skupiny s pozorovaným aktivním ošetře-

ním lze najít subjekt ze skupiny s pozorovaným kontrolním ošetřením, jenž má stejnou pravděpodobnost být zařazen do skupiny s aktivním ošetřením jako uvažovaný aktivně ošetřený subjekt. Proto při párování (stále jako v předchozí kapitole jsou cílem odhady  $\tau_{ATT}$ ) je ke každému subjektu s pozorovaným aktivním ošetřením z čeho vybírat jeho kontrafaktuál ze skupiny s pozorovaným kontrolním ošetřením.

Tento předpoklad lze posoudit vizuálně na základě histogramů odhadnutých propenzitních skóre pro skupinu s pozorovaným aktivním a pozorovaným kontrolním ošetřením. Přitom nejde o to, aby histogramy měly stejný tvar, ale aby měly „ostře kladné“ délky sloupců nad stejným intervalem/intervaly. Pokud by došlo jen k částečnému překryvu, pak by finální výsledky kauzální analýzy nebylo možné zobecnit na celou původní populaci, z níž pocházel datový soubor, ale jen na její část odpovídající oblasti překryvu propenzitních skóre.

Histogram odhadnutých propenzitních skóre pro datový soubor k akci „Nakup nad“ splňuje předpoklad překryvu, jak je vidět na obrázku 5.10. Intervaly s kladnými četnostmi propenzitních skóre pro skupinu s aktivním ošetřením se kryjí s intervaly s kladnými četnostmi propenzitních skóre pro skupinu s kontrolním ošetřením.

### 5.3.2 Vyváženost párování

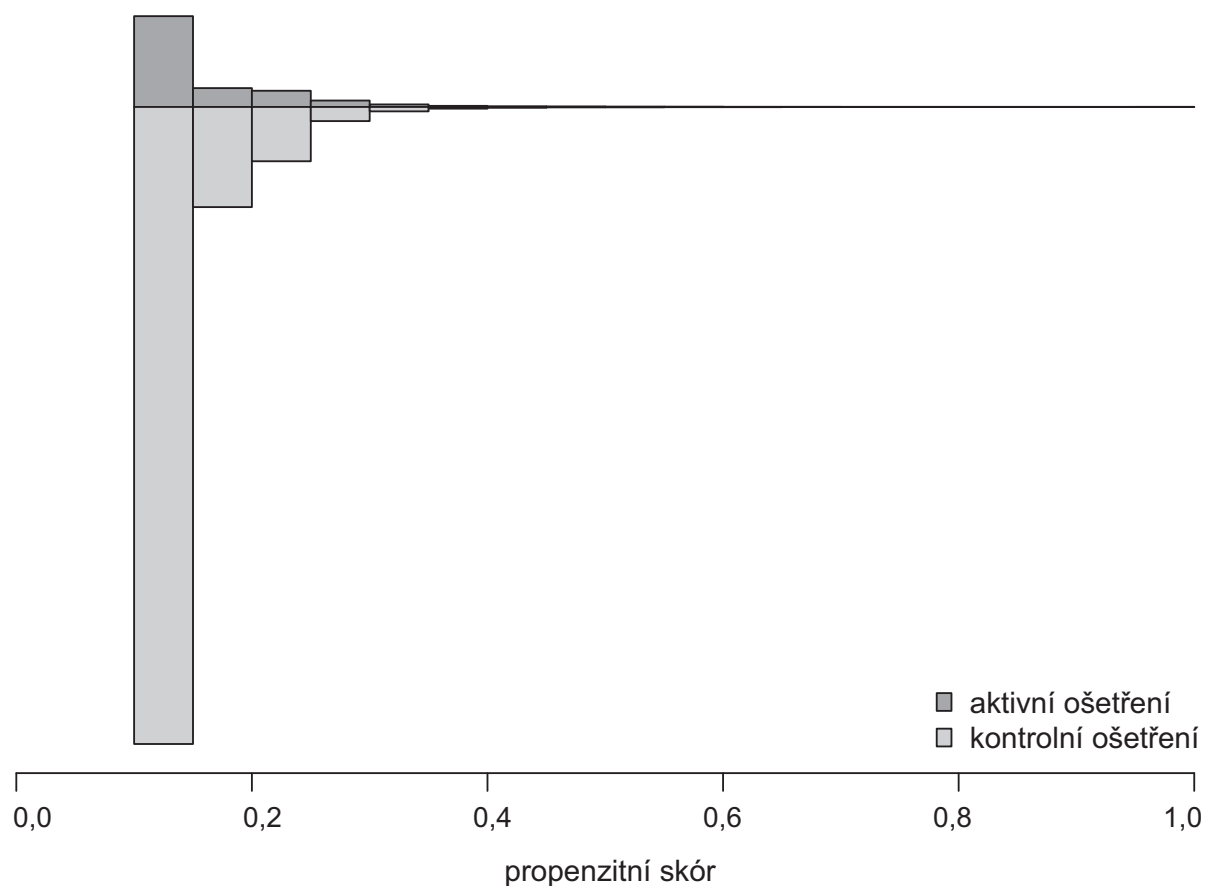
#### Soubor napárováný metodou nejbližšího souseda na základě propenzitních skóre

Párovací algoritmus nejbližšího souseda nyní, na rozdíl od kapitoly 5.2, měří vzdálenost mezi subjekty na základě odhadnutých propenzitních skóre. (Lze použít stejné párovací balíky jazyka **R** jako v předchozí kapitole.)

Vzhledem k tomu, že propenzitní skóre nabývají hodnot z omezeného intervalu  $\langle 0, 1 \rangle$  (mnoho vzdáleností je pak „hodně podobných“), je výhodné tento interval natáhnout pomocí logitu (vzdálenosti se natáhnou, pořadí ovšem zůstanou zachována a párování proběhne na základě hodnot z nekonečné škály). Označuje-li  $e(x)$  propenzitní skór, pak logit propenzitního skóru je

$$\text{logit}(e(x)) = \log \frac{e(x)}{1 - e(x)}.$$

## Histogram propenzitních skóre



**Obrázek 5.10:** Histogram pro posouzení předpokladu překryvu odhadnutých propenzitních skóre. Každý sloupec histogramu propenzitních skóre odhadnutých pro skupinu s aktivním ošetřením „zrcadlí“ (libovolně dlouhý) sloupec propenzitních skóre odhadnutých pro skupinu s kontrolním ošetřením. (Při aktuálním rozlišení a velikosti obrázku jsou obtížně rozpoznatelné velmi malé četnosti pro vysoké hodnoty skóre.)

Propenzitní skór má vyvažovací schopnost, napárované skupiny by tedy měly být s ohledem na zavádějící kovariáty vyvážené. Propenzitní skór byl ovšem pouze odhadován, je tedy potřeba vyváženost ověřit.

Pro párování metodou nejbližšího souseda na základě logitu propenzitního skóre (bez opakování, s deterministicky nastaveným zacházením se shodami) jsou k dispozici hodnoty *SMD* a další statistiky v tabulce 5.9. Hodnoty *SMD* jsou pro všechny sledované kovariáty menší než hraniční hodnota 0,1, napárované skupiny lze tedy s ohledem na kovariáty považovat za vyvážené a efekt ošetření lze odhadnout porovnáním odezev mezi oběma skupinami.

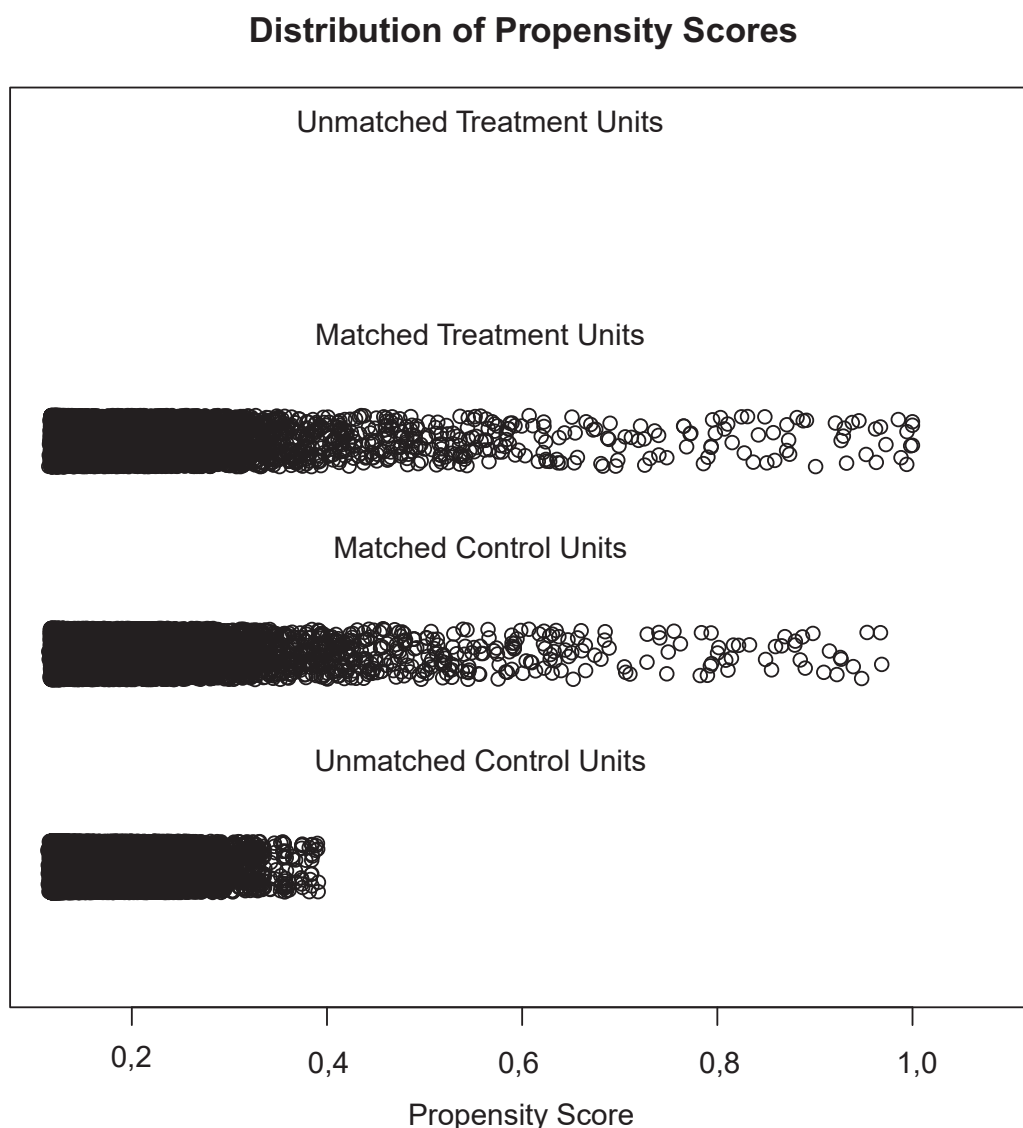
**Tabulka 5.9:** Číselné charakteristiky kovariát v souboru napárovaném metodou nejbližšího souseda, kde jsou metrikou vzdálenosti logity propenzitních skóre. Ve sloupci označeném hodnotou „1“ je všech 14 910 subjektů s pozorovaným aktivním ošetřením. Ve sloupci označeném hodnotou „0“ je 14 910 napárovaných kontrafaktuálů ze skupiny s kontrolním ošetřením. Rozdíl v rozloženích všech kovariát mezi oběma napárovanými skupinami je posouzen na základě hodnot *SMD*.

	0		1		SMD
	<i>n</i>	14 910	<i>n</i>	14 910	
GS (mean (sd))	17 000,94	(32 032,46)	17 070,47	(31 547,31)	0,002 19
NPurchases (mean (sd))	27,67	(51,43)	28,02	(54,93)	0,006 65
NVisits (mean (sd))	221,33	(295,92)	216,61	(297,27)	0,015 94
CreditsVol (mean (sd))	865,08	(2 495,70)	979,32	(4 634,41)	0,030 69
OpenRate (mean (sd))	0,21	(0,30)	0,21	(0,29)	0,011 04
Platinum (mean (sd))	0,28	(0,45)	0,28	(0,45)	0,009 12
City2 (mean (sd))	0,67	(0,47)	0,68	(0,47)	0,012 02

Výsledek párování lze posoudit i orientačně pomocí grafů, jež porovnávají, kolik subjektů a které subjekty z kontrolní a aktivní skupiny se k párování využily. Lze také porovnat tvar histogramů propenzitních skóre ve skupinách s aktivním i kontrolním ošetřením, a to před párováním a po párování.

Na obrázku 5.11 je tzv. jitter graf, z něhož je vidět, že kolika subjektům ze skupiny s aktivním ošetřením se párovacím algoritmem povedlo najít „dvojče“ ze skupiny s kontrolním ošetřením a také kolik subjektů (a které) ze skupiny s kontrolním ošetřením zůstalo nevyužito. Tento obrázek souvisí také s histogramem na obr. 5.10, z něhož je vidět, že četnosti pro některé intervaly propenzitních skóre ve skupině s kontrolním ošetřením jsou podstatně větší než ve skupině s aktivním ošetřením.

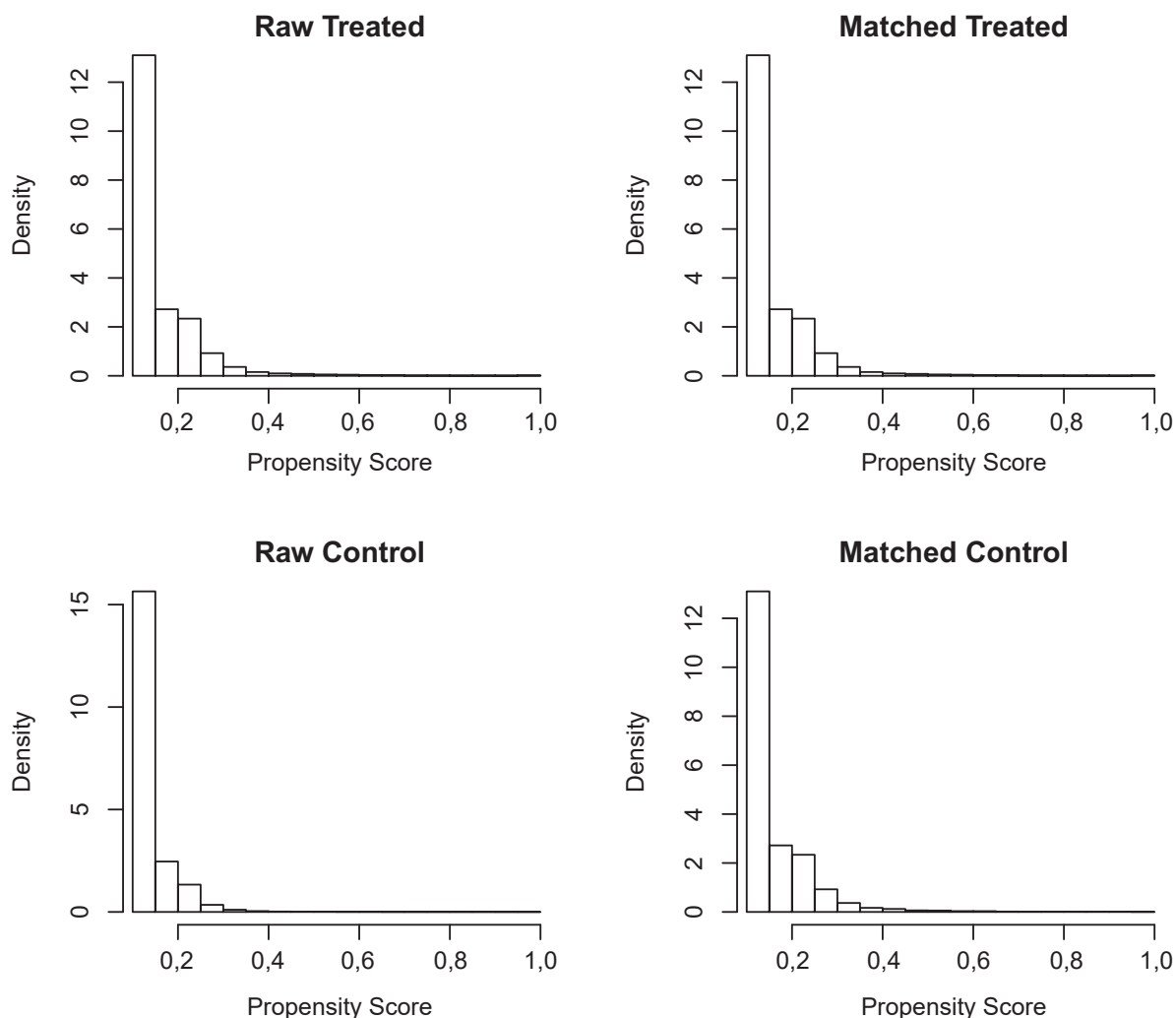
Pro interpretaci v kontextu řešeného problému měření efektu akce „Nakup nad“ v rámci subpopulace těch, kdo se akce účastnili, je výhodné, že ke všem subjektům s aktivním ošetřením jsou napárovány subjekty s kontrolním ošetřením, jak je zřejmé z obrázku 5.11. Takto zůstává subpopulace s aktivním ošetřením, ke které se budou vztahovat závěry, nezměněna.



**Obrázek 5.11:** Z grafu je vidět, že žádný subjekt ze skupiny s aktivním ošetřením nezůstal nenapárovaný. Dále je vidět, že k párování nevyužité subjekty ze skupiny s kontrolním ošetřením jsou ty s malou hodnotou propenzitního skóru. Naopak s rostoucí hodnotou propenzitního skóru subjektů v obou skupinách ubývá a zejména pro velké hodnoty propenzitních skóru je málo subjektů s kontrolním ošetřením.

Další grafy pro orientační posouzení úspěšnosti párovací procedury jsou na obrázku 5.12. Jedná se o histogramy odhadnutých propenzitních skóru, a to před párováním a po párování zvlášť ve skupině s aktivním a ve skupině s kontrolním ošetřením.

Shoda histogramů odhadnutých propenzitních skóru ve skupině s aktivním ošetřením před párováním a po něm zajišťuje již zmíněnou možnost zobecnění závěrů ohledně efektů akce na celou subpopulaci aktivně ošetře-



**Obrázek 5.12:** Histogramy odhadnutých propenzitních skóreů. Nahoře jsou histogramy pro skóreů subjektů s aktivním ošetřením před párováním (Raw Treated) a po něm (Matched Treated). Dole jsou histogramy pro skóreů subjektů s kontrolním ošetřením před párováním (Raw Control) a po něm (Matched Control). Při posuzování podobnosti histogramů je potřeba reflektovat různé škálování na svislé ose.

ných (účastníků akce). Velká podobnost histogramů mezi napárovanými skupinami naznačuje, že napárovaná „dvojčata“ ze skupiny s kontrolním ošetřením jsou vhodnými náhradami nepozorovatelných kontrafaktuálů k subjektům s pozorovaným aktivním ošetřením.

Nepodobnost histogramů před párováním a po něm ze skupiny s kontrolním ošetřením nepředstavuje žádný problém, jen ukazuje, která část subjektů ze skupiny s kontrolním ošetřením zůstala pro párování nevyužita.

## Soubor napárovaný metodou nejbližšího souseda na základě propenzitních skóru s nastavenou hodnotou měřítka pro okolí

Párování metodou nejbližšího souseda na základě propenzitních skóru bylo zcela uspokojivé jak s ohledem na překryv, tak s ohledem na vyváženost zavádějících kovariát, takže není žádný podstatný důvod model vylepšovat. Ovšem na obrázku 5.11 je vidět, že několik málo subjektů s velkými propenzitními skóry ze skupiny s aktivním ošetřením bylo zřejmě napárováno s objekty ze skupiny s kontrolním ošetřením, jež měly skóry o něco menší. Tomu lze zabránit tak, že se párovacímu algoritmu napevno nastaví maximální „povolená“ vzdálenost pro to, jaké subjekty lze spárovat.

Tato vzdálenost se nazývá měřítko pro okolí (*caliper*) a je podrobněji popsána v podkapitole 3.5.1. Obvykle se velikost měřítka pro okolí nastavuje na 0,2násobek směrodatné odchylky té proměnné, podle které se měří vzdálenost. Jelikož předchozí párovací procedura měřila vzdálenost na logitové škále propenzitního skóru, měřítko okolí je pak  $caliper = 0,2 * sd(\text{logit}(e(x)))$  (na tuto skutečnost je potřeba pamatovat při zadávání argumentů párovacích funkcí v **R**).

Výsledkem párovací procedury podle metody nejbližšího souseda s metrikou logitu propenzitních skóru a nastaveným měřítkem pro okolí na hodnotu  $0,2 * sd(\text{logit}(e(x)))$  je napárovaný datový soubor, z něhož bylo ve skupině s aktivním ošetřením proti původnímu pozorovanému souboru vyloučeno pouze 43 subjektů (z původních 14 910, což je výborný výsledek). Vyváženost párování lze opět posoudit na základě hodnot *SMD* všech kovariát, jež jsou uvedeny v tabulce 5.10 a všechny jsou menší než hraniční hodnota 0,1.

Pro pohodlné porovnání vyváženosti mezi párováním pomocí metody nejbližšího souseda na základě logitu propenzitního skóru s měřítkem okolí, bez měřítka okolí a původním pozorovaným souborem slouží graf na obrázku 5.13. Kovariáty zavádějících proměnných jsou seřazeny od nejhůře vyvážených k nejlépe vyváženým na základě původního pozorovaného souboru.

Je vidět, že párování s měřítkem i bez měřítka dává velmi podobné výsledky, vyváženost je mírně lepší u párování s měřítkem.



**Tabulka 5.10:** Číselné charakteristiky kovariát v souboru napárovaném metodou nejbližšího souseda, kde metrika vzdálenosti je logit propenzitního skóru a měřítko pro okolí je  $0,2 \cdot \text{sd}$  logitu propenzitního skóru. V sloupci označeném hodnotou „1“ je 14 867 takových subjektů (z původních 14 910) s pozorovaným aktivním ošetřením, ke kterým se ve vzdálenosti povolené měřítkem podařilo najít jejich kontrafaktuály. Počet subjektů s aktivním ošetřením, ke kterým se nepodařilo najít subjekty ze skupiny s kontrolním ošetřením, je 43.

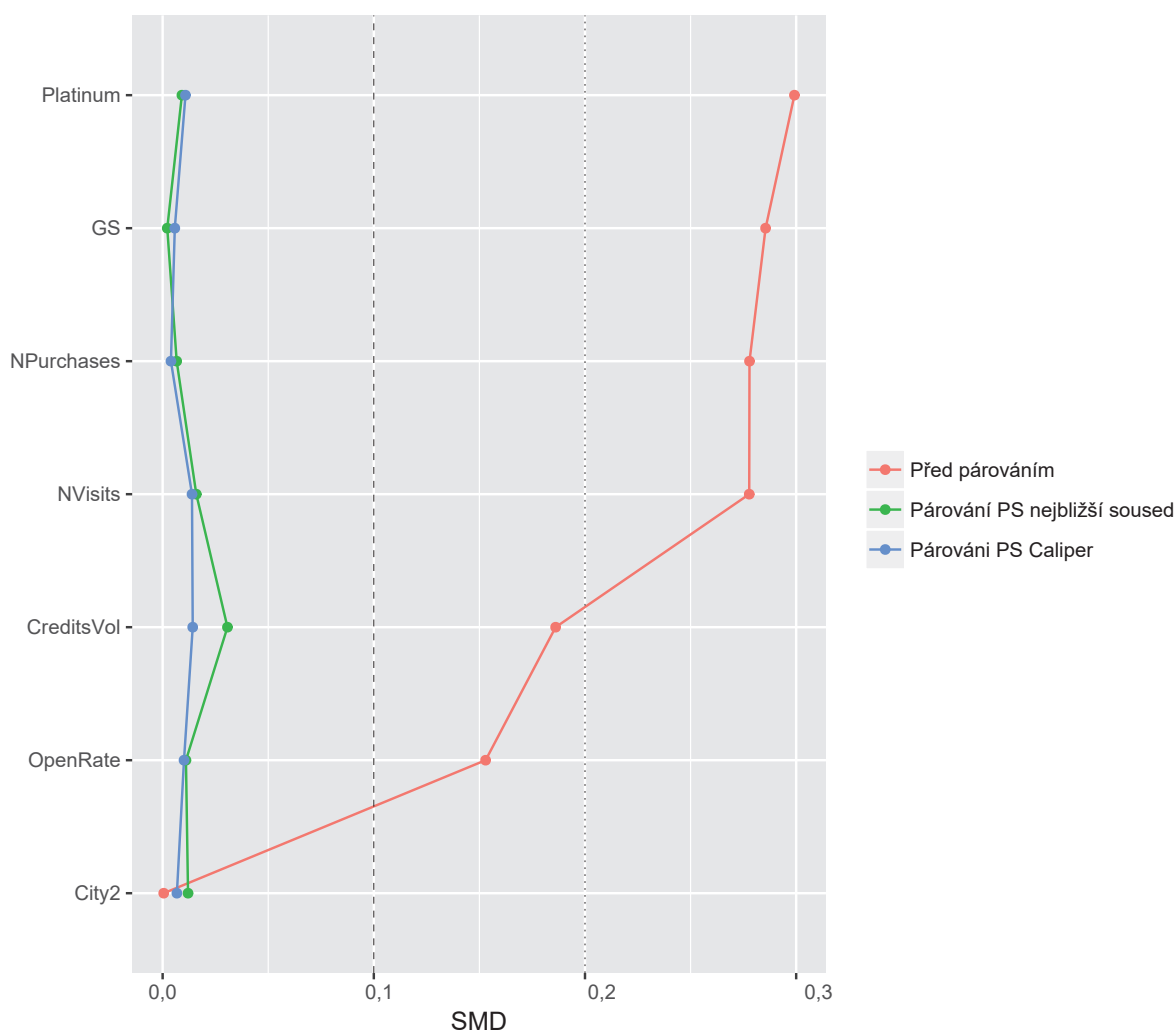
	0		1		SMD
	<i>n</i>	14 867	14 867		
GS (mean (sd))	16 729,28	(31 429,55)	16 553,39	(29 673,34)	0,005 75
NPurchases (mean (sd))	27,35	(50,81)	27,15	(51,48)	0,003 93
NVisits (mean (sd))	218,32	(293,96)	214,24	(293,79)	0,013 90
CreditsVol (mean (sd))	832,38	(2 374,32)	866,51	(2 417,26)	0,014 24
OpenRate (mean (sd))	0,21	(0,30)	0,21	(0,29)	0,010 11
Platinum (mean (sd))	0,28	(0,45)	0,28	(0,45)	0,010 81
City2 (mean (sd))	0,67	(0,47)	0,68	(0,47)	0,006 75

### 5.3.3 Odhady efektů

V případě párování na základě logitu propenzitních skóre metodou nejbližšího souseda bez nastaveného měřítka okolí je skupina zákazníků s aktivním ošetřením v napárovaném souboru totožná se skupinou zákazníků s aktivním ošetřením v původním datovém souboru. Proto jsou také popisné statistiky pro skupinu s aktivním ošetřením v tabulkách 5.9 a 5.2 stejné. Níže odhadnuté efekty akce „Nakup nad“ na zisky typu  $\tau_{ATT}$  tedy lze zobecnit na celou subpopulaci uživatelů, kteří se účastnili akce.

Pro skupinu s aktivním ošetřením odpovídající párování na základě propenzitních skóre s nastaveným měřítkem okolí jsou již ovšem popisné statistiky jiné (viz tabulku 5.10). Proto také efekty  $\tau_{ATT}$  odhadnuté na základě párování s měřítkem okolí nelze zobecnit na celou populaci uživatelů, ale pouze na takovou subpopulaci uživatelů s aktivním ošetřením, pro něž byl v této párovací proceduře nalezen pár s kontrolním ošetřením. Takových uživatelů s aktivním ošetřením bylo 14 867 z původních 14 910.

Abadie-Imbensovy odhady směrodatných chyb odhadů efektů, na základě kterých jsou odvozeny intervalové odhady efektů, bohužel nepostihují variabilitu způsobenou odhadováním samotných propenzitních skóre.



**Obrázek 5.13:** Graf pro porovnání vyváženosti mezi párovacími procedurami na základě propenzitních skóre s měřítkem, bez měřítka a původním nenapárovaným datovým souborem. Hodnota 0,1 na ose  $SMD$  udává hraniční hodnotu, pod kterou je kovariáta považována za vyváženou.

Výsledné odhady kauzálních efektů akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  na základě párování pomocí propenzitních skóre jsou uvedeny v tabulkách 5.11 a 5.12, jejich interpretace v kontextu měření efektivity akce na podporu prodeje je provedena v podkapitole 5.5.

## 5.4 Vážení propenzitními skóre

Metoda vážení propenzitními skóre (IPTW, 3.6.2) umožní odpovědět na výzkumné otázky VO4 až VO6, které se vztahují k celé populaci všech uživatelů společnosti  $F$  (tedy ne pouze k subpopulaci uživatelů, již se účastnili akce). Tyto otázky se ptají na to, jaké by byly v rámci populace

**Tabulka 5.11:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci subpopulace účastníků akce získané na základě párování pomocí propenzitních skóru

Výsledky pro párování dle propenzitních skóru					
	průměr pro skupinu s aktivním ošetřením	průměr pro kontrafaktuály s kontrolním ošetřením	rozdíl průměrů	SE	interval spolehlivosti pro efekt
$Y_1$	142,096	271,641	-129,545	4,071	(-137,52; -121,57)
$Y_2$	-21,934	252,031	-273,964	4,929	(-283,62; -264,30)
$Y_3$	417,200	259,480	157,720	6,352	(145,27; 170,17)

**Tabulka 5.12:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci subpopulace účastníků akce získané na základě párování pomocí propenzitních skóru a nastaveného měřítka okolí

Výsledky pro párování dle propenzitních skóru s měřítkem okolí					
	průměr pro skupinu s aktivním ošetřením	průměr pro kontrafaktuály s kontrolním ošetřením	rozdíl průměrů	SE	interval spolehlivosti pro efekt
$Y_1$	139,986	269,172	-129,186	4,065	(-137,15; -121,22)
$Y_2$	-18,035	249,950	-267,984	4,063	(-275,95; -260,02)
$Y_3$	422,083	257,772	164,311	5,456	(153,62; 175,00)

všech uživatelů společnosti  $F$  zisky  $Y_1$ ,  $Y_2$  a  $Y_3$ , kdyby se akce na podporu prodeje „Nakup nad“ nikdo neúčastnil a následně, jaký by byl skutečný efekt akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci celé populace.

Metody párování bohužel umožňují odhadovat pouze efekty typu  $\tau_{ATT}$ , nebo případně  $\tau_{ATC}$ . Vztahují se tedy vždy jen k nějaké subpopulaci z původní populace všech pozorovaných subjektů. Metoda IPTW umožňuje odhadnout i efekty typu  $\tau_{ATE}$ . Místo původní pozorované populace lze touto metodou uměle nasimulovat pseudopopulaci, v rámci které se ošetření (účast/neúčast v akci) přiřadí jednotlivým subjektům „jakoby“ náhodně. Pak lze porovnáním skupin s aktivním a kontrolním ošetřením zjistit, jaký je efekt ošetření na celou populaci.

V kontextu akce „Nakup nad“ jde tedy o to porovnat, jaké by byly zisky, kdyby se akce „povinně“ účastnili všichni uživatelé, se situací, kdy by se akce neúčastnil nikdo. Jedná se tedy o efekt akce na celou populaci uživatelů společnosti  $F$ . Princip odhadu  $ATE$  je graficky ilustrován na obr. 5.4.

Z podkapitoly 5.3 jsou již propenzitní skóry odhadnuty, stejně tak je již ověřen dostatečný překryv odhadnutých propenzitních skóru. Na základě odhadnutých propenzitních skóru se následně přepočítají váhy pro každý pozorovaný subjekt tak, že každý subjekt s pozorovaným aktivním ošetřením a odhadnutým propenzitním skórem  $\hat{e}(x)$  dostane váhu  $\frac{1}{\hat{e}(x)}$  a každý subjekt s pozorovaným kontrolním ošetřením a odhadnutým propenzitním skórem  $\hat{e}(x)$  dostane váhu  $\frac{1}{1-\hat{e}(x)}$ .

Po převážení pozorovaných subjektů odhadnutými váhami by skupina s pozorovaným aktivním ošetřením i skupina s pozorovaným kontrolním ošetřením měly mít stejné rozložení kovariát. Toto lze posoudit na základě tabulky 5.13.

**Tabulka 5.13:** Uvedené počty subjektů s aktivním i kontrolním ošetřením se vztahují k pseudopopulaci nasimulované převažováním původně pozorovaných subjektů. Všechny hodnoty  $SMD$  jsou podstatně menší než hraniční hodnota 0,1, nasimulované pseudopopulace jsou tedy vyvážené.

	0	1	SMD
$n$	102 747,20	102 203,50	
GS (mean (sd))	11 328,73 (25 090,68)	11 120,42 (20 385,92)	0,009
NPurchases (mean (sd))	18,12 (38,92)	17,60 (36,01)	0,014
NVisits (mean (sd))	156,18 (236,87)	153,41 (234,30)	0,012
CreditsVol (mean (sd))	509,73 (2 294,52)	469,13 (2 070,41)	0,019
OpenRate (mean (sd))	0,17 (0,27)	0,17 (0,26)	0,004
Platinum (mean (sd))	0,17 (0,38)	0,17 (0,38)	0,007
City2 (mean (sd))	0,68 (0,47)	0,68 (0,47)	< 0,001

Pro výpočet  $SMD$  každé kovariáty  $X$  byly použity její vážený průměr a vážený rozptyl ve skupině s aktivním i kontrolním ošetřením. Pro skupinu s aktivním ošetřením je vážený průměr  $M_{e1}$  a vážený rozptyl  $S_{e1}^2$  počítán

dle těchto vzorců:

$$M_{e1} = \frac{\sum_{i: W_i=1} \frac{X_i}{\hat{e}_i}}{\sum_{i: W_i=1} \frac{1}{\hat{e}_i}}$$

$$S_{e1}^2 = \frac{\sum_{i: W_i=1} \frac{1}{\hat{e}_i} (X_i - M_{e1})^2}{\frac{N_1-1}{N_1} \sum_{i: W_i=1} \frac{1}{\hat{e}_i}},$$

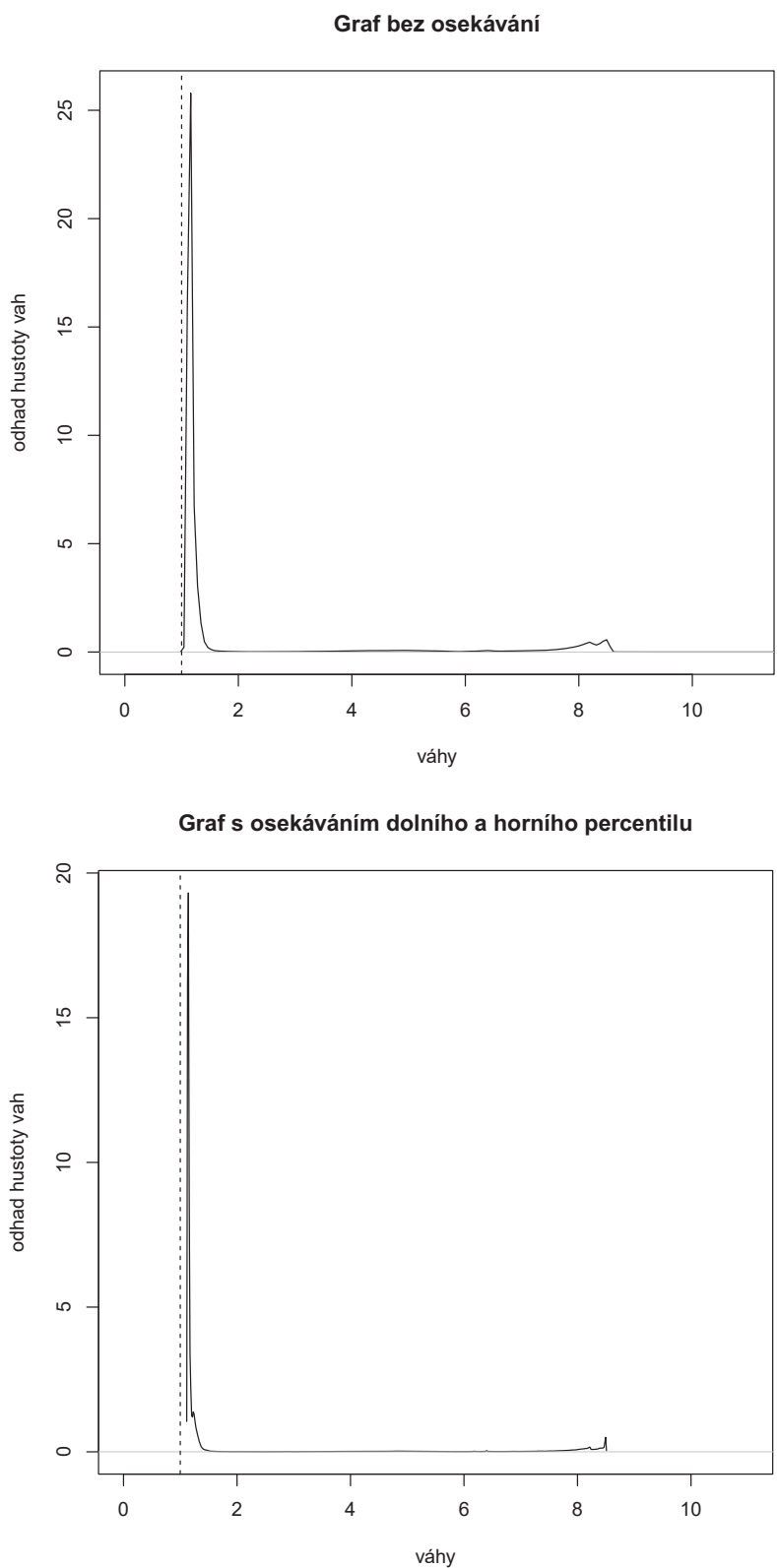
kde  $N_1$  je počet účastníků s aktivním ošetřením. (Pro skupinu s kontrolním ošetřením analogicky.)

Funkce v jazyce **R** potřebné pro použití metody IPTW jsou dostupné např. v rámci balíku `ipw` (příloha B).

I v případě, že jsou obě skupiny nasimulované pseudopopulace dobře vyvážené, mohou být odhady efektů nespolehlivé (s velkou variabilitou) v důsledku velmi vlivných subjektů. Jedná se o ty subjekty, jimž byla přiřazena extrémně velká váha. Proto je žádoucí takovéto subjekty identifikovat a případně je odstranit například pomocí osekávání (*truncation*). (To může způsobit vychýlení odhadu efektu, ale sníží jeho variabilitu.)

Pro představu o rozložení vah (zajímavé jsou zejména extrémně velké váhy) slouží jednak první řádek tabulky popisných statistik vah 5.14 a dále graf hustot vah, jenž je na obrázku 5.14 nahoře. Váhy jsou zdola omezeny hodnotou 1, takže minimální pozorovaná hodnota vah odpovídá nejmenší možné hodnotě 1. Maximální pozorovaná váha je ovšem až 31,66, přičemž horní kvartil je pouze 1,23. Tato skutečnost (i hodnota průměru rovná dvěma) naznačuje, že problematických pozorování s příliš velkými váhami bude spíše málo a na „vzdáleném“ pravém konci rozložení vah. Totéž je vidět i z tvaru hustoty vah na pravém konci jejího rozložení. Z toho lze usoudit, že negativní dopad vlivných subjektů na odhad efektů zřejmě není závažný.

Přesto je rozumné vlivné subjekty s velkými váhami odstranit, což je možné zařídit například osekáváním (3.6.3) na prvním a devadesátém devátém kvantilu. V druhém řádku tabulky 5.14 a v dolním grafu na obrázku 5.14 je již oseknutý soubor vah, kde byla každému subjektu, jehož váha překročila hodnotu 8,49, jeho váha uměle stlačena na hodnotu 8,49 (totéž pro první kvantil). Všechny subjekty tedy zůstaly v souboru zachovány, jen se některým jejich váha „osekla“.



**Obrázek 5.14:** V grafu bez osekávání je na vodorovné ose škálování pouze po hodnotu 10, ačkoliv nejvlivnějších 20 pozorování nabývá hodnoty z intervalu  $(8,6; 31,66)$ . Jejich četnosti však byly tak malé, že se v grafu hustot nezobrazovaly.

**Tabulka 5.14:** Přehled kvantilů a dalších popisných statistik: v prvním řádku pro celý soubor vah (odpovídající všem původně pozorovaným subjektům); ve druhém řádku pro soubor vah oseknutý na prvním a devadesátém devátém kvantilu. V souboru bez osekávání je maximální váha 31,66; v souboru s osekáváním je maximum pouze 8,49.

	Min.	Dol. kv.	Med.	Prům.	Hor. kv.	Max.
celý soubor vah	1,00	1,14	1,15	2,00	1,23	31,66
oseknutý soubor vah	1,13	1,14	1,15	2,00	1,23	8,49

Na nasimulované pseudopopulaci je již možné odhadnout efekty ošetření na odezvy v rámci celé populace uživatelů společnosti  $F$ . Efekt účasti v akci „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  bude odhadnut pomocí<sup>32</sup> lineárního marginálního strukturálního modelu (MSM) pomocí vztahu (3.6.3):

$$E(Y(w)) = \psi_0 + \psi_1 w, \quad w = 0, 1,$$

přičemž průměrný efekt ošetření  $\tau_{ATE}$  odpovídá parametru  $\psi_1 = E(Y(1)) - E(Y(0))$ . V pozici  $Y$  v 3.6.3 jsou teď postupně zisky společnosti  $Y_1$ ,  $Y_2$  a  $Y_3$ . Odhadnuté efekty jsou vloženy pro pseudopopulaci bez osekávání do tabulky 5.15 a pro pseudopopulaci upravené osekáváním do tabulky 5.16.

**Tabulka 5.15:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci celé populace společnosti  $F$ . Odhady jsou získány pomocí metody IPTW, ve sloupci SE jsou robustní sandwichové odhady variability odhadu efektu a váhami nasimulovaná pseudopopulace je bez osekávání.

Výsledky získané pomocí IPTW pro pseudopopulaci bez osekávání					
	průměr pro pseudo- populaci s aktivním ošetř.	průměr pro pseudo- populaci s kontrol. ošetř.	rozdíl průměrů $\hat{\psi}_1$	SE	interval spolehlivosti pro efekt
$Y_1$	121,60	273,36	-151,76	2,45	(-156,56; -146,95)
$Y_2$	-12,86	258,49	-271,35	2,41	(-276,08; -266,62)
$Y_3$	427,61	266,18	161,43	4,39	(152,82; 170,04)

<sup>32</sup> MSM je výhodné, zejména pokud je kromě zavádějících proměnných do modelu zařazen i vliv dalších prediktorů (nezavádějících!). V řešené úloze žádný další prediktor není, stačilo by tedy pouze porovnat průměry obou pseudopopulací. Ovšem funkce `svyglm` v jazyce **R** je univerzálně předchystaná i pro modely s dalšími prediktory, a je tedy pohodlné ji využít.

**Tabulka 5.16:** Bodové odhady (rozdíl průměrů) a intervalové odhady kauzálního efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci celé populace společnosti  $F$ . Odhady jsou získány pomocí metody IPTW, ve sloupci SE jsou robustní sandwichové odhady variability odhadu efektu a váhami nasimulovaná pseudopopulace byla upravena osekáváním.

Výsledky získané pomocí IPTW pro pseudopopulaci s osekáváním					
	průměr pro pseudo- populaci s aktivním ošetř.	průměr pro pseudo- populaci s kontrol. ošetř.	rozdíl průměrů $\hat{\psi}_1$	SE	interval spolehlivosti pro efekt
$Y_1$	121,65	273,33	-151,68	2,45	(-156,47; -146,89)
$Y_2$	-12,93	258,49	-271,42	2,41	(-276,15; -266,69)
$Y_3$	427,52	266,19	161,33	4,40	(152,71; 169,95)

Z porovnání bodových i intervalových odhadů efektů mezi tabulkami 5.15 a 5.16 je vidět, že osekávání nebylo potřeba, jelikož odhady se téměř neliší.

Interpretace odhadnutých efektů v kontextu měření efektivity akce na podporu prodeje je v podkapitole 5.5.

## 5.5 Interpretace odhadů, diskuze a omezení

Smyslem akce na podporu prodeje „Nakup nad“ je zvýšit zisky společnosti. Součástí akce je ovšem „rozdávání“ kreditů, které má na jedné straně uživatele povzbudit k výraznějšímu nakupování (a v důsledku má vést ke zvyšování zisků společnosti), na druhé straně ale uplatněné kredity představují pro společnost náklady. Vzhledem k tomu, že společnost  $F$  pouze přeprodává zboží a služby, její zisky jsou tvořeny již tak dost nízkou marží. Pro výhodnost akce na podporu prodeje je tedy kritické zejména uplatnění rozdaných kreditů.

Marketéři firmy by zajisté nejraději vyřešili problém „Co by bylo se ziskem, kdyby akce nebyla?“. Ovšem s ohledem na charakter ošetření  $W$  (v této kapitole je definováno jako účast/neúčast v akci) se lze pouze ptát, co by bylo, kdyby se uživatelé akce neúčastnili. Tato otázka je ekvivalentní s otázkou, co by bylo se zisky, kdyby společnost  $F$  nerozdala žádný kredit. Jelikož zejména rozdané kredity představují pro společnost  $F$  riziko



nevýhodnosti akce, takto zjednodušená formulace problému výhodnosti akce je postačující.

Při seznámení se s vlastnostmi všech uživatelů společnosti (tabulka 5.2) je vidět, že ti, kteří se akce účastnili, jsou celkově vzhledem k pozorované historii i dalším stavovým vlastnostem ti „utrácivější“. Skutečnost, že během celého sledovaného období (pokrývá období konání akce a následující období uplatnění kreditů) nakonec vygenerovali větší zisk než skupina těch, kdo se akce neúčastnili, tedy vůbec není překvapivá. Je však těžké říct, zda tohoto pro společnost  $F$  lepšího výsledku dosáhli kvůli konání akce samotné (jež vyvolala účast v akci), nebo kvůli tomu, že celkově více utrácí.

Tento problém vedl k formulování výzkumných otázek **VO1**, ..., **VO6**, které lze dále roztrždit podle toho, ke které populaci, a nebo, ke kterému období se vztahují.

- **Podle vztažené populace**

Výzkumné otázky **VO1**, **VO2** a **VO3** se vztahují k subpopulaci těch uživatelů společnosti  $F$ , kteří se akce „Nakup nad“ skutečně účastnili. Zjišťují dopad akce výhradně na tuto subpopulaci, a to s ohledem na zisky společnosti definované prostřednictvím proměnných  $Y_1$ ,  $Y_2$  a  $Y_3$ .

Výzkumné otázky **VO4**, **VO5** a **VO6** se vztahují k celé populaci všech uživatelů společnosti  $F$ . Zjišťují dopad akce na celou populaci, a to s ohledem na zisky společnosti definované prostřednictvím proměnných  $Y_1$ ,  $Y_2$  a  $Y_3$ .

- **Podle období**

Jednotlivá období související s akcí „Nakup nad“ jsou znázorněna na obr. 5.1.

Společnost  $F$  zajímají zisky vzniklé v době uplatňování kreditů, které jsou malé až záporné. Nižší zisky jsou generovány jednak kvůli uplatněným kreditům samotným a dále kvůli nižší míře nakupování bezprostředně po konání akce. Nižší míra nakupování je způsobena tím, že zejména účastníci akce v předchozím období hodně utráceli, aby získali kredity, a v následném období již nakupují méně. Není ale

jasné, jak by se chovali, kdyby se akce neúčastnili (nelze je srovnat s druhou skupinou „neúčastníků“, protože ti celkově méně nakupují). S tímto obdobím souvisí výzkumné otázky **VO1**, **VO2**, **VO4** a **VO5**.

Pak jsou pro společnost zajímavé celkové čisté zisky za celé sledované období. Do nich se promítají zisky jak za období akce s intenzivnějším nakupováním, tak za období s uplatňováním kreditů, které je celkově na zisky slabší. S celým sledovaným obdobím souvisí výzkumné otázky **VO3** a **VO6**.

### 5.5.1 Co lze usuzovat na základě observačních dat?

Ještě před použitím metod kauzální analýzy lze orientačně, na základě porovnání pozorovaných skupin účastníků a neúčastníků akce, očekávat následující zisky:

Pro hrubé zisky  $Y_1$  v období uplatnění kreditů:

Kdyby se skuteční účastníci akce hypoteticky akce neúčastnili, tak by pravděpodobně nakupovali více než ti, kdo se akce neúčastnili ve skutečnosti (jelikož jsou utráčivější). Potom by rozdíl mezi jejich pozorovanými hrubými zisky a hrubými zisky, které by vygenerovali, kdyby se akce neúčastnili (kontrafaktuály), měl být větší než mezi pozorovanými účastníky a neúčastníky (souvislost s **VO1**). Tento rozdíl reprezentuje ztrátu (kvůli zápornému znaménku) společnosti na hrubých ziscích způsobenou účastí v akci v rámci skupiny účastníků. Je zejména způsobena tím, že tomuto období předcházelo období intenzivních nákupů indukovaných akcí.

Pro čisté zisky  $Y_2$  v období uplatnění kreditů:

U skutečných účastníků akce, kdyby se hypoteticky akce neúčastnili (kontrafaktuály), lze očekávat, že i jimi vytvořené čisté zisky by měly být větší než u pozorované skupiny skutečných „neúčastníků“ (souvisí s **VO2**). Rozdíl mezi pozorovanými čistými zisky a čistými zisky, které by vznikly, kdyby se skuteční účastníci akce hypoteticky neúčastnili, reprezentuje ztrátu (kvůli zápornému znaménku) společnosti na čistých ziscích způsobenou účastí v akci v rámci skupiny účastníků. Ztráta je způsobena nejenom nižší intenzitou nakupování v tomto období, ale také uplatňováním kreditů.

Pro čisté zisky  $Y_3$  v celém sledovaném období:

Kdyby se skuteční účastníci akce této akce hypoteticky neúčastnili (kontrafaktuály), tak by pravděpodobně za celé sledované období i tak vygenerovali

společnosti větší čisté zisky než pozorovaná skupina „skutečných neúčastníků“ (souvisí s **VO3**). Rozdíl mezi celkovými čistými zisky, které za celé období přinesli firmě skuteční účastníci akce, a hypotetickými čistými zisky skutečných účastníků, kdyby se akce neúčastnili, by tedy měl být menší než rozdíl mezi pozorovanou skupinou účastníků a neúčastníků.

Na subpopulaci účastníků akce lze tedy na základě pozorovaných rozdílů mezi skutečnými účastníky a neúčastníky a s ohledem na to, že skuteční účastníci akce jsou utrácivější než skuteční neúčastníci, očekávat: větší ztrátu u  $Y_1$  a  $Y_2$  způsobenou akcí, než se jeví z observačních dat, menší zisk u  $Y_3$  způsobený akcí, než se jeví z observačních dat.

Orientačním východiskem pro posuzování efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  v rámci celé populace všech uživatelů společnosti  $F$  jsou hodnoty celkového průměru  $Y_1$ ,  $Y_2$  a  $Y_3$  na uživatele, uvedené v tabulce 5.1 (souvisí s **VO4**, **VO5** a **VO6**). Vzhledem k tomu, že se jedná o vážené průměry, jež reflektují různé velikosti subpopulace uživatelů a neuživatelů akce, očekávání ohledně kauzálního efektu nejsou přímočará.

### 5.5.2 Co bylo zjištěno kauzální analýzou?

Pro zodpovězení výzkumných otázek **VO1**, **VO2** a **VO3** byly použity párovací algoritmy, a to metoda nejbližšího souseda na základě Mahalanobisovy vzdálenosti, optimální metoda párování na základě Mahalanobisovy vzdálenosti a metoda nejbližšího souseda na základě propenzitních skóru.

Bodové odhady efektu akce „Nakup nad“ na zisky  $Y_1$ ,  $Y_2$  a  $Y_3$  se při řešení metodou nejbližšího souseda na základě Mahalanobisovy vzdálenosti a optimální metodou párování na základě Mahalanobisovy vzdálenosti téměř neliší (liší se řádově v desítkách haléřů). Bodové odhady metodou nejbližšího souseda na základě propenzitního skóru se od předchozích mírně liší (řádově v jednotkách korun). Ovšem 95% intervalové odhady všech efektů u každé ze tří metod pokrývají bodové odhady zbylých metod. Vzhledem k tomu, že výsledky byly odhadovány na souboru, který je totožný s populací, budou dále interpretovány pouze bodové odhady efektů. Tento přístup je v souladu s potřebami společnosti  $F$  – tu zajímá, jaký byl skutečný (kauzální) dopad akce „Nakup nad“ na její listopadové zisky.

Pro interpretaci kauzálního efektu budou upřednostněny výsledky získané optimální metodou párování.

- Ve srovnání s metodou nejbližšího souseda na základě Mahalanobisovy vzdálenosti vedou výsledky optimálního párování na „globální optimum“, tedy nejmenší možnou celkovou vzdálenost napárovaných dvojic.
- Metoda nejbližšího souseda na základě propenzitního skóru má i přes svoji velkou výhodu párování na základě skaláru (propenzitního skóru) nevýhodu, jež spočívá ve skutečnosti, že tento skór je potřeba odhadnout. Toto vnáší do algoritmu párování „neurčitost“, která není plně pod kontrolou.

Všechny výsledky však, až na jednu výjimku,<sup>33</sup> směřují ke stejné interpretaci kauzálního efektu akce „Nakup nad“ na zisky.

**VO1** Hrubý zisk  $Y_1$  společnosti v období uplatnění kreditů vychází v průměru na účastníka akce na 142,1 Kč; kdyby se tito účastníci akce ve skutečnosti akce neúčastnili, pak by hrubý zisk v průměru na účastníka vycházel na 274,5 Kč (tedy o 1,1 Kč více než u skupiny těch, kdo se akce skutečně neúčastnili). Efekt účasti v akci „Nakup nad“ na hrubý zisk během období uplatnění kreditů vychází tedy v průměru na účastníka akce na  $-132,4$  Kč. **Tato ztráta je skutečně o něco větší než rozdíl mezi pozorovanou skupinou účastníků a neúčastníků ( $-131,2$  Kč). Odpovídá tedy očekávání popsanému v podkapitole 5.5.1.**

**VO2** Čistý zisk  $Y_2$  společnosti v období uplatnění kreditů vychází v průměru na účastníka akce na  $-21,9$  Kč. Kdyby se tito účastníci akce ve skutečnosti akce neúčastnili, pak by čistý zisk v průměru na účastníka vycházel na 255,5 Kč, což je o téměř 4 Kč méně než u skupiny těch, kdo se akce skutečně neúčastnili! Toto je překvapivý výsledek a neodpovídá očekávání z předchozí podkapitoly. Efekt účasti v akci „Nakup nad“ na čistý zisk během období uplatnění kreditů vychází tedy v průměru na účastníka akce jen na  $-277,5$  Kč (ve srovnání s  $-281,3$  Kč,

<sup>33</sup> U odhadu pomocí propenzitního skóru je efekt akce na  $Y_1$   $-129,5$  Kč, což je méně než  $-131,1$  (rozdíl mezi pozorovanou skupinou účastníků a neúčastníků). Ostatní metody párování odhadovaly naopak více.

což je rozdíl mezi pozorovanými skupinami účastníků a neúčastníků). **Tedy ztráta v čistém zisku je menší, než naznačovaly úvahy v podkapitole 5.5.1. Toto je pro společnost  $F$  obzvláště příjmný výsledek, protože obava, že skutečná ztráta je díky zkreslujícímu vlivu zavádějících proměnných větší než pozorovaná hodnota  $-281$  Kč, se ukázala jako lichá.** Možným vysvětlením je uplatňování kreditů z jiných akcí na podporu prodeje, než je akce „Nakup nad“. Tyto jiné kredity mohou uplatňovat všichni uživatelé, ale účastníci akce „Nakup nad“ jich mají zřejmě z dřívějšíka nastrádáno více.

**VO3** Čistý zisk  $Y_3$  společnosti v celém sledovaném období vychází v průměru na účastníka akce na 417,2 Kč. Kdyby se tito účastníci akce ve skutečnosti akce neúčastnili, pak by čistý zisk v celém sledovaném období vycházel v průměru na účastníka na 264,1 Kč, což je o 3 Kč méně než u skupiny těch, kdo se akce skutečně neúčastnili! Tento výsledek opět neodpovídá očekávání z předchozí podkapitoly. Celkový efekt účasti v akci „Nakup nad“ na čistý zisk tedy vychází v průměru na účastníka až na 153,1 Kč (ve srovnání se 150 Kč, což je rozdíl mezi pozorovanými skupinami účastníků a neúčastníků). **Čistý zisk z akce „Nakup nad“ je tedy dokonce větší a ne menší, než naznačovaly úvahy (a obavy společnosti) v podkapitole 5.5.1. Celkově tedy lze říci, že akce se společností  $F$  vyplácí a samotná účast v akci vydělala v průměru na účastníka navíc 153,1 Kč.**

Všechny uvedené částky vychází z tabulek 5.7 a 5.1.

Pro zodpovězení výzkumných otázek **VO4**, **VO5** a **VO6** byla použita metoda vážení propenzitními skóry (IPTW) bez osekávání a s osekáváním. Oba přístupy vedly k téměř stejným závěrům (odlišnosti jsou v jednotkách haléřů), takže je prakticky irelevantní, která metoda bude pro interpretaci upřednostněna. Níže uvedené hodnoty budou vycházet z tabulky 5.15 pro IPTW bez osekávání.

**VO4** Kdyby se nikdo z uživatelů neúčastnil akce „Nakup nad“, (což je ekvivalentní s tím, že by společnost nevydala žádné kredity), pak

by hrubý zisk  $Y_1$  společnosti v období uplatnění kreditů vycházel na 273,36 Kč (téměř se neliší od pozorované skupiny neúčastníků akce, u kterých je průměr  $Y_1$  roven hodnotě 273,34 Kč). Ovšem kdyby se naopak všichni této akce účastnili, pak by hrubý zisk společnosti v období uplatnění kreditů vycházel na 121,6 Kč (menší ve srovnání s pozorovanými účastníky akce, u kterých je průměr  $Y_1$  roven hodnotě 142,1 Kč). Efekt akce „Nakup nad“ na zisky  $Y_1$ , kdyby se jí hypoteticky všichni uživatelé účastnili, by byl  $-151,8$  Kč. Ztráta na celé populaci by tedy byla větší než ztráta na subpopulaci skutečných účastníků akce.

**VO5** Kdyby se nikdo z uživatelů neúčastnil akce „Nakup nad“, pak by čistý zisk  $Y_2$  společnosti v období uplatnění kreditů vycházel na 258,5 Kč (podobně vychází průměr u pozorované skupiny neúčastníků akce, který je 259,4 Kč). Ovšem kdyby se naopak všichni této akce účastnili, pak by čistý zisk společnosti v období uplatnění kreditů vycházel na  $-12,9$  Kč („větší“, byť stále záporný zisk ve srovnání s pozorovanými účastníky akce, u kterých je průměr  $Y_2$  roven hodnotě  $-21,9$  Kč). Efekt akce „Nakup nad“ na zisky  $Y_2$ , kdyby se jí hypoteticky všichni uživatelé účastnili, by byl  $-271,4$  Kč. Ztráta na celé populaci by tedy již byla menší než ztráta na subpopulaci skutečných účastníků akce.

**VO6** Kdyby se nikdo z uživatelů neúčastnil akce „Nakup nad“, pak by čistý zisk  $Y_3$  společnosti v celém sledovaném období vycházel na 266,2 Kč (podobně vychází průměr u pozorované skupiny neúčastníků akce, který je 267,2 Kč). Ovšem kdyby se naopak všichni uživatelé této akce účastnili, pak by čistý zisk společnosti v období uplatnění kreditů vycházel na 427,6 Kč (o 10 Kč větší zisk ve srovnání s pozorovanými účastníky akce, u kterých je průměr  $Y_3$  roven hodnotě 417,2 Kč). Efekt akce „Nakup nad“ na zisky  $Y_3$ , kdyby se jí hypoteticky všichni uživatelé účastnili, by byl 161,3 Kč. Celkový pozitivní efekt akce na celé populaci by tedy byl větší než na subpopulaci skutečných účastníků akce.

Z praktického pohledu je pro společnost  $F$  užitečné vědět, že kdyby se žádný z jejích uživatelů akce „Nakup nad“ neúčastnil, (tedy kdyby nevydala

kredity a nedošlo k uplatnění žádného kreditu), pak by její čistý zisk na uživatele v celém sledovaném období vycházel na 266,18 Kč. Skutečný pozorovaný průměrný zisk na uživatele v celém sledovaném období je 288,97 Kč. Rozdíl těchto hodnot dává společnosti jasnou představu o efektu účasti v akci na její čisté příjmy.

**Celkově lze na základě odpovědí na výzkumné otázky VO1 až VO6 konstatovat, že akce „Nakup nad“ se společnosti  $F$  vyplácí a to dokonce více, než naznačují pozorované rozdíly v ukazatelích ziskovosti mezi skupinami účastníků a neúčastníků akce.**

### 5.5.3 Omezení pro aplikaci použitých metod

Odpověď na výzkumné otázky VO1 až VO6 byla získána pomocí párovacích metod a IPTW metod kauzální analýzy. Tyto metody však mají svá omezení. Omezení aplikace použitých metod obecně i v kontextu řešeného problému efektivity akce „Nakup nad“ plynou

- 1) jednak ze samotného definování příslušnosti ke skupině s aktivním a kontrolním ošetřením,
- 2) dále z předpokladů použitých metod
- 3) a z nároků na kvalitu a rozsah dat i výpočetní možnosti použitého softwaru.

#### ad 1)

Charakter akce na podporu prodeje neumožňuje pomocí párovacích metod zjišťovat odpověď na otázku „Jaký by byl zisk, kdyby se akce „Nakup nad“ nekonala?“ (v celé populaci nebo subpopulaci účastníků akce). V rámci této akce se lze pouze ptát: „Jaký by byl zisk, kdyby se uživatelé akce neúčastnili?“ (v celé populaci nebo subpopulaci účastníků akce). Důvod je ten, že i kdyby se akce „Nakup nad“ nekonala, jednotliví uživatelé by i tak mohli uskutečnit nákupy nad 500 Kč.

Situace by mohla být příznivější, kdyby podmínky účasti v akci nebylo možné splnit jindy než v období konání akce (např. akce na podporu prodeje v kapitole 6). Vzhledem k charakteru akce „Nakup nad“ a k nákladům společnosti plynoucím z uplatněných kreditů je ovšem i zjednodušená

formulace otázky „Jaký by byl zisk, kdyby se uživatelé akce neúčastnili?“ prakticky relevantní.

### **ad 2)**

Pro korektní použití párovacích metod je potřeba, aby byly splněny předpoklady uvedené v 5.1.2. Zejména je obtížné vypořádat se s předpokladem nezavádějícího mechanismu přiřazování, jelikož splnění tohoto předpokladu nelze testovat. Lze jej pouze obhájit na základě znalosti kontextu řešeného problému, a je tedy téměř vždy možné jeho platnost rozporovat. Nelze totiž prokázat, že neexistuje žádná další zavádějící proměnná, která do modelu není zahrnuta.

Tato nevýhoda „neověřitelnosti“ předpokladů párovacích metod se nevztahuje pouze k řešené problematice akce „Nakup nad“, ale je obecná. Nikdy nelze prokázat, že všechny potřebné zavádějící proměnné jsou do modelu zahrnuty.

### **ad 3.)**

Algoritmy použitých metod vyžadují plnou vyplněnost datových matic, protože jen případy s údaji ke všem sledovaným proměnným mohou být do párovacího algoritmu zahrnuty. V kontextu akce „Nakup nad“ tento požadavek nepředstavuje problém, jelikož data pocházejí z on-line slevového portálu, kde jsou veškerá potřebná data automaticky zaznamenávána. V jiných situacích pro „děravé“ datové matice připadají v úvahu metody dopočítávání chybějících hodnot, pokud to míra vyplněnosti datové matice umožňuje.

Další omezení použití uvedených metod představuje malý rozsah datového souboru jak ve skupině s aktivním, tak ve skupině s kontrolním ošetřením. Obvykle je skupina s kontrolním ošetřením četnostně větší, kritický bývá rozsah datového souboru skupiny s aktivním ošetřením. Subjektům s aktivním ošetřením se hledají „dvojčata“ ve skupině s kontrolním ošetřením, a to tak, aby byla co nejpodobnější v hodnotách kovariát. Toto někdy není na celém definičním oboru kovariát možné splnit (s tím souvisí předpoklad překryvu), pak lze posuzovanou populaci zúžit na takovou subpopulaci, kde předpoklad překryvu platí. Potom ale lze interpretaci výsledků vztahovat jen na zúženou subpopulaci.



Naopak příliš velké datové soubory (řádově v milionech) mohou být výpočetně velmi náročné (zejména pro metody optimálního párování) a dostupné algoritmy mohou mít problém s konvergencí nebo jsou extrémně náročné na výpočetní čas. V tomto směru je velmi slibný nedávno vyvinutý algoritmus „near-optimal generalised full matching“ (Sävje, Higgins a Sekhon, 2017).

Nakonec může omezení představovat i dostupnost softwaru k párovacím procedurám. Jelikož jsou metody kauzální analýzy poměrně nové a výpočetně (ne ideově) náročné, nejsou implementovány v běžných pro firemní analytiku používaných softwarech. Na druhou stranu jazykové prostředí **R**, které disponuje mnoha balíky s různými párovacími a dalšími algoritmy, je zdarma a lze jej ve firemní praxi jednoduše implementovat. V zásadě stačí převést data z používané firemní databáze do potřebného datového formátu a disponovat výkonným serverem.

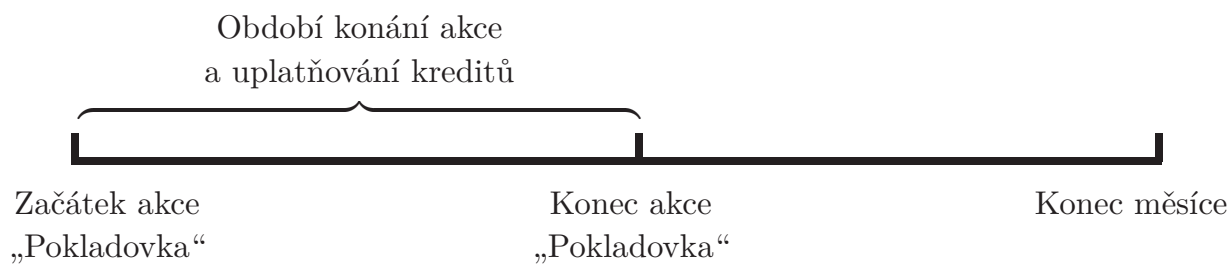


## Kapitola 6

# Aplikace metod odhadu kauzálního efektu na observačních datech s neregulárním mechanismem přiřazování

V této kapitole bude zjišťován **vliv účasti v akci** na podporu prodeje ve hře „Pokladovka“ – **na průměrné útraty** připadající na uživatele, a to jednak po dobu dvou týdnů trvání akce a dále i na útraty během celého měsíce, v němž akce probíhala. Podkapitola 6.1 se věnuje akci „Pokladovka“, která se konala v období od 1. 2. 2018 do 14. 2. 2018. V podkapitole 6.2 je popsána akce „Pokladovka“, která se konala v období od 1. 5. 2018 do 14. 5. 2018 (obrázek 6.1).

Uživatelé tím, že se zapojili do akce „Pokladovka“, získávali kredit. Ten mohli ve formě slevy uplatnit po dobu trvání akce na nákup zboží vybraných kategorií, přičemž pro každého hrajícího uživatele byla nabídka kategorií pro uplatnění kreditu personalizována s ohledem na jeho nákupní historii. Uplatněné kredity sice představují pro společnost  $F$  náklady (odpovídají částce, o kterou se společnosti snižují zisky z prodeje), na druhou stranu samotná akce „Pokladovka“ by měla vést ke zvýšení tržeb.



**Obrázek 6.1:** Časový průběh akce „Pokladovka“

Pro posouzení výhodnosti této akce (z pohledu společnosti  $F$ ) je tedy potřeba odhadnout, jaké by byly útraty uživatelů za sledovaný měsíc, kdyby se hra „Pokladovka“ nehrála. Pokud by účast/neúčast ve hře byla jednotlivým uživatelům přiřazena náhodně a uživatelé by byli „donuceni“ se tímto přiřazením řídit, pak by bylo možno efekt účasti v „Pokladovce“ na útraty uživatelů jednoduše v průměru odhadnout rozdílem útrat ve skupině hráčů a skupině nehráčů. Bohužel ale účast ve hře je volbou uživatele, a to v závislosti na jeho nepozorovatelných osobních vlastnostech, které mohou mít vliv jak na chuť zapojit se do hry, tak i na výši útrat. V důsledku tedy subpopulace hráčů a subpopulace nehráčů nejsou stejné s ohledem na nepozorovatelné osobní vlastnosti, a nerepresentují tedy stejnou populaci. Proto nemusejí být případné rozdíly v útratě mezi těmito dvěma skupinami způsobeny pouze vlivem účasti ve hře „Pokladovka“, ale také nepozorovatelnými zavádějícími proměnnými uživatelů.

Na rozdíl od akce „Nakup nad“ lze u „Pokladovky“ očekávat, že mezi zavádějícími proměnnými budou i nepozorovatelné osobnostní psychologické charakteristiky související s „hráčským naturelem“ uživatelů, jelikož tyto charakteristiky hra o poklad (tedy kredity a překvapení v podobě personalizované nabídky) využívá. Získání kreditů v akci „Nakup nad“ bylo naopak za „zásluhy“, tedy dostatečně velký nákup.

Jedná se zde o typický problém observačních studií s neregulárním mechanismem přiřazování, kdy zavádějící proměnné nejsou k dispozici. Pak lze efekt ošetření odhadnout například pomocí instrumentální proměnné, pokud je k dispozici.

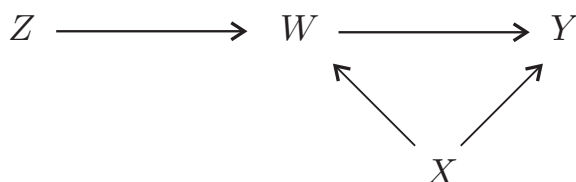
Výzkumné otázky, které se vážou k uvedeným problémům, budou rozděleny do dvou skupin podle toho, k jaké populaci se vztahují. Vliv akce na útraty bude zkoumán zvlášť na populaci těch únorových uživatelů služeb společnosti  $F$ , kteří „Pokladovku“ ještě nikdy nehráli, a zvlášť na populaci všech květnových uživatelů společnosti bez omezení dle jejich hráčské historie (tedy včetně těch, kdo s podobnou hrou již měli dřívější zkušenost). Vzhledem k tomu, že dle dlouhodobých zkušeností společnosti s podobnými akcemi účast ve hře zvedá útraty uživatelů i během krátkého období po ukončení akce, bude zkoumán vliv účasti v akci na útraty uživatelů i během celého měsíce, kdy se akce konala, nejen po dobu konání akce samotné.

Výzkumné otázky lze tedy formulovat následovně:

- VO7.** Jaký je skutečný efekt účasti v akci „Pokladovka“ na útraty uživatelů bez dřívější hráčské zkušenosti během období konání akce?
- VO8.** Jaký je skutečný efekt účasti v akci „Pokladovka“ na útraty uživatelů bez dřívější hráčské zkušenosti během celého měsíce v jehož první polovině se akce konala?
- VO9.** Jaký je skutečný efekt účasti v akci „Pokladovka“ na útraty uživatelů během období konání akce? (Populace uživatelů není omezena s ohledem na hráčskou historii.)
- VO10.** Jaký je skutečný efekt účasti v akci „Pokladovka“ na útraty uživatelů během celého měsíce, v jehož první polovině se akce konala? (Populace uživatelů není omezena s ohledem na hráčskou historii.)

Pro odhad efektu hry „Pokladovka“ na útraty uživatele bude využita instrumentální proměnná, kterou v této kapitole reprezentuje „povzbuzující email“. Všichni uživatelé (v rámci specifikované populace) byli o konání a pravidlech hry „Pokladovka“ v předstihu informováni ve svém profilu na stránkách společnosti  $F$  (informace o hře byla zveřejněna také v newsletteru společnosti a na Facebooku společnosti). Některým uživatelům byl ale navíc doručen personalizovaný email, který povzbuzoval k účasti ve hře. Výběr uživatelů pro zaslání povzbuzujícího emailu byl náhodný a pravděpodobnost obdržení emailu byla nastavena na 50 %.

Pro ilustraci vztahů mezi instrumentální proměnnou, ošetřením, odezvou a zavádějícími proměnnými při platných předpokladech slouží obrázek 6.2.



**Obrázek 6.2:** Vzhledem k současnému vlivu zavádějících proměnných  $X$  jak na ošetření  $W$  (účast v akci „Pokladovka“), tak na odezvu  $Y$  (útraty) nelze efekt ošetření na odezvu odhadovat pouze porovnáním odezev mezi skupinami s aktivním a kontrolním ošetřením. Proto se tento efekt odhadne pomocí instrumentální randomizované proměnné  $Z$  (zaslání mailu), která kovariátami  $X$  není ovlivněna (žádná šipka z  $X$  nevede do  $Z$ ) a na  $Y$  působí pouze prostřednictvím  $W$  (žádná šipka ze  $Z$  nevede přímo do  $Y$ ).

## 6.1 Vliv akce „Pokladovka“ na útraty uživatelů, kteří dosud podobnou hru nehráli

### 6.1.1 Popis datového souboru

Výzkumné otázky **VO7** a **VO8** budou řešeny v rámci populace těch uživatelů společnosti  $F$ , kteří byli aktivní (tedy během ledna 2018 – měsíce předcházejícího únorové akci „Pokladovka“ – byli aspoň jednou na stránkách společnosti  $F$ ), odebírají její newsletter a nikdy ve své historii „Pokladovku“ nehráli. Takových uživatelů bylo 464 605. Zkoumaný efekt ošetření lze na této populaci upřesnit jako efekt první zahrané „Pokladovky“.

Účast uživatele v akci na podporu prodeje „Pokladovka“ je v roli binárního ošetření  $W \in \{1, 0\}$ . Hodnota 1 (Pokladovka ano) znamená aktivní ošetření, tedy skutečnost, že si uživatel „vyzvedl poklad“. Znamená to, že otevřel truhlu ve svém profilu a našel personalizovanou nabídku i s kreditem, který mohl na nákup z pokladové nabídky během konání akce uplatnit. Hodnota 0 (Pokladovka ne) znamená kontrolní ošetření, tedy že si uživatel „poklad nevyzvedl“ a kredit nezískal.

Instrumentální proměnná  $Z \in \{1, 0\}$  je reprezentována doručením povzbuzujícího emailu. Hodnota 1 (mail ano) znamená, že uživateli byl email doručen a hodnota 0 (mail ne) znamená, že uživateli email doručen nebyl.

Separátně budou modelovány dvě odezvy  $Y_1$  a  $Y_2$ . Odezva  $Y_1 \in \langle 0, \infty \rangle$  představuje celkovou útratu uživatele v korunách po dobu konání hry (od 1. 2. 2018 do 14. 2. 2018) a odezva  $Y_2 \in \langle 0, \infty \rangle$  představuje celkovou útratu uživatele v korunách během celého února 2018. V každé z proměnných  $Y_1, Y_2$  mohou být kromě případné útraty za zboží z kategorie navázané na kredity i útraty za jiné kategorie zboží, jelikož smyslem akce je zvýšit celkové prodeje napříč všemi kategoriemi zboží. Důvod, proč je zajímavé modelovat kromě  $Y_1$  i  $Y_2$ , plyne ze zkušeností společnosti s dlouhodobě pozorovaným přesahem efektu podobných akcí i na krátké období po ukončení akce.

Absolutní a relativní četnosti uživatelů kategorizované podle toho, zda hráli/nehráli pokladovku a zda obdrželi/neobdrželi povzbuzující mail, jsou v kontingenčních tabulkách 6.1.

**Tabulka 6.1:** Kontingenční tabulky absolutních a relativních četností pro proměnné: ošetření  $W$  (účast ve hře s hodnotami „Pokladovka ne“ a „Pokladovka ano“) a instrumentální proměnnou  $Z$  (obdržení povzbuzujícího mailu s hodnotami „mail ne“ a „mail ano“)

<b>Absolutní četnosti uživatelů</b>			
	mail ne	mail ano	Sum
Pokladovka ne	231 291	228 466	459 757
Pokladovka ano	782	4 066	4 848
Sum	232 073	232 532	464 605

<b>Relativní četnosti uživatelů</b>			
	mail ne	mail ano	Sum
Pokladovka ne	0,497 8	0,491 7	0,989 6
Pokladovka ano	0,001 7	0,008 8	0,010 4
Sum	0,499 5	0,500 5	1,000 0

### 6.1.2 Předpoklady identifikace efektu ošetření pomocí instrumentální proměnné

Identifikovat vliv akce „Pokladovka“ na útraty uživatelů prostřednictvím instrumentální proměnné je možné pouze za splnění předpokladů ze sekce 4.1.3. Některé z těchto předpokladů lze posoudit statistickými nástroji, jiné nejsou testovatelné a je možné pouze obhájit je s ohledem na kontext aplikační úlohy.

#### 1. SUTVA

Tento předpoklad požaduje, aby potenciální odezvy každého uvažovaného subjektu nebyly ovlivněny ošetřením jakéhokoliv jiného subjektu. V kontextu akce „Pokladovka“ by tento předpoklad mohl být porušen, pokud by kategorie produktů, v rámci které může uvažovaný uživatel utratit vyhrané kredity, byla již vyčerpána dřívějšími účastníky hry. Tedy potenciální útrata (odezva) uvažovaného hráče by byla ovlivněna účastí ve hře (ošetření) jiných uživatelů. Dle zkušeností společnosti k tomuto nedochází, jelikož nabídka produktů je natolik bohatá, že její vyčerpání během hry (v rámci personalizovaných kategorií) je téměř nemožné. SUTVA dále předpokládá, že neexistuje více verzí pro jednotlivé úrovně ošetření. Toto je zaručeno

pravidly hry – uživatel se hry buď aktivně účastní a získá kredit, nebo se hry neúčastní. SUTVA musí dále platit i ve vztahu instrumentu (povzbuzující mail) k ošetření (účast ve hře). Zde ale porušení můžeme téměř vyloučit.

## 2. Randomizace instrumentu

Tento předpoklad je jednoznačně splněn, jelikož výběr těch uživatelů z výše definované populace dosavadních nehráčů, kteří obdrželi povzbuzující mail, byl proveden náhodně na základě generátoru náhodných čísel z alternativního rozložení s parametrem 0,5.

## 3. Monotonie

Předpoklad  $W_i(1) \geq W_i(0)$  zajišťuje, že v uvažované populaci nejsou „defiers“. Přítomnost „defiers“ by v kontextu řešené úlohy znamenala, že existuje uživatel, který mail nedostal, a právě proto hru „natruc“ zahrál, ačkoliv jinak by nehrál (to je těžko představitelné, a to tím spíše že takový uživatel ani neví, že patří do skupiny těch, pro něž je  $Z = 0$ ). A nebo, že existuje takový uživatel, který povzbuzující mail dostal, a právě proto „natruc“ hru nehrál. Vzhledem k tomu, že pracujeme s populací, která je definována jako populace aktivních uživatelů společnosti  $F$ , lze předpokládat, že uživatelé chtějí být informováni o různých nabídkách produktů a akcích společnosti, a „rebelská protireakce“ uživatelů je tedy krajně nepravděpodobná.

## 4. Restrikce na výlučnost

Tento velmi důležitý předpoklad požaduje, aby instrument (povzbuzující mail) nemohl ovlivnit odezvu (útraty) jinak než výlučně skrze ošetření (účast ve hře). Skutečnost, že uvažovaný instrument byl randomizovaný, ještě nezajišťuje, že nemůže na odezvu působit i přímo. (V kontextu 2SLS je pouze zajištěno, že náhodné chyby obou stupňů s instrumentem nesouvisejí). Marketéři společnosti  $F$  o přímém vlivu mailu na útraty pochybují a zdůvodňují to skutečností, že jejich aktivní uživatelé jsou vystaveni tak velkému množství jiných jejich reklamních sdělení (prostřednictvím mailů, FB, stránek s profilem uživatele apod.), že přímý vliv povzbuzujícího mailu na útraty lze vyloučit. (Resp. pokud by byl, pak by byl zcela marginální.)



## 5. Relevance

Předpoklad  $E[W_i(1) - W_i(0)] \neq 0$  zajišťuje, že instrument souvisí (ideálně silně) s ošetřením, tedy v kontextu akce „Pokladovka“ povzbuzující mail souvisí s účastí ve hře. Ve výše definované populaci dosavadních hráčů lze ovšem očekávat velký podíl „never-takers“, a tedy nepřekvapí, že instrument se ukáže jako slabý, což je zřejmé z následujících statistik:

Podíl hráčů mezi těmi, kdo dostali povzbuzující mail, je 0,017 49. To je sice 5,2 krát více než podíl hráčů mezi těmi, kdo povzbuzující mail nedostali, (těch je 0,003 37), přesto tento podíl považujeme za nízký. (Uvedené hodnoty odpovídají sloupcově podmíněným četnostem z tabulky 6.2.) Také korelace mezi proměnnou „povzbuzující“ mail a proměnnou „účast ve hře“ je malá:  $\text{cor}(Z, W) = 0,069\,458\,28$ . Tedy i když je předpoklad relevance splněn, instrument se jeví jako slabý.

Celkově lze považovat předpoklady za splněné (obhajitelné), a kauzální analýza pomocí instrumentální proměnné je tedy v uvedeném kontextu možná. S ohledem na slabý instrument nelze ovšem očekávat malou směrodatnou chybu odhadu efektu (i když rozsah souboru 464 605 je velký, jak je vidět z tabulky 6.1).

**Tabulka 6.2:** Kontingenční tabulka sloupcově podmíněných relativních četností: Mezi uživateli, kteří obdrželi povzbuzující mail, se hry účastnilo 1,7 %; mezi uživateli bez povzbuzujícího mailu se hry účastnilo pouze 0,33 %. Sloupcově podmíněná účast ve hře je tedy 5,19krát větší u těch uživatelů, kteří obdrželi povzbuzující mail.

Sloupcově podmíněné relativní četnosti uživatelů		
	mail ne	mail ano
Pokladovka ne	0,996 63	0,982 51
Pokladovka ano	0,003 37	0,017 49
Sum	1,000 00	1,000 00

### 6.1.3 Odhady efektu pomocí instrumentální proměnné

Východiskem pro odhad efektu účasti ve hře „Pokladovka“ na útraty uživatelů jsou údaje o celkových útratách uživatelů  $Y_1$  po dobu trvání akce

kategorizované podle účasti ve hře  $W$  a obdržení mailu  $Z$  (tabulka 6.3). Obdobně kategorizované údaje o celkových útratách uživatelů  $Y_2$  během celého měsíce jsou v tabulce 6.4. Obě tabulky jsou doplněné o tabulky počtů uživatelů, kteří uskutečnili nákup.

**Tabulka 6.3:** Tabulka nahoře udává celkové útraty uživatelů po dobu trvání akce kategorizovaně podle instrumentu (doručení mailu) a ošetření (účasti ve hře). Jelikož mnoho zákazníků během uvedeného období nic nekoupilo, je na doplnění uvedena i tabulka dole. Ta udává, kolik uživatelů po dobu trvání akce utratilo kladnou částku. Všech zákazníků je 464 605, těch, co nic během akce nenakoupili je 441 648 (= 464 605 – 22 957).

**Y1 : Celkové útraty všech uživatelů po dobu trvání akce**

	mail ne	mail ano	Sum
Pokladovka ne	11 303 069	11 203 168	22 506 237
Pokladovka ano	384 295	993 233	1 377 528
Sum	11 687 364	12 196 401	23 883 765

**Y1 : Počty uživatelů, kteří utráceli po dobu trvání akce**

	mail ne	mail ano	Sum
Pokladovka ne	10 819	10 519	21 338
Pokladovka ano	359	1 260	1 619
Sum	11 178	11 779	22 957

**Tabulka 6.4:** Tabulka nahoře udává celkové útraty uživatelů během celého měsíce kategorizovaně podle instrumentu (doručení mailu) a ošetření (účasti ve hře). Jelikož mnoho zákazníků během uvedeného období nic nekoupilo, je na doplnění uvedena i tabulka dole. Ta udává, kolik uživatelů během celého měsíce utratilo kladnou částku. Všech zákazníků je 464 605, těch, co nic během celého měsíce nenakoupili, je 424 180 (= 464 605 – 40 425).

**Y2 : Celkové útraty všech uživatelů během celého měsíce**

	mail ne	mail ano	Sum
Pokladovka ne	22 551 490	22 216 763	44 768 253
Pokladovka ano	533 458	1 436 343	1 969 801
Sum	23 084 948	23 653 106	46 738 054

**Y2 : Počty uživatelů, kteří utráceli během celého měsíce**

	mail ne	mail ano	Sum
Pokladovka ne	19 479	18 982	38 461
Pokladovka ano	421	1 543	1 964
Sum	19 900	20 525	40 425

Uživatelů, kteří nedostali povzbuzující mail, je 232 073, uživatelů kteří dostali povzbuzující mail, je 232 532 (tabulka 6.1). Tyto dvě skupiny jsou tedy čtenostně srovnatelné. Obdobně je vidět z tabulky 6.3, že není velký rozdíl mezi útratami  $Y_1$  těch, kdo mail dostali a kdo mail nedostali v rámci skupiny nehráčů. ( $11\,203\,168 \doteq 11\,303\,069$ ; dokonce ve skupině bez mailu, která je o něco menší, je o něco větší objem útrat než ve skupině s mailem.) Ovšem u hráčů jsou útraty zjevně větší ve skupině s mailem než ve skupině bez mailu ( $993\,233 \gg 384\,295$ ). (Analogické vztahy s obdobnou interpretací platí i pro počty „skutečně nakupujících“, tabulka 6.3 dole.) Zdá se tedy, že povzbuzující mail indukoval účast ve hře a zvedl útraty uživatelů.

Kdyby byla účast ve hře randomizovaná, pak by bylo jednoduché efekt účasti v akci „Pokladovka“ na útraty uživatelů po dobu konání akce odhadnout přímo z tabulky. Stačilo by porovnat průměrné útraty hráčů a nehráčů (jednalo by se o odhad pro  $E(Y | W = 1) - E(Y | W = 0)$ ). Dle údajů v tabulce 6.3 tento rozdíl vychází na  $\frac{1\,377\,528}{4\,848} - \frac{22\,506\,237}{459\,757} = 235,2$  Kč. Tuto částku ovšem nelze interpretovat jako kauzální efekt akce „Pokladovka“ na výši útrat, jelikož není jasné, zda „ochota“ hrát hru a „ochota“ utrácet není způsobená nějakými zavádějícími proměnnými, které nemáme pod kontrolou. Proto se nabízí odhadnout efekt „Pokladovky“ na útraty prostřednictvím instrumentální proměnné (povzbuzujícího mailu), která je již randomizovaná.

Nelze však ani odhadnout efekt hry na útraty uživatelů výlučně porovnáním průměrných útrat mezi těmi, kdo mail obdrželi, a těmi, kdo jej neobdrželi. (Jednalo by se o odhad pro  $E(Y | Z = 1) - E(Y | Z = 0) = ITT$ .) Důvodem je skutečnost, že ne všichni uživatelé jsou typu „compliers“. Mezi uživateli jsou také „always-takers“ (ti, co mail nedostali, ale hry se účastnili, a to ne z důvodu „rebelství“, jelikož přítomnost „defiers“ je v 6.1.2 vyloučena) a „never-takers“ (ti, co mail dostali, ale hry se neúčastnili, opět s vyloučením možnosti „defiers“). „Always-takers“ a „never-takers“ žádnou informaci o vlivu „Pokladovky“ na útraty nenesou.

Proto bude odhad efektu účasti ve hře „Pokladovka“ na útraty  $Y_1$  i  $Y_2$  odhadnut pomocí univerzální metody 2SLS. Vzhledem k tomu, že žádné další vysvětlující proměnné (ne zavádějící!) nejsou uvažovány, lze hledaný efekt odhadnout jednoduše také pomocí Waldovy statistiky odhadu či na základě rozkladu ITT .

**Waldův odhad efektu  $W$  na  $Y_1$** 

Waldův odhad efektu účasti ve hře (ošetření  $W$ ) na celkové útraty po dobu trvání akce (odezva  $Y_1$ ) prostřednictvím povzbuzujícího mailu (instrumentální proměnná  $Z$ ) lze podle A.1.7 získat z hodnot kovarianční matice v tabulce 6.5 jako podíl odhadnutých kovariancí

$$\hat{\tau}_{LATE} = \frac{\text{côv}(Y_1, Z)}{\text{côv}(W, Z)} = \frac{0,522\,424\,87}{0,003\,529\,038} = 148,036\,1$$

Účast ve hře tedy zvyšuje celkové útraty na jednotlivého uživatele o 148 Kč, avšak pouze v subpopulaci uživatelů typu „complier“. Ve srovnání s hodnotou 232,2 Kč, jež byla spočítána pouze jako porovnání průměrných útrat hráčů a nehráčů, je vidět, že odhad efektu účasti v akci „Pokladovka“ na výši útrat je v subpopulaci „compliers“ podstatně menší ( $148 < 232$ ). Zároveň ale stále není známo, jaký je efekt akce v rámci subpopulace „non-compliers“.

**Tabulka 6.5:** Kovarianční matice pro proměnné  $Y_1$ ,  $W$  a  $Z$

	$Z$	$W$	$Y_1$
$Z$	0,250 0	0,003 5	0,522 4
$W$	0,003 5	0,010 3	2,428 5
$Y_1$	0,522 4	2,428 5	175 178,488 0

**Odhad efektu  $W$  na  $Y_1$  pomocí 2SLS**

Pomocí 2SLS lze realizaci odhadu téhož efektu ošetření  $W$  (prostřednictvím  $Z$ ) na odezvu  $Y_1$  získat postupně ve dvou krocích:

- Nejdříve se v prvním stupni prostřednictvím instrumentu  $Z$  odhadnou hodnoty ošetření  $W$ .
- Tyto odhady  $\hat{W}$  následně ve druhém stupni slouží jako „náhrada“ původně pozorovaných hodnot ošetření  $W$  při modelování odezvy  $Y_1$ . (Tím se odstraní nežádoucí vliv zavádějících kovariát na ošetření  $W$ .)

Podrobně je postup popsán v příloze A.1.

**Tabulka 6.6:** Výstupy 2SLS modelu pro  $Y_1$ . Bodové odhady efektu ošetření ve sloupci „Odhad“ jsou v horní i dolní tabulce stejné (a totožné s Waldovým odhadem). Směrodatné chyby odhadu „SE“ efektu ošetření se ovšem v tabulkách liší. V horní tabulce je směrodatná chyba odhadu spočítána pouze z druhého stupně, a tudíž nepostihuje variabilitu odhadu ze stupně prvního. Přesnější odhad směrodatné chyby je pak v dolní tabulce, která zohledňuje variabilitu obou stupňů. Tento robustní odhad směrodatné chyby lze přímo získat použitím příkazů z balíku **AER**. Z praktického pohledu by bylo žádoucí, aby směrodatné chyby vycházely menší. To by ale musel být k dispozici silnější instrument.

	Odhad	SE	$t$ hodnota	$\Pr(>  t )$
(Posunutí)	49,861 9	1,096 0	45,50	0,000 0
$\hat{W}$	148,036 1	86,998 5	1,70	0,088 8

	Odhad	SE	$t$ hodnota	$\Pr(>  t )$
(Posunutí)	49,861 9	1,094 4	45,559 7	0,000 0
$W$	148,036 1	86,876 4	1,704 0	0,088 4

Číselná realizace odhadu efektu „účasti ve hře“ na „útratu po dobu trvání hry“ pomocí 2SLS vychází stejně jako v případě odhadu pomocí Waldovy statistiky a je rovna hodnotě 148,036 1 Kč, jak je vidět z tabulky 6.6.

Pro odhad efektu byl použit balík **AER** jazyka **R** (popis balíku je v příloze **B**).

### Odhad efektu $W$ na $Y_1$ pomocí rozkladu **ITT**

Z aplikačního hlediska není příjemné, že odhadnutý efekt účasti v akci „Pokladovka“ na útraty nelze zobecnit na celou populaci všech uživatelů, ale jen na subpopulaci „compliers“. Při tradičním využití 2SLS se tato skutečnost v interpretacích nezohledňovala. Proto bude pro názornost efekt účasti v akci na útraty odhadnut ještě jednou a to pomocí **ITT** tak, aby bylo zřejmé, že odhadovaný efekt se týká pouze subpopulace „compliers“ v rámci populace únorových uživatelů bez dřívější hráčské zkušenosti.

Odhad efektu ošetření  $W$  (prostřednictvím  $Z$ ) na odezvu  $Y$  pomocí **ITT** je popsán v 4.1.7, tedy  $\tau_{LATE} = \frac{ITT}{P_c}$ . Bude tedy potřeba zjistit efekt **ITT** (průměrný rozdíl v útratách mezi těmi, kdo mail dostali, a těmi, kdo mail nedostali) a podíl „compliers“ (podíl těch, kdo účastí/neúčastí ve hře reagovali v souladu s tím, zda dostali/nedostali povzbuzující mail).

Odhad efektu  $ITT = E(Y | Z = 1) - E(Y | Z = 0)$  získáme z hodnot útrat uživatelů v tabulce 6.1 průměrovaných absolutními četnostmi uživatelů v tabulce 6.3:

$$I\hat{T}T = \frac{12\,196\,401}{232\,532} - \frac{11\,687\,364}{232\,073} = 2,089\,697 \text{ Kč.}$$

Právě odhadnutý efekt dvou korun je ale nezajímavý. Motivem k rozeslání „povzbuzujících“ mailů nebylo zjistit dopad mailů skrze účast v akci na útraty, ale jejich prostřednictvím zjistit dopad akce samotné na útraty. Efekt akce „Pokladovka“ (v modelu indukovaný mailem) na výši útrat lze ovšem pomocí instrumentu měřit pouze na subpopulaci „compliers“.

Podíl „compliers“ lze získat z pozorovaných četností v tabulce 6.2 postupně. Nejdříve je potřeba získat odhad podílu „never-takers“ a „always-takers“ a následně odečtením jejich podílů od celku zůstane podíl uživatelů typu „compliers“.

„Never-takers“ jsou ti, kteří se akce neúčastní nikdy, dokonce ani když dostanou povzbuzující mail. Jejich podíl lze díky randomizaci  $Z$  vyjádřit jako  $P_n = P(W_i = 0 | Z = 1)$  a odhad pak je  $\hat{P}_n = 0,982\,51$ .

„Always-takers“ jsou ti, kteří se akce účastní vždy, dokonce i když nedostanou povzbuzující mail. Jejich podíl lze vyjádřit jako  $P_a = P(W_i = 1 | Z = 0)$  a odhad pak je  $\hat{P}_a = 0,003\,37$ .

„Compliers“ jsou ti, kdo reagují v souladu s podnětem (doručením/nedoručení mailu) svou účastí/neúčastí v akci. Jejich podíl lze vyjádřit jako  $P_c = 1 - P_a - P_n$  a odhad pak je  $\hat{P}_c = 1 - 0,982\,51 - 0,003\,37 = 0,014\,12$

Tedy pouze 1,4 % uživatelů společnosti (bez dřívějších hráčských zkušeností) lze považovat za „compliers“ a pouze k nim se vztahuje odhad efektu akce „Pokladovka“ na útraty  $\hat{\tau}_{LATE}$ :

$$\hat{\tau}_{LATE} = \frac{2,089\,697}{0,014\,12} = 148 \text{ Kč.}$$

148 korun je v rámci subpopulace „compliers“ kauzální efekt účasti ve hře „Pokladovka“ na výši útrat po dobu konání akce. Ve srovnání s velmi odlišnou hodnotou 235 Kč (rozdíl v útratách mezi „hráči“ a „nehráči“) je zřejmé, že alespoň v rámci subpopulace „compliers“ randomizovaná instrumentální proměnná pomohla identifikovat skutečný efekt akce na podporu prodeje „Pokladovka“.

### Odhad efektu $W$ na $Y_2$

Výzkumná otázka **VO2** se týká vlivu akce na podporu prodeje „Pokladovka“ ( $W$ ) na útraty uživatelů bez předchozí hráčské zkušenosti **během celého měsíce**, kdy se „Pokladovka“ hrála ( $Y_2$ ). Instrumentální proměnná  $Z$  reprezentována povzbuzujícím mailem i ošetření  $W$  – účast v akci „Pokladovka“ – zůstávají stejné jako v předchozí podkapitole. Pouze odezva  $Y_2$ , tedy útraty uživatele během celého měsíce, se od předchozí podkapitoly liší. Proto předpoklady i metody řešení zůstávají stejné a stručně budou uvedeny jen výstupy metod, jež vedou k odhadu efektu účasti ve hře „Pokladovka“ na útraty  $Y_2$ .

Pro orientační seznámení se se zkoumanou závislostí bude nejdříve uveden rozdíl mezi celkovou měsíční útratou hráčů a nehráčů (v rámci zkoumané populace). Ten vychází v průměru na jednoho účastníka na základě hodnot z tabulek 6.1 a 6.4

$$\frac{1\,969\,801}{4\,848} - \frac{44\,768\,253}{459\,757} = 308,94 \text{ Kč ve prospěch hráčů.}$$

Těchto téměř 309 Kč opět nelze kvůli nepozorovaným zavádějícím proměnným interpretovat jako kauzální vliv účasti ve hře „Pokladovka“ na celkové útraty během celého měsíce. Proto bude kauzální efekt odhadnut pouze na populaci „compliers“ – jako v případě celkových útrat po dobu konání hry.

Waldův odhad efektu ošetření  $W$  („účast ve hře“) prostřednictvím instrumentální proměnné  $Z$  („povzbuzující mail“) na odezvu  $Y_2$  („celkové útraty během celého měsíce“) lze dopočítat z hodnot kovarianční matice v tabulce 6.7 jako podíl odhadnutých kovariancí

$$\hat{\tau}_{LATE} = \frac{\text{côv}(Y_2, Z)}{\text{côv}(W, Z)} = \frac{0,561\,751\,292}{0,003\,529\,038} = 159,179\,7.$$

**Tabulka 6.7:** Kovarianční matice pro proměnné  $Y_2$ ,  $W$  a  $Z$

	$Z$	$W$	$Y_2$
$Z$	0,250 0	0,003 5	0,561 7
$W$	0,003 5	0,010 3	3,190 0
$Y_2$	0,561 7	3,190 0	374 192,512 1

Účast ve hře tedy zvyšuje celkové měsíční útraty připadající na jednotlivého uživatele ze subpopulace „compliers“ o 159 Kč. Na stejnou subpopulaci „compliers“, kteří ve své historii „Pokladovku“ nehráli, se vztahuje i 148korunový efekt účasti ve hře na útraty pouze během konání akce.

Z toho vyplývá, že kdyby se nikdo ze skupiny uživatelů typu „compliers“ akce „Pokladovka“ neúčastnil, pak by v průměru společnost  $F$  vydělala na jednoho z nich o 159 Kč měsíčně méně.

Stejný odhad efektu účasti ve hře „Pokladovka“ na měsíční útraty lze získat i pomocí 2SLS, jak ve vidět z tabulky 6.8, doplněné o směrodatné chyby odhadu. Interpretačně zajímavý je také odhad  $ITT$  a z něj následně odvozený odhad efektu.  $\hat{ITT} = 2,247\,003$  Kč, podíl „compliers“ zůstává stejný jako při odhadu efektu  $W$  na  $Y_2$  a tedy  $\hat{\tau}_{LATE} = \frac{2,247\,003}{0,014\,116\,14} = 159,179\,7$ .

**Tabulka 6.8:** 2SLS model pro  $Y_2$ . Bodový odhad efektu ošetření je totožný s Waldovým odhadem. Směrodatná chyba „SE“ je již odhadnuta robustně.

	Odhad	SE	$t$ hodnota	$\Pr(>  t )$
(Posunutí)	98,936 4	1,600 2	61,828 3	0,000 0
$W$	159,179 7	127,023 0	1,253 2	0,210 1

## 6.2 Vliv akce „Pokladovka“ na útraty uživatelů z populace bez omezení dle hráčské historie

### 6.2.1 Popis datového souboru

Výzkumné otázky **VO9** a **VO10** budou řešeny v rámci populace těch uživatelů společnosti  $F$ , kteří byli aktivní (t.j. byli aspoň jednou na stránkách společnosti  $F$ ) během dubna 2018 a odebírají newsletter společnosti  $F$ . V této populaci mohou být i hráči dřívějších akcí typu „Pokladovka“. Všech uživatelů takto definované populace bylo 484 425. V této kapitole bude zkoumán vliv účasti v květnové „Pokladovce“ na útraty uživatelů buď pouze během období konání akce, nebo během celého měsíce.

Proměnné účast v akci „Pokladovka“ (ošetření  $W$ ) a doručení povzbuzujícího mailu (instrumentální proměnná  $Z$ ) zůstávají definovány stejně jako v předchozí kapitole. Jediný rozdíl spočívá v tom, že náhodně generované adresy pro zaslání povzbuzujících mailů se generují v rámci celé



populace uživatelů (a ne s omezením na dosavadní „nehráče“ jako v předchozí podkapitole). Odezvy  $Y_1$  a  $Y_2$  odpovídají útratám během konání květnové akce „Pokladovka“ a útratám během celého měsíce května, jinak jsou definovány také stejně jako v předchozí podkapitole.

Absolutní a relativní četnosti uživatelů kategorizované podle toho, zda hráli, či nehráli květnovou „Pokladovku“, a zda obdrželi, či neobdrželi povzbuzující mail, jsou v kontingenčních tabulkách 6.9.

**Tabulka 6.9:** Kontingenční tabulky absolutních a relativních četností pro proměnné: ošetření  $W$  (účast ve hře) a instrumentální proměnná  $Z$  (obdržení povzbuzujícího mailu) v rámci populace bez omezení

Absolutní četnosti uživatelů			
	mail ne	mail ano	Sum
Pokladovka ne	238 358	233 312	471 670
Pokladovka ano	3 851	8 904	12 755
Sum	242 209	242 216	484 425

Relativní četnosti uživatelů			
	mail ne	mail ano	Sum
Pokladovka ne	0,492 0	0,481 6	0,973 7
Pokladovka ano	0,007 9	0,018 4	0,026 3
Sum	0,500 0	0,500 0	1,000 0

### 6.2.2 Předpoklady identifikace efektu ošetření pomocí instrumentální proměnné

Vzhledem k tomu, že v předchozí podkapitole byla populace uživatelů omezena na ty, kteří „Pokladovku“ nikdy nehráli, nebylo překvapující, že instrumentální proměnná byla slabá. V této podkapitole není populace uživatelů nijak jejich hráčskou minulostí omezena, lze tedy očekávat, že instrument bude silnější.

Posouzení předpokladů pro identifikaci efektu ošetření pomocí instrumentální proměnné zůstává obdobné jako v podkapitole 6.1.2, pouze pátý **předpoklad relevance instrumentu** je potřeba posoudit znovu. Tento předpoklad požaduje, aby  $E[W_i(1) - W_i(0)] \neq 0$ .

Střední hodnoty  $E[W_i(1)]$  a  $E[W_i(0)]$  lze odhadnout prostřednictvím podmíněných četností z tabulky 6.10. Podíl hráčů mezi těmi, kdo dostali povzbuzující mail, je 0,036 76, což je sice 2,3krát více než podíl hráčů mezi těmi, kdo povzbuzující mail nedostali, těch je 0,015 90, ale hodnoty jsou stále velmi nízké. Také hodnota korelačního koeficientu mezi proměnnou „povzbuzující mail“ a proměnnou „účast ve hře“ je malá:  $\text{cor}(Z, W) = 0,065\ 143\ 95$ . Tato instrumentální proměnná je tedy opět slabá.

**Tabulka 6.10:** Kontingenční tabulka sloupcově podmíněných relativních četností: Mezi uživateli, již obdrželi povzbuzující mail, se hry účastnilo 3,67 %; mezi uživateli bez povzbuzujícího mailu se hry účastnilo pouze 1,59 %. Sloupcově podmíněná účast ve hře je tedy 2,31krát větší u těch uživatelů, kteří obdrželi povzbuzující mail. Hodnoty se vztahují k populaci bez omezení.

Sloupcově podmíněné relativní četnosti uživatelů		
	mail ne	mail ano
Pokladovka ne	0,984 10	0,963 24
Pokladovka ano	0,015 90	0,036 76
Sum	1,000 00	1,000 00

### 6.2.3 Odhady efektu pomocí instrumentální proměnné

Pro posouzení výhodnosti květnové akce „Pokladovka“ jsou v tabulce 6.11 údaje o útratách uživatelů během konání akce samotné (vztahují se k odezvě  $Y_1$ ) a v tabulce 6.12 jsou údaje o útratách během celého května (vztahují se k odezvě  $Y_2$ ). Kromě informací o celkových útratách jsou tabulky doplněny i údaji o počtu skutečně nakupujících.

Uživatelů, již nedostali povzbuzující mail, je 242 209, uživatelů, již dostali povzbuzující mail, je 242 216. Tyto dvě skupiny jsou tedy četnostně srovnatelné. Z tabulky 6.11 je dále vidět, že v rámci skupiny nehráčů také není téměř žádný rozdíl mezi útratami (po dobu konání akce) těch, kdo mail dostali a nedostali ( $11\ 452\ 837 \doteq 11\ 108\ 377$ ). Ovšem ve skupině hráčů lze pozorovat zřetelný rozdíl mezi objemem útrat těch, kdo byli mailem povzbuzeni, a těmi bez mailu ( $1\ 909\ 63 \gg 957\ 528$ ). Analogické vztahy s obdobnou interpretací platí i pro počty skutečně nakupujících.

**Tabulka 6.11:** Tabulka nahoře udává celkové útraty uživatelů po dobu trvání akce kategorizovaně podle instrumentu a ošetření. Jelikož mnoho zákazníků během uvedeného období nic nekoupilo, je na doplnění uvedena i tabulka dole. Ta udává, kolik uživatelů po dobu trvání akce utratilo kladnou částku. Všech zákazníků je 484 425, tedy těch, co nic během akce nenakoupili, je 463 112 (= 484 425 – 21 313). Údaje se vztahují k populaci bez omezení.

**Y1 : Celkové útraty všech uživatelů po dobu trvání akce**

	mail ne	mail ano	Sum
Pokladovka ne	11 452 837	11 108 377	22 561 214
Pokladovka ano	957 528	1 909 632	2 867 160
Sum	12 410 365	13 018 009	25 428 374

**Y1 : Počty uživatelů, kteří utráceli v době trvání akce**

	mail ne	mail ano	Sum
Pokladovka ne	9 255	8 611	17 866
Pokladovka ano	1 147	2 300	3 447
Sum	10 402	10 911	21 313

**Tabulka 6.12:** Tabulka nahoře udává celkové útraty uživatelů během celého měsíce kategorizovaně podle instrumentu a ošetření. Jelikož mnoho zákazníků během uvedeného období nic nekoupilo, je na doplnění uvedena i tabulka dole. Ta udává, kolik uživatelů během celého měsíce utratilo kladnou částku. Všech zákazníků je 484 425, tedy těch, co nic během celého měsíce nenakoupili, je 441 783 (= 484 425 – 42 642). Údaje se vztahují k populaci bez omezení.

**Y2 : Celkové útraty všech uživatelů za celý měsíc**

	mail ne	mail ano	Sum
Pokladovka ne	25 931 311	24 920 722	50 852 033
Pokladovka ano	1 658 311	3 454 467	5 112 778
Sum	27 589 622	28 375 189	55 964 811

**Y2 : Počty uživatelů, kteří utráceli během celého měsíce**

	mail ne	mail ano	Sum
Pokladovka ne	19 516	18 363	37 879
Pokladovka ano	1 554	3 209	4 763
Sum	21 070	21 572	42 642

Orientačně lze efekt akce na útraty **po dobu konání** akce posoudit porovnáním průměrné útraty hráčů a nehráčů po dobu konání akce, což na základě údajů v tabulkách 6.11 a 6.9 vychází:

$$\frac{2\,867\,160}{12\,755} - \frac{22\,561\,214}{471\,670} = 176,954\,5 \text{ Kč ve prospěch hráčů.}$$

Tento rozdíl by ovšem bylo chybné interpretovat jako odhad efektu akce „Pokladovka“ na výši útrat, jelikož účast ve hře nebyla randomizovaná a možné zavádějící proměnné mohly způsobit vychýlení tohoto odhadu. Jelikož tyto zavádějící proměnné nejsou k dispozici, bude odhad efektu akce získán prostřednictvím randomizovaného instrumentu ve formě povzbuzujícího mailu.

Dále pro hrubou představu o efektu akce na útraty **během celého měsíce** slouží následující porovnání průměrných měsíčních útrat hráčů a nehráčů (potřebné údaje jsou v tabulkách 6.12 a 6.9):

$$\frac{5\,112\,778}{12\,755} - \frac{50\,852\,033}{471\,670} = 293,032\,3 \text{ Kč ve prospěch hráčů.}$$

Také hodnota 293 Kč je vychýlena kvůli nepozorovatelným zavádějícím proměnným a bude revidována prostřednictvím stejného instrumentu.

### Waldův odhad efektu $W$ na $Y_1$ a $Y_2$

Waldův odhad efektu účasti ve hře (ošetření  $W$ ) na útraty po dobu trvání akce (odezva  $Y_1$ ) i na útraty během celého měsíce (odezva  $Y_2$ ) prostřednictvím povzbuzujícího mailu (instrumentální proměnná  $Z$ ) je získán na základě odhadnutých kovariancí<sup>34</sup> v tabulce 6.13.

$$\hat{\tau}_{LATE, Y_1} = \frac{\hat{c}ov(Y_1, Z)}{\hat{c}ov(W, Z)} = \frac{0,626\,802\,715}{0,005\,215\,282} = 120,185\,8$$

$$\hat{\tau}_{LATE, Y_2} = \frac{\hat{c}ov(Y_2, Z)}{\hat{c}ov(W, Z)} = \frac{0,809\,991\,1}{0,005\,215\,282} = 155,311\,1$$

Účast ve hře tedy zvyšuje **útraty během konání akce** u uživatele ze subpopulace „compliers“ o 120 Kč a **útraty během celého měsíce** až o 155 Kč.

<sup>34</sup> Tabulka zaokrouhluje hodnoty na čtyři desetinná místa, výpočty jsou ale co nejpřesnější kvůli srovnání s odhady pomocí jiných metod (2SLS a rozklad ITT).

**Tabulka 6.13:** Vlevo je kovarianční matice pro proměnné  $Z$ ,  $W$  a  $Y_1$ ; vpravo je kovarianční matice pro proměnné  $Z$ ,  $W$  a  $Y_2$ .

	$Z$	$W$	$Y_1$		$Z$	$W$	$Y_2$
$Z$	0,250 0	0,005 2	0,626 8	$Z$	0,250 0	0,005 2	0,810 0
$W$	0,005 2	0,025 6	4,536 6	$W$	0,005 2	0,025 6	7,512 5
$Y_1$	0,626 8	4,536 6	266 645,505 4	$Y_2$	0,810 0	7,512 5	574 934,610 5

### Odhad efektu $W$ na $Y_1$ a $Y_2$ pomocí 2SLS

Odhady efektu akce „Pokladovka“ na útraty  $Y_1$  a  $Y_2$  získané Waldovou metodou se neliší od odhadů získaných pomocí 2SLS. Tyto odhady efektů jsou ve výstupech 2SLS modelů (tabulka 6.14) doplněny navíc směrodatnými chybami odhadu. Bohužel tyto chyby jsou dost velké vinou slabého instrumentu a v případě zobecňování na superpopulaci by vedly k velmi širokým intervalům spolehlivosti.

**Tabulka 6.14:** Výstupy 2SLS modelu nahoře pro  $Y_1$ , dole pro  $Y_2$  již s adjustovaným sandwichovým odhadem směrodatné chyby „SE“. Bodové odhady efektu ošetření ve sloupci „Odhad“ jsou totožné s Waldovými odhady.

model pro $Y_1$	Odhad	SE	$t$ hodnota	$\Pr(>  t )$
(Posunutí)	49,327 4	2,011 7	24,520 0	0,000 0
$W$	120,185 8	71,033 0	1,692 0	0,090 7
model pro $Y_2$	Odhad	SE	$t$ hodnota	$\Pr(>  t )$
(Posunutí)	111,439 0	2,953 6	37,730 2	0,000 0
$W$	155,311 1	104,289 5	1,489 2	0,136 4

### Odhad efektu $W$ na $Y_1$ a $Y_2$ pomocí rozkladu $ITT$

Odhady efektu květnové „Pokladovky“ získané pomocí Waldova odhadu či pomocí 2SLS se vztahují pouze k subpopulaci „compliers“, tedy k subpopulaci uživatelů, již jsou ovlivnitelní doručením/nedoručením povzbuzujícího mailu. Z aplikačního hlediska by bylo užitečné vědět aspoň to, jak je tato subpopulace velká. Jistou informační hodnotu má zároveň i hodnota  $ITT$  efektu.

*ITT* efekt odhadujeme jako průměrný rozdíl v útratách  $Y_1$  a  $Y_2$  mezi těmi, kdo mail dostali, a těmi, kdo mail nedostali. (Níže uvedené odhady *ITT* efektu lze získat z tabulek 6.9, 6.11 a 6.12.)

$$\begin{aligned} \hat{ITT}_{Y_1} &= \frac{13\,018\,009}{242\,216} - \frac{12\,410\,365}{242\,209} = 2,507\,206 \text{ Kč} \\ \hat{ITT}_{Y_2} &= \frac{28\,375\,189}{242\,216} - \frac{27\,589\,622}{242\,209} = 3,239\,958 \text{ Kč} \end{aligned}$$

Pro zjištění podílu „compliers“ v populaci je potřeba znát i podíly „never-takers“ a „always-takers“ – a ty lze odhadnout pomocí četností v tabulce 6.10:

Odhad podílu „never-takers“ je  $\hat{P}_n = 0,963\,239\,4$

Odhad podílu „always-takers“ je  $\hat{P}_a = 0,0158\,994\,9$

Odhad podílu „compliers“ je  $\hat{P}_c = 1 - 0,963\,239\,4 - 0,0158\,994\,9 = 0,020\,861\,09$

V rámci populace uživatelů společnosti  $F$  (bez omezení dle hráčské historie) tedy pouze 2 % reagují zapojením/nezapojením se do hry „Pokladovka“ v souladu s doručením/nedoručením povzbuzujícího mailu. Odhady efektu „Pokladovky“ na útraty  $Y_1$  a  $Y_2$  v rámci subpopulace „compliers“ pak pomocí *ITT* vycházejí:

$$\begin{aligned} \hat{\tau}_{LATE,Y_1} &= \frac{\hat{ITT}_{Y_1}}{P_c} = \frac{2,507\,206}{0,020\,861\,09} = 120,185\,7 \text{ Kč} \\ \hat{\tau}_{LATE,Y_2} &= \frac{\hat{ITT}_{Y_2}}{P_c} = \frac{3,239\,958}{0,020\,861\,09} = 155,311\,1 \text{ Kč} \end{aligned}$$

Interpretace výsledků je v diskuzi 6.3.

### 6.3 Interpretace odhadů, diskuze a omezení

Smyslem akce na podporu prodeje „hra Pokladovka“ je zvýšit zisky společnosti. Součástí akce je i rozdávání a následné uplatňování kreditů, které ovšem pro společnost představují náklady. V akci „Nakup nad“ se kredity rozdávaly za „zásluhy“, tedy za dostatečně velký nákup, který již sám o sobě generoval společnosti zisk. V akci „Pokladovka“ se kredity „rozdávají“ pouze s vidinou budoucího zisku indukovaného až uplatněním kreditů. Proto je tím spíše žádoucí správně odhadnout skutečný efekt akce

„Pokladovka“ na útraty, aby s ohledem na hodnotu kreditů bylo jasné, zda se akce společnosti vyplácí.

Přímo spočítané rozdíly v průměrných útratách mezi hráči a nehráči „Pokladovky“ jsou sice kladné ve prospěch hráčů, ale není jasné, zda ti, co hru hráli, by v průměru i bez hry neutratili více než nehráči. Tedy rozdíl v průměrných útratách může být způsoben i něčím jiným než hrou samotnou. Tento skrytý efekt zavádějících proměnných nelze ošetřit metodami z kapitoly 5, protože tyto zavádějící proměnné jsou nepozorovatelné. Proto byla použita instrumentální proměnná v podobě „povzbuzujícího mailu“, jejímž prostřednictvím lze skutečný kauzální efekt akce „Pokladovka“ na útraty zjistit.

### 6.3.1 Interpretace odhadů a diskuze

#### Populace bez dřívější zkušenosti s hrou „Pokladovka“

Výzkumné otázky VO7 a VO8 se vztahovaly k populaci takových uživatelů společnosti, kteří nikdy v historii „Pokladovku“ nehráli. Zjišťovaný vliv akce na útraty v rámci této populace tak lze chápat jako efekt „první zahrané Pokladovky“.

Použitý instrument se ukázal slabým, pouze 1,4 % populace bylo typu „complier“ (uživatel, který účastí/neúčastí ve hře reagoval v souladu s tím zda dostal/nedostal povzbuzující mail). Celkově bylo skutečných hráčů velmi málo ve srovnání s nehráči (4848 hráčů, 459757 nehráčů), takže je zřejmě obtížné najít lepší instrument, který by povzbuzoval ke hře silněji (a neměl přímý vliv na útraty) než instrument použitý.

Vzhledem k tomu, že výsledky není potřeba zobecňovat na superpopulaci, ale týkají se přímo populace samotné (výběr je totožný s populací), větší směrodatná chyba odhadu (způsobená slabým instrumentem) výrazně nevadí.

Odhadnutý kauzální efekt „Pokladovky“ na útraty během doby konání akce je menší než orientační odhad získaný pouze porovnáním průměrných útrat hráčů a nehráčů (148 Kč proti orientačním 235 Kč). Potvrdilo se tedy, že vyšší útraty hráčů jsou výrazně ovlivněny dalšími nepozorovanými zavádějícími proměnnými, ovšem tento odhad se týká pouze hráčů typu „compliers“, ne všech uživatelů zkoumané populace. Ač je obecně problematické při použití instrumentálních proměnných dělat úsudky o celé populaci

pouze na základě výsledků získaných pro subpopulaci „compliers“, v kontextu řešeného problému efektivity akcí na podporu prodeje to představuje jen částečný problém.

Náklady společnosti vznikají jen u skutečných hráčů – těch je 4 848. A zejména u hráčů je pro společnost důležité vědět, kolik z jejich útrat lze přisoudit hře samotné. (Srovnáním s cenou kreditů pak společnost může posoudit výhodnost akce.) Stačí tedy dopočítat, kolik uživatelů z 4 848 hráčů je typu „complier“, a pro tuto skupinu pak budou platit odhadnuté efekty hry na útraty  $Y_1$  a  $Y_2$ .

Mezi 4 848 hráči jsou jednak ti, kdo dostali povzbuzující mail a jsou typu „complier“, pak ti co dostali mail a jsou typu „always-taker“, a nakonec ti, kdo sice nedostali mail, ale jsou také typu „always-taker“. Efekty jsou odhadnuté dobře jen pro první skupinu z nich, tedy ty, kdo dostali mail a jsou typu „complier“ a jejich počet je potřeba určit. Z důvodu přehlednosti budou ve výpočtech použity symboly.

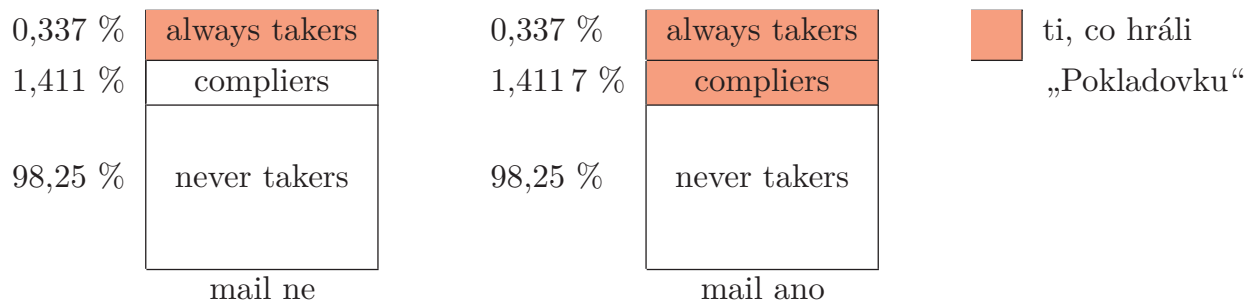
$n_1 = 232\,532$  označuje počet uživatelů, kteří dostali mail;  $n_2 = 232\,073$  označuje počet uživatelů, již mail nedostali;  $\hat{P}_n = 0,982\,514\,2$  je odhadnutý podíl „never-takers“ jak ve skupině s mailem, tak ve skupině bez mailu;  $\hat{P}_a = 0,003\,369\,629$  je odhadnutý podíl „always-takers“ jak ve skupině s mailem, tak ve skupině bez mailu a  $\hat{P}_c = 1 - \hat{P}_a - \hat{P}_n = 0,014\,116\,14$  je odhadnutý podíl „compliers“, také v obou skupinách. Počet všech hráčů lze tedy vyjádřit jako  $n_1 \cdot \hat{P}_c + n_1 \cdot \hat{P}_a + n_2 \cdot \hat{P}_a = 4\,848$ .

Všech „compliers“ je  $n_1 \cdot \hat{P}_c + n_2 \cdot \hat{P}_c = 6\,558$ . Průměrný efekt akce na útraty odpovídá právě této subpopulaci se 6 558 uživateli, kteří se liší pouze tím, zda dostali, nebo nedostali povzbuzující mail. S ohledem na nepozorovatelné zavádějící proměnné se ovšem díky randomizaci instrumentu neliší. Odhadnutý kauzální efekt tedy platí i pro skupinu hráčů (typu „complier“) a jejich počet je  $n_1 \cdot \hat{P}_c = 3\,282$ .

Z 4 848 hráčů tedy skutečný efekt hry na útraty lze určit u 3 282 hráčů. U zbylých 1 566 hráčů stále není jasné, jaká část jejich útrat je vyvolána hrou samotnou. Jedná se o hráče „always-takers“, kteří zřejmě rádi hrají, a pokud jsou náchylnost ke hře a k utrácení svázané pozitivně, pak by efekt hry na jejich útraty byl ještě menší než u „compliers“.

Přehledně jsou výsledky uvedených úvah shrnuty do tabulky 6.15 a na obrázku 6.3.





**Obrázek 6.3:** Schematické rozložení subpopulací „always-takers“, „never-takers“ a „compliers“ v rámci skupiny s doručeným povzbuzujícím mailem a bez mailu. Podíly odpovídají populaci uživatelů bez dřívějších zkušeností s hrou „Pokladovka“.

**Tabulka 6.15:** Odhady kauzálních efektů pomocí instrumentální proměnné platí pro subpopulaci „compliers“.

Populace bez dřívější zkušenosti s hrou „Pokladovka“				
Odezva	Rozdíl průměrných útrat hráčů a nehráčů	Odhadnutý kauzální efekt hry pro „compliers“	Počet hráčů	Počet hráčů typu „compliers“
$Y_1$	235,20	148,04	4 848	3 282
$Y_2$	308,94	159,18	4 848	3 282

Z tabulky 6.15 je zřejmé, že **kauzální vliv účasti ve hře** „Pokladovka“ na útraty (ať už během období hry samotné, nebo během celého měsíce) v rámci subpopulace účastníků, kteří jsou ovlivnitelní povzbuzujícím mailem („compliers“), **je pozitivní**. V případě útrat během období konání hry ( $Y_1$ ) je tento efekt 148 Kč, v případě útrat během celého měsíce je tento odhad 159 Kč. Tyto výsledky lze interpretovat tak, že kdyby nikdo ze subpopulace účastníků, kteří jsou ovlivnitelní povzbuzujícím mailem, tuto hru nehrál (resp. kdyby se tato akce na podporu prodeje nekonala), utratili by tito uživatelé o 148 Kč (resp. 159 Kč) v průměru na osobu méně.

Kauzální interpretaci naopak neumožňovaly přímo spočítané rozdíly v útratách mezi hráči a nehráči. Pro útraty během hry samotné tento rozdíl vycházel 235 Kč a pro útraty během celého měsíce pak rozdíl vycházel 309 Kč. Tyto rozdíly ovšem **nadhodnocovaly** skutečný kauzální efekt účasti ve hře, jelikož neměly pod kontrolou zavádějící proměnné. Odhady

pomocí instrumentu „povzbuzující mail“ vliv zavádějících proměnných naopak odstranily.

Správně odhaduté efekty hry lze také přímo uplatnit na počtu 3 282 hráčů, kteří jsou typu „complier“. Díky této znalosti může společnost přesněji zvažovat výhodnost akce na podporu prodeje „Pokladovka“.

### Populace bez omezení dle hráčské historie

Výzkumné otázky **VO9** a **VO10** se vztahovaly k populaci všech uživatelů společnosti, bez ohledu na jejich hráčskou historii.

I na této populaci, i když zahrnovala dřívější hráče, se použitý instrument ukázal jako slabý, pouze 2,1 % populace bylo typu „complier“. Ovšem i v této populaci bylo skutečných hráčů velmi málo ve srovnání s nehráči (12 755 hráčů, 471 670 nehráčů).

Celkové výsledky analogické úvahám pro populaci bez dřívější hráčské zkušenosti jsou shrnuty v tabulce 6.16.

**Tabulka 6.16:** Odhady kauzálních efektů pomocí instrumentální proměnné platí pro subpopulaci „compliers“.

Populace bez omezení dle hráčské historie				
Odezva	Rozdíl průměrných útrat hráčů a nehráčů	Odhadnutý kauzální efekt hry pro „compliers“	Počet hráčů	Počet hráčů typu „compliers“
$Y_1$	176,95	120,18	12 755	5 053
$Y_2$	293,03	155,31	12 755	5 053

Z tabulky je vidět, že i pro populaci uživatelů bez omezení dle hráčské historie platí, že rozdíly průměrných útrat  $Y_1$  či  $Y_2$  mezi hráči a nehráči nadhodnocovaly efekt akce „Pokladovka“ na útraty. Nicméně i po korekci na zavádějící proměnné prostřednictvím instrumentální proměnné zůstal tento efekt pozitivní. V případě útrat během období konání akce je pozitivní vliv „Pokladovky“ na útraty 120 Kč, pozitivní vliv na útraty během celého měsíce je až 155 Kč. Stále tedy „Pokladovka“ zvyšuje útraty, ovšem toto lze tvrdit jen pro uživatele typu „complier“, z nichž 5 053 skutečně hru hrálo a vygenerovalo náklady v podobě kreditů.

Celkově lze konstatovat, že výzkumné otázky **VO7** až **VO10** byly zodpovězeny, ovšem jen s omezením na subpopulaci uživatelů, kteří jsou

ovlivnitelní povzbuzujícím mailem („compliers“). V rámci této subpopulace pomohla randomizovaná instrumentální proměnná identifikovat skutečný efekt akce na podporu prodeje „Pokladovka“, a tak přispěla ke zpřesnění měření její efektivity. **Akce „Pokladovka“ je v této subpopulaci pro společnost  $F$  přínosná, jelikož efekt účasti ve hře na výši útrat je pozitivní (v rámci subpopulace „compliers“). Nicméně nárůst útrat v důsledku účasti ve hře je menší, než by bylo možné očekávat pouze na základě porovnání útrat skupiny hráčů a nehráčů.**

### 6.3.2 Omezení pro aplikaci metody instrumentální proměnné

Aplikace metod instrumentální proměnné je limitována mnoha aspekty, které jsou uvedeny v následujících odstavcích.

- Omezení plynoucí z předpokladů v 6.1.2:  
Zásadním omezením metody instrumentální proměnné je samotná volba instrumentální proměnné. Její hodnota musí být jednotlivým subjektům přiřazena náhodně (nebo pseudonáhodně), musí mít vliv na ošetření a pouze skrze ošetření smí ovlivňovat závislou proměnnou a dále nesmí být ovlivněna zavádějícími kovariáty. Obecně je nalezení proměnné splňující všechny tyto podmínky velmi obtížné a pro použití metody zcela limitující. V konkrétním případě akce „Pokladovka“ jsou tyto podmínky splněny a zdůvodnění je v podkapitolách 6.1.2 a 6.2.2.
- Omezení plynoucí ze slabého instrumentu:  
Při slabém instrumentu je pouze malá část populace typu „compliers“, tedy jen malý podíl případů v datovém souboru nese informaci o odhadovaném kauzálním efektu. Důsledkem je pak velký rozptyl odhadů efektu, a tedy jejich nestabilita. V případě zkoumaného vlivu akce „Pokladovka“ se použitý instrument ukázal jako slabý. Velký rozptyl odhadu efektů nepředstavuje v tomto případě problém proto, že výběrový soubor je totožný s populací. Odhadnutý efekt tedy není potřeba zobecňovat. Obecně ale je právě zobecnění odhadu efektu z výběrového souboru „compliers“ na celou populaci „compliers“ při slabém instrumentu problematické.

- Omezení plynoucí z redukce populace na subpopulaci „compliers“: Skutečnost, že odhadnutý kauzální efekt lze pomocí instrumentálních proměnných vztáhnout pouze na subpopulaci „compliers“, je největší nevýhodou této metody. Tyto subjekty navíc nelze přímo identifikovat, pouze lze odhadnout jejich celkový podíl v celé populaci. Proto ani nelze popsat, jaké další vlastnosti (z případných dalších sledovaných proměnných) subjekty typu „compliers“ mají. Ovšem při silném instrumentu je i podíl „compliers“ v populaci velký. V kontextu akce „Pokladovka“ je pro společnost  $F$  zajímavá pouze subpopulace hráčů, protože pouze oni představují pro společnost náklady v podobě uplatněných kreditů. V podkapitole 6.3.1 je ukázáno, že z hráčů je velká část typu „compliers“, tedy nárůst útrat velké části hráčů lze v průměru pomocí metody instrumentálních proměnných odhadnout.
- Požadavky na software:  
Metoda instrumentálních proměnných neklade velké nároky na softwarové vybavení. V okleštěné podobě při porozumění metodě lze použít jakýkoliv software, který má implementovanou regresní analýzu (tedy stačí i Excel). Pro kvalitní odhad variability odhadu efektů je ovšem potřeba sofistikovanější software. Zde je opět výhodné použít jazykové prostředí **R** ze stejných důvodů jako v podkapitole 5.5.3.

## Kapitola 7

# Závěr

Metody kauzální analýzy umožňují řešit mnohé problémy z oblasti marketingu a managementu, přesto se v těchto oblastech téměř vůbec nevyužívají, a to jak v českém, tak i ve světovém kontextu. V českém kontextu navíc nejsou dostatečně popularizovány ani samotné metody kauzální analýzy, a tak veškeré rešerše v této publikaci jsou téměř výhradně na základě zahraničních literárních pramenů.

Jedním z cílů této publikace bylo prozkoumat možnosti aplikací metod kauzální analýzy v oblasti podpory prodeje. V publikaci jsou tyto metody představeny, je navržena česká terminologie tam, kde není etablována, a zejména je demonstrováno použití těchto metod v oblasti měření efektivity podpory prodeje. Publikace tak vyplňuje mezeru v instruktážní literatuře, jak prakticky odhadovat kauzální efekty. Obvykle je v tomto typu literatury uveden aforismus „Korelace neimplikuje kauzalitu“, následně je za účelem zjištění efektu doporučeno použití randomizovaného experimentu a nakonec výklad mnoha statistických a ekonometrických metod začíná: „Za předpokladu, že  $X$  má vliv na  $Y$  platí. . . “ Jenže v mnoha aplikačních oblastech včetně marketingu a managementu data z randomizovaných experimentů k dispozici nejsou a čtenář se nedozví, že i v případě dat z observačních studií (tento charakter má většina dat ve firemní praxi) lze kauzální efekty odhadovat a jaké metody jsou k dispozici. Publikace tedy může sloužit jako doplnění běžné instruktážní literatury o metody kauzální analýzy s příklady použití v oblasti marketingu.

Dalším cílem publikace bylo přispět k řešení problému spojeného s odhadováním základních tržeb při akcích na podporu prodeje, a odhadování efektu akcí na podpory prodeje na sledované ukazatele ziskovosti. Tradičně používané odhady základních tržeb (tedy tržeb, jaké by byly, kdyby se akce na podporu prodeje nekonala) jsou velmi zkreslené zejména v důsledku mnoha v čase se měnících vlivů, které běžně používané odhady nemají pod

kontrolou. Základní tržby (v období konání akce nebo v období následném) se obvykle predikují na základě tržeb v období před konáním akce a to pomocí velmi primitivních i velmi sofistikovaných prediktivních modelů. Žádné z těchto modelů ale nemohou zachytit vlivy, které se „dějí“ v době konání akce na podporu prodeje a mohou ukazatele ziskovosti ovlivnit také. Navíc přesnější (dynamické lineární) modely jsou natolik sofistikované, že je obtížné je v běžné firemní praxi uplatnit.

V publikaci je představen nový přístup k odhadování základních tržeb a následně odhadování efektu akcí na podporu prodeje na různé ukazatele ziskovosti. Odhady základních tržeb nejsou predikovány pomocí „minulých“ tržeb předcházejících období konání akce. Místo toho je hodnota základních tržeb predikována z dat získaných ve stejném čase, kdy se akce na podporu prodeje konala. Na všechny subjekty, ať se již akce zúčastnily, nebo ne, působí vnější vlivy (chování konkurence, ekonomická situace regionu a pod.) stejně. Porovnáním těchto dvou skupin subjektů, s kontrolou možných zavádějících proměnných (!) lze tedy efekt akcí odhadovat bez vychýlení z důvodu v čase proměnlivých vlivů, jež tradičně používané metody nemají pod kontrolou.

Další výhodou tohoto přístupu je jeho snadná implementace ve firemní praxi. V intuitivní rovině jsou navržené metody velmi srozumitelné a nevyžadují žádné speciální statistické vzdělání. Jakýkoliv pracovník analytického oddělení či pracovník, který má na starosti firemní databáze, může tyto metody poměrně jednoduše implementovat. Předpokladem je vhodný softwarový nástroj. V 5. a 6. kapitole této publikace byl pro odhady efektů konkrétních akcí na podporu prodeje společnosti  $F$  využíván jazyk **R**, který je zdarma i pro firemní využití.

Posouzení přesnosti navržených odhadů efektů akcí na podporu prodeje pomocí metod kauzální analýzy (i ve srovnání s přesností odhadů pomocí tradičních metod) není na observačních datech možné, protože nikdo neví, jaká je objektivní pravda o tom, jaké by byly tržby, kdyby se akce na podporu prodeje nekonala. Porovnat odhady získané tradičním přístupem (pre-post analýzy) a navržené odhady pomocí kauzální analýzy lze (v případě dostupnosti všech údajů, které oba přístupy na vstupu požadují), ovšem při jejich odlišnosti nelze statisticky rozhodnout, který odhad je lepší.

---

Konkrétní příklady použití navržených odhadů jsou demonstrovány na datech společnosti  $F$ , která provozuje slevový portál. V kapitole 5 je odhadován vliv konkrétní akce na podporu prodeje (akce „Nakup nad“) na zisky připadající na jednoho zákazníka (uživatele slevového portálu), přičemž v datové matici byly k dispozici takové proměnné, které umožňují využití metod kauzální analýzy s tzv. regulárním mechanismem přiřazování. Součástí odhadu vlivu akce na zisky je i odhad „základních zisků“ (v kontextu řešeného problému reprezentují analogii základních tržeb) pomocí vybraných metod kauzální analýzy. Z pohledu společnosti  $F$  byl zajímavý zejména odhad efektu akce na zisky v období po ukončení akce. V tomto období jsou zisky nižší jednak z důvodu „předzásobení“ v době konání akce, jednak v důsledku uplatňování kreditů (nákladů společnosti) získaných v době akce. Dále bylo potřeba posoudit výhodnost akce v celém sledovaném období (pokrývajícím jak dobu konání akce, tak následné období). Pouhé porovnání zisků skupin účastníků a neúčastníků akce by vedlo k vychýleným odhadům efektu akce z důvodu zavádějících proměnných. Po aplikování zvolených kauzálních metod se prokázalo, že ztráta v období po ukončení akce je menší, než se jevila z prostého porovnání skupin a celkové zisky společnosti za celé sledované období byly dokonce v důsledku akce na podporu prodeje větší, než se jevílo z porovnání skupin účastníků a neúčastníků. Výhodnost pořádání akce na podporu prodeje „Nakup nad“ (i s náklady na uplatněné kredity) byla prokázána.

V kapitole 6 (akce „Pokladovka“) je odhadován vliv takové akce na podporu prodeje na útraty připadající na jednoho zákazníka, při které není možné sledovat zavádějící proměnné. Nelze tedy použít metody s regulárním mechanismem přiřazování, a proto publikace jako řešení prezentuje možnosti využití instrumentálních proměnných v kontextu kauzální analýzy (interpretace výsledků je jiná než u tradičního ekonometrického přístupu k instrumentálním proměnným). Pozitivní vliv akce na velikost útrat byl prokázán (ačkoliv podstatně menší, než se jevil na základě posouzení útrat skupiny hráčů a nehráčů), ovšem pouze na subpopulaci tzv. „compliers“. Tato metoda neumí odhadnout efekt akce na celé populaci uživatelů společnosti  $F$ , což je její hlavní nedostatek. Přesto přináší alespoň částečné poznání.

Metody kauzální analýzy se v poslední dekádě bouřlivě vyvíjejí, přesto je stále hodně toho, co zatím řešit neumějí. Zejména zatím není k dispozici odbornou komunitou akceptovaný přístup, který by odhadoval kauzální efekt spojitého ošetření. V této publikaci bylo (napříč všemi kapitolami) prezentováno pouze binární ošetření. Problém nepředstavuje ani víceúrovňové kvalitativní ošetření, jelikož většinu metod kauzální analýzy lze pro více úrovní ošetření přizpůsobit. Pokud má ovšem ošetření spojitý charakter, současným, často nedostačujícím řešením, je pouze jeho kategorizace.

Další výzvou jsou velká firemní data, a to nejen ve smyslu velkého počtu případů (zde je problémem výpočetní náročnost a nároky na kapacity serverů), ale také ve smyslu velkého počtu proměnných (tradiční statistické metody očekávají podstatně větší počet případů než proměnných, což u velkých společností často neplatí). Zde je potřeba propojit metody kauzální analýzy se strojovým učením a metodami pro volbu proměnných. Opačný problém představuje naopak nedostatek dat, kdy hodnoty potřebných proměnných nejsou k dispozici.



# Summary

The methods of causal analysis make it possible to deal with a number of issues in the fields of marketing and management, yet they are rarely used for these purposes, both in the Czech and global context. In the Czech context, moreover, the methods are not sufficiently popularized and that is why this publication is almost exclusively based on foreign literary sources.

One of the goals of this publication was to explore the possibilities of applying the methods of causal analysis in the area of sales support. To meet this goal, the publication introduces these methods, proposes Czech terminology as it yet needs to be established, and most importantly, demonstrates the use of the methods in measuring the effectiveness of sales support. The publication thus tries to fill a gap in the instructional literature on how to practically estimate causal effects. In the available literature, the aphorism “correlation does not imply causation” is usually mentioned. This is fairly often accompanied by a recommendation of a randomized experiment in order to determine the respective effect and followed by a standard introduction to a number of statistical and econometric methods which tend to begin in the following way: “Assuming that  $X$  has an effect on  $Y$  the following is true. . . .” However, in many application areas, including marketing and management, data from randomized experiments are not available and the reader of this literature will sadly not learn that even with data gained from observational studies (this is the nature of most data in business practice), causal effects can be estimated and there are methods available to do so. Having the methods of causal analysis and examples of use in marketing as its main focus, this publication presents a highly relevant addition to the current instructional literature.

Another goal of this publication was to address the issues associated with estimating baseline sales when these are supported by sales promotion events, and to deal with the issues of estimating the effect of sales promotion events on the monitored profitability indicators. The traditionally used estimates of baseline sales (i.e. the level of sales that would have been

achieved in the absence of the sales promotion events) are highly biased, mainly due to the many time-varying effects that commonly used estimators do not control. Baseline sales (achieved in the period of sales promotion or in a subsequent period) are usually predicted on the basis of sales in the period before the event, using both very primitive and very sophisticated predictive models. However, none of these models can capture the influencing factors active at the time of the sales promotion event which also can affect the profitability indicators. In addition, the sophistication of more accurate (dynamic linear) models makes them difficult to apply in everyday business practice.

This publication thus presents a new approach to estimating baseline sales and subsequently, estimating the effect of sales promotion events on various profitability indicators. Here, the estimators of baseline sales are not predicted using the “past” sales, i.e. preceding the sales promotion event. Instead, the value of baseline sales is predicted from data obtained at the time when the sales promotion event was taking place. All subjects, whether they have participated in the event or not, are equally affected by external influences (competitive behaviour, economic situation in the region, etc.). By comparing these two groups of subjects while controlling for the possible confounding variables (!), the effect of events can be estimated without bias due to time-varying influences, which the traditionally used methods do not control.

Another advantage of this approach is its easy implementation in business practice. On an intuitive level, the proposed methods are very clear and do not require any special statistical training. Any employee of the analytics department or an employee running the company’s database can implement them relatively easily via a appropriate software tool. In chapters 5 and 6, the **R** language, which is available free of charge for general company use, was employed to estimate the effects of specific sales promotion events of the company  $F$ .

Sadly, the assessment of the accuracy of the proposed estimators of the effects of sales promotion events while employing the methods of causal analysis (even when compared to the accuracy of estimators using traditional methods) is not possible when using observational data. This is because it is not known what level the sales would have reached if the sales

---

promotion event had not taken place. Although it is possible to compare the estimates obtained by the traditional approach (pre-post analysis) and the estimates proposed via the causal analysis (if all the input data required by both approaches are available), it is still not possible to statistically determine which estimator is better should the estimates differ.

Specific examples of the use of the proposed estimators are demonstrated on the data of company  $F$ , which operates a discount portal. Chapter 5 estimates the impact of a specific sales promotion (“Purchase Over”) event on profits per customer (i.e. a user of the discount portal), where the data matrix included variables that allow the use of methods of causal analysis with the so-called regular assignment mechanism. A component of the estimator of the effect of the event on profits is the estimator of “baseline profits” (in the context of the addressed problem they represent the analogy to baseline sales) using the selected methods of causal analysis. From the company  $F$ ’s point of view, the estimate of the effect of the event on profits in the period following the event was particularly interesting. During this period, profits are lower, both due to “pre-supply” at the time of the event and due to the application of credits (company costs) gained at the time of the event. Furthermore, it was necessary to assess the benefits of the event in the entire period observed (covering both the duration of the event and the subsequent period). Merely comparing the profits of groups participating and non-participating in the event would, due to confounding variables, lead to biased estimators of the effect of the event. After applying the chosen causal methods, it was shown that the loss in the period following the event is smaller than it appeared to be from a simple comparison of the groups, and that the company’s total profits for the whole period were, as a result of the sales promotion event, greater than what would the mere comparison between participants and non-participants suggest. The advantage of organizing an event to support the sale of “Purchase Over” (even with the costs associated with applied credits) has been proven.

In chapter 6 (“Treasury event”), a sales promotion event effect on single customer spending when confounding variables are not accessible, is estimated. As the methods with the regular assignment mechanism cannot be used, the publication therefore presents as a solution the possibility of

using instrumental variables in the context of causal analysis (interpretation of results is different from the one used in the traditional econometric approach to instrumental variables). The positive effect of the event on the amount of customer spending was proved (although it was significantly smaller than it appeared to be on the basis of the assessment of the money spent by the group of players and non-players), but only on the subpopulation of the so-called “compliers”. Thus, this method cannot estimate the effect of the event on the entire population of company  $F$ 's users, which is its main drawback. Nevertheless, it does generate valuable partial findings.

The methods of causal analysis have been rapidly evolving over the last decade, yet there is still much they cannot address. In particular, an approach approved by the professional community that could be used to estimate the causal effect of continuous treatment is not yet available. In this publication, only binary treatment was presented. Multi-level qualitative treatment would be viable as well, as most causal analysis methods can be adapted for multiple levels of treatment. However, if the treatment has a continuous character, the current, often insufficient approach, can only offer its categorization.

Another challenge is presented by the big company data, not only in terms of a high number of cases (there is a problem with computational complexity and server capacity requirements), but also in terms of a large number of variables (traditional statistical methods expect a significantly greater number of cases than variables, which is not often true with large companies). In this case, it is necessary to connect the methods of causal analysis with machine learning and with the methods for the selection of variables. The opposite problem, on the other hand, is the lack of data, where the values of the required variables are not available.

# Literatura

- ABADIE, A. a G. W. Imbens. (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, 74(1), 235–267.  
<https://doi.org/10.1111/j.1468-0262.2006.00655.x>
- ABRAHAM, M. M. a L. M. Lodish. (1993). An Implemented System for Improving Promotion Productivity Using Store Scanner Data. *Marketing Science*, 12(3), 248–269.  
<https://doi.org/10.1287/mksc.12.3.248>
- AILAWADI, K., Harlam, B. A., Cesar, J., a D. Trounce. (2006). Promotion Profitability for a Retailer: The Role of Promotion, Brand, Category, and Store Characteristics. *Journal of Marketing Research*, 43(4), 518–535.  
<https://doi.org/10.1509/jmkr.43.4.518>
- AILAWADI, K. a S. Gupta. (2014). Sales Promotions. In R. S. Winer and S. A. Neslin (Eds), *History of Marketing Science* (pp. 463–497). Singapore.  
[https://doi.org/10.1142/9789814596480\\_0017](https://doi.org/10.1142/9789814596480_0017)
- ANGRIST, J. D., Imbens, G. W. a D. B. Rubin. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434), 444–455.  
<https://doi.org/10.1080/01621459.1996.10476902>
- ANGRIST, J. D. a J. S. Pischke. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton University Press.  
<https://doi.org/10.1515/9781400829828>
- ARONOW, P. M. a A. Carnegie. (2013). Beyond LATE: Estimation of the Average Treatment Effect with an Instrumental Variable. *Political*

- Analysis*, 21(4), 492–506.  
<https://doi.org/10.1093/pan/mpt013>
- ATAMAN, B., Mela, C. F. a H. Van Heerde. Building Brands. (2008).  
*Marketing Science*, 27(6), 1036–1054.  
<https://doi.org/10.1287/mksc.1080.0358>
- BELLONI, A., Chernozhukov, V. a Ch. Hansen. (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, 28(2), 29–50.  
<https://doi.org/10.1257/jep.28.2.29>
- BLATTBERG, R. C. a R. Briesch. (2012). Sales Promotions. In Ö. Özalp & R. Phillips (Eds.), *The oxford handbook of pricing management* (pp. 585–619). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199543175.013.0024>
- BLATTBERG, R. C., Kim, B.-D. a J. Ye. (1996). Defining Baseline Sales in a Competitive Environment. *Seoul Journal of Business*, 2(1), 1–38.
- BOSTON CONSULTING GROUP. (2014). Making big data work: Retailing. URL [https://www.bcgperspectives.com/content/articles/retail\\_pricing\\_making\\_big\\_data\\_work\\_retailing/](https://www.bcgperspectives.com/content/articles/retail_pricing_making_big_data_work_retailing/).
- BOSTON CONSULTING GROUP. (2015). How retailers can improve promotion effectiveness. URL <https://www.bcg.com/publications/2015/retail-pricing-how-retailers-can-improve-promotion-effectiveness.aspx>.
- BOSTON CONSULTING GROUP. (2018). Improving promotional effectiveness. URL <https://www.bcg.com/expertise/industries/retail/improving-promotional-effectiveness.aspx>.
- BOUND, J., Jaeger, D. A., a R. B. Baker. (1995). Problems with Instrumental Variable Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association*, 90, 443–450.  
<https://doi.org/10.1080/01621459.1995.10476536>

- 
- BREUER, P., Elliott, B. a S. Rickert. (2013). *The power of advanced analytics in revenue management*. McKinsey&Company. URL [https://www.mckinsey.com/~media/mckinsey/dotcom/client\\_service/retail/articles/perspectives%20-%20winter%202013/power\\_of\\_advanced\\_analytics\\_in\\_revenue\\_management\\_periscope%20vf.ashx](https://www.mckinsey.com/~media/mckinsey/dotcom/client_service/retail/articles/perspectives%20-%20winter%202013/power_of_advanced_analytics_in_revenue_management_periscope%20vf.ashx).
- BROWN, K., Merrigan, P., a J. Royer. (2018). Estimating Average Treatment Effects With Propensity Scores Estimated With Four Machine Learning Procedures: Simulation Results in High Dimensional Settings and With Time to Event Outcomes. *Econometrics: Econometric & Statistical Methods – General eJournal*. <https://doi.org/10.2139/ssrn.3272396>
- BRYNJOLFSSON, E., Hitt, L. M., a H. H. Kim. (2011). *Strength in numbers: How does data-driven decisionmaking affect firm performance?* <https://doi.org/10.2139/ssrn.1819486>
- BUDÍKOVÁ, M., Králová, M. a B. Maroš. (2010). *Průvodce základními statistickými metodami*. Grada Publishing, Praha.
- COOK, T. D. (2008). ‘Waiting for Life to Arrive’: A History of the Regression Discontinuity Design in Psychology, Statistics and Economics. *Journal of Econometrics*, 142(2), 636–654. <https://doi.org/10.1016/j.jeconom.2007.05.002>
- DATTA, H., Knox, G. a B. J. Bronnenberg. (2018). Changing Their Tune: How Consumers’ Adoption of Online Streaming Affects Music Consumption and Discovery. *Marketing Science*, 37(1), 5–21. <https://doi.org/10.1287/mksc.2017.1051>
- DAVENPORT, T. H., Harris, J. G., a R. Morison. (2010). *Analytics at Work: Smarter Decisions, Better Results*. Harvard Business Press.
- DIAMOND, A. a J. S. Sekhon. (2013). Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *The Review of Economics and Statistics*, 95(3), 932–945. [https://doi.org/10.1162/REST\\_a\\_00318](https://doi.org/10.1162/REST_a_00318)

- FRÖLICH, M. (2007). Nonparametric IV Estimation of Local Average Treatment Effects with Covariates. *Journal of Econometrics*, 139, 35–75.  
<https://doi.org/10.1016/j.jeconom.2006.06.004>
- GEDENK, K., Neslin, S. A. a K. L. Ailawadi. (2010). Sales Promotion. *Retailing in the 21st century*. Springer, Berlin Heidelberg, 393–407.  
[https://doi.org/10.1007/978-3-540-72003-4\\_24](https://doi.org/10.1007/978-3-540-72003-4_24)
- GÓMEZ, M. I., Rao, V. R. a E. W. McLaughlin. (2007). Empirical Analysis of Budget and Allocation of Trade Promotions in the US Supermarket Industry. *Journal of Marketing Research*, 44(3), 410–424.  
<https://doi.org/10.1509/jmkr.44.3.410>
- GREENLAND, S., Pearl, J. a J. M. Robins. (1999). Causal Diagrams for Epidemiologic Research. *Epidemiology*, 10, 37–48.  
<https://doi.org/10.1097/00001648-199901000-00008>
- GU, X. a P. R. Rosenbaum. (1993). Comparison of Multivariate Matching Methods: Structures, Distances and Algorithms. *Journal of Computational and Graphical Statistics*, 2(4), 405–420.  
<https://doi.org/10.1080/10618600.1993.10474623>
- HAHN, J., Todd, P. a W. Van Der Klaauw. (2001). Identification and Estimation of Treatment Effects with a Regression Discontinuity Design. *Econometrica*, 69(1), 201–209.  
<https://doi.org/10.1111/1468-0262.00183>
- HAMILTON, B. H. a J. A. Nickerson. (2003). Correcting for Endogeneity in Strategic Management Research. *Strategic Organization*, 1(1), 51–78.  
<https://doi.org/10.1177/1476127003001001218>
- HARTMANN, W., Nair, H. S. a S. Narayanan. (2011). Identifying Causal Marketing Mix Effects Using a Regression Discontinuity Design. *Marketing Science*, 30(6), 1079–1097.  
<https://doi.org/10.1287/mksc.1110.0670>
- HECKMAN, J. J., Ichimura, H., a P. E. Todd. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job



- 
- Training Programme. *Review of Economic Studies*, 64(4), 605–654.  
<https://doi.org/10.2307/2971733>
- HIRANO, K., Imbens, G. W. a G. Ridder. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, 71(4), 1161–1189.  
<https://doi.org/10.1111/1468-0262.00442>
- HO, D. E., Imai, K., King, G. a E. A. Stuart. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1–28. URL  
<http://www.jstatsoft.org/v42/i08/>.  
<https://doi.org/10.18637/jss.v042.i08>
- HOLLAND, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945–960.  
<https://doi.org/10.1080/01621459.1986.10478354>
- HONG, H. a D. Nekipelov. (2010). Semiparametric Efficiency in Nonlinear LATE Models. *Quantitative Economics*, 1(2010), 279–304.  
<https://doi.org/10.3982/QE43>
- HUBER, M., Lechner, M. a G. Mellace. (2016). The Finite Sample Performance of Estimators for Mediation Analysis Under Sequential Conditional Independence. *Journal of Business a Economic Statistics*, 34(1), 139–160.  
<https://doi.org/10.1080/07350015.2015.1017644>
- IMAI, K., Keele, L. a T. Yamamoto. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25(1), 51–71.  
<https://doi.org/10.1214/10-STS321>
- IMBENS, G. W. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*, 86(1), 4–29.  
<https://doi.org/10.1162/003465304323023651>

- IMBENS, G. W. a J. D. Angrist. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), 467–475.  
<https://doi.org/10.2307/2951620>
- IMBENS, G. W., a T. Lemieux. (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics*, 142(2), 615–635.  
<https://doi.org/10.1016/j.jeconom.2007.05.001>
- IMBENS, G. W. a D. B. Rubin. (2009). *Causal Inference in Statistics and the Social Sciences* (1 ed.). Cambridge and New York: Cambridge University Press.
- IMBENS, G. W. a D. B. Rubin. (2015). *Causal Inference for Statistics, Social, and biomedical Sciences An Introduction*. New York: Cambridge University Press.  
<https://doi.org/10.1017/CBO9781139025751>
- IMBENS, G. W. a J. M. Wooldridge. (2009). Recent Developments in the Econometrics of Program Evaluation, *Journal of Economic Literature*, 47(1), 5–86.  
<https://doi.org/10.1257/jel.47.1.5>
- JETTA, K. a E. W. Rengifo. (2011). A Model to Improve the Estimation of Baseline Retail Sales. *Journal of CENTRUM Cathedra*, 4(1), 10–26.  
<https://doi.org/10.7835/jcc-berj-2011-0048>
- KIMES, S. E. a U. Dholakia. (2011). Restaurant Daily Deals: Customers' Responses to Social Couponing. *Cornell Hospitality Report*, 11(20), 6–18. URL <https://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?article=1071&context=chrpubs>.
- KLEIBER, Ch. a A. ZEILEIS. (2008). *Applied Econometrics with R*. Springer-Verlag, New York. URL <https://CRAN.R-project.org/package=AER>.  
<https://doi.org/10.1007/978-0-387-77318-6>
- KOTLER, P. T. a G. Armstrong. (2014). *Principles of Marketing*. Pearson education.

- 
- KOTLER, P. T. a K. L. Keller (2011). *Marketing Management*. Pearson Prentice Hall, 14th edition.
- LECHNER, M. (2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In Lechner, M. a F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies* (pp. 43–58). Heidelberg and New York: Physica; Mannheim: Centre for European Economic Research.  
[https://doi.org/10.1007/978-3-642-57615-7\\_3](https://doi.org/10.1007/978-3-642-57615-7_3)
- LEE, D. S. a T. Lemieux. (2010). Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, 48(2), 281–355.  
<https://doi.org/10.1257/jel.48.2.281>
- MAGID, M. A. a L. M. Lodish. (1990). Getting the Most out of Advertising and Promotion. *Harvard Business Review*, (May–June issue), 50–60.
- MANSKI, Ch. F. (2007). *Identification for Prediction and Decision*. Harvard University Press, Cambridge and London.
- McAFEE, A. a E. Brynjolfsson. (2012). Big Data: The Management Revolution. *Harvard Business Review*, (October issue), 60–68.
- MONETA, A. a F. Russo. (2014). Causal Models and Evidential Pluralism in Econometrics. *Journal of Economic Methodology*, 21(1), 54–76.  
<https://doi.org/10.1080/1350178X.2014.886473>
- NAGYOVÁ, J. (1999). *Marketingová komunikace není pouze reklama*. VOX, Praha.
- PAUWELS, K., Hanssens, D. M. a S. Siddarth. (2002). The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity. *Journal of Marketing Research*, 39(4), 421–439.  
<https://doi.org/10.1509/jmkr.39.4.421.19114>
- PEARL, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, New York and Melbourne.

- PETERSEN, M., Porter, K., Gruber, S., Wang, Y. a M. Laan. (2010). Diagnosing and Responding to Violations in the Positivity Assumption. *Statistical methods in medical research*, 21, 31–54.  
<https://doi.org/10.1177/0962280210386207>
- R CORE TEAM (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.  
URL <https://www.R-project.org/>.
- ROBINS, J. M. (1986). A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Period – Application to Control of the Healthy Worker Survivor effect. *Mathematical Modelling*, 7, 1393–1512.  
[https://doi.org/10.1016/0270-0255\(86\)90088-6](https://doi.org/10.1016/0270-0255(86)90088-6)
- ROBINS, J. M., Hernán, M. A. a B. A. Brumback. (2000). Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology*, 11, 550–560.  
<https://doi.org/10.1097/00001648-200009000-00011>
- ROSENBAUM, P. R. (1989). Optimal Matching for Observational Studies. *Journal of the American Statistical Association*, 84(408), 1024–1032.  
<https://doi.org/10.1080/01621459.1989.10478868>
- ROSENBAUM, P. R. (1995). *Observational Studies*. Springer Science+Business Media, New York.  
<https://doi.org/10.1007/978-1-4757-2443-1>
- ROSENBAUM, P. R. (2002). Covariance Adjustment in Randomized Experiments and Observational Studies. *Statistical Science*, 17(3), 286–327.  
<https://doi.org/10.1214/ss/1042727942>
- ROSENBAUM, P. R. a D. B. Rubin. (1983b). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55.  
<https://doi.org/10.1093/biomet/70.1.41>

- 
- ROSENBAUM, P. R. a D. B. Rubin. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician*, 39(1), 33–38.  
<https://doi.org/10.1080/00031305.1985.10479383>
- RUBIN, D. B. (1973a). Matching to Remove Bias in Observational Studies. *Biometrics*, 29(1), 159–183.  
<https://doi.org/10.2307/2529684>
- RUBIN, D. B. (1973b). The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies. *Biometrics*, 29(1), 184–203.  
<https://doi.org/10.2307/2529685>
- RUBIN, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5), 688–701.  
<https://doi.org/10.1037/h0037350>
- RUBIN, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, 2(1), 1–26.  
<https://doi.org/10.3102/10769986002001001>
- RUBIN, D. B. (1979). Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74(366), 318–328.  
<https://doi.org/10.2307/2286330>
- RUBIN, D. B. (1980). Comment on: Randomization aAnalysis of Experimental Data: The Fisher Randomization tTest. *Journal of the American Statistical Association*, 75(371), 591–593.  
<https://doi.org/10.2307/2287653>
- RUBIN, D. B. (1990). Formal Mode of Statistical Inference for Causal Effects. *Journal of Statistical Planning and Inference*, 25(3), 279–292.  
[https://doi.org/10.1016/0378-3758\(90\)90077-8](https://doi.org/10.1016/0378-3758(90)90077-8)

- RUBIN, D. B. a N. Thomas. (2000). Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association*, 95(450), 573–585. <https://doi.org/10.1080/01621459.2000.10474233>
- RUBIN, D. B. a R. P. Waterman. (2006). Estimating the Causal Effects of Marketing Interventions Using Propensity Score Methodology. *Statistical Science*, 21(2), A Special Issue on Statistical Challenges and Opportunities in Electronic Commerce Research, 206–222. <https://doi.org/10.1214/088342306000000259>
- SÄVJE, F., Higgins, M. J. a J. S. Sekhon. (2017). Generalized Full Matching, arXiv 1703.03882. URL <https://arxiv.org/abs/1703.03882>.
- SEKHON, J. S. (2011). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*, 42(7), 1–52. URL <http://www.jstatsoft.org/v42/i07/>. <https://doi.org/10.18637/jss.v042.i07>
- SILVA, M. (2012). *Business models to create Customer Loyalty and Generate Profitability: Groupon's Case* [dizertační práce]. URL <https://repositorio.ucp.pt/bitstream/10400.14/12038/1/MasterThesisMartaSilva2012.pdf>.
- SUCHÁNEK, P. a M. KRÁLOVÁ. (2016a). Vztah spokojenosti zákazníka a výkonnosti podniku. *Trendy v podnikání – Business Trends*, Plzeň: Fakulta ekonomická ZČU v Plzni, 6(2), 3–11.
- SUCHÁNEK, P. a M. KRÁLOVÁ. (2016b). Concept of Customer Satisfaction within the Structural Equation Modeling – Measurement Part. In Olexová, C. a M. Janošková. *Marketing Management, Trade, Financial and Social Aspects of Business: Conference Proceedings of the 4th International Scientific Conference*, 231–237, Košice: Faculty of Business Economics with seat in Košice.
- THISTLETHWAITE, D. L. a D. T. Campbell. (1960). Regression-Discontinuity Analysis: An Alternative to the Ex-post Facto

- 
- Experiment. *Journal of Educational Psychology*, 51(6), 309–317.  
<https://doi.org/10.1037/h0044319>
- Van der KLAUW, W. (2008). Regression Discontinuity Analysis: A Survey of Recent Developments in Economics, *Labour*, 22(2), 219–245.  
<https://doi.org/10.1111/j.1467-9914.2008.00419.x>
- Van der WAL, W. M. a R. B. Geskus (2011). ipw: An R Package for Inverse Probability Weighting. *Journal of Statistical Software*, 43(13), 1–23. URL <http://www.jstatsoft.org/v43/i13/>.  
<https://doi.org/10.18637/jss.v043.i13>
- VYTLACIL, E. (2002). Independence, Monotonicity and Latent Variable Models: an Equivalence Result, *Econometrica*, 70, 331–341.  
<https://doi.org/10.1111/1468-0262.00277>
- WANG, R. (2011). *What CIO's And CMO's Should Know About Daily Deal Sites Such As Groupon*. Forbes. URL <https://www.forbes.com/sites/ciocentral/2011/04/08/what-cios-and-cmos-should-know-about-daily-deal-sites-such-as-groupon/#769bb1c047b7>.
- YOSHIDA, K. a J. Bohn. (2018). *tableone: Create 'Table 1' to Describe Baseline Characteristics*. URL <https://CRAN.R-project.org/package=tableone>.
- ZEILEIS, A. (2006). Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software*, 16(9), 1–16. URL <http://www.jstatsoft.org/v16/i09/>.  
<https://doi.org/10.18637/jss.v016.i09>





## Příloha A

# Instrumentální proměnné

## A.1 Tradiční ekonometrický přístup

Instrumentální proměnná řeší problém vychýlených odhadů v modelech s následující specifikací:

$$Y = \alpha_0 + \alpha_1 W + \varepsilon_2 \quad (\text{A.1.1})$$

kde  $\text{cov}(W, \varepsilon_2) \neq 0$ .  $W$  je endogenní proměnná a v kontextu kapitoly 4.1 reprezentuje ošetření. Parametr  $\alpha_1$  pak odpovídá efektu ošetření  $W$  na odezvu  $Y$ .

Potom pro odhad parametru  $\alpha_1$  pomocí metody nejmenších čtverců (OLS) platí:

$$\begin{aligned} \hat{\alpha}_{1,OLS} &= \frac{\text{cov}(Y, W)}{\text{var}(W)} = \frac{\text{cov}(\alpha_0 + \alpha_1 W + \varepsilon_2, W)}{\text{var}(W)} \\ &= \frac{\alpha_1 \text{cov}(W, W) + \text{cov}(W, \varepsilon_2)}{\text{var}(W)} \\ &= \alpha_1 + \frac{\text{cov}(W, \varepsilon_2)}{\text{var}(W)} \\ E(\hat{\alpha}_{1,OLS}) &= \alpha_1 + E\left(\frac{\text{cov}(W, \varepsilon_2)}{\text{var}(W)}\right) \end{aligned} \quad (\text{A.1.2})$$

Odhad  $\hat{\alpha}_{1,OLS}$  je kvůli nenulové kovarianci  $\text{cov}(W, \varepsilon_2)$  vychýlený. Toto vychýlení lze odstranit prostřednictvím instrumentální proměnné, která je v souladu s kapitolou 4.1 značená  $Z$ . Místo modelu A.1.1 budeme nyní uvažovat „dvojstupňový model“ a odhad parametru  $\alpha_1$  pak získáme pomocí metody 2SLS (Two-Stage-Least-Square):

$$\text{Druhý stupeň: } Y = \alpha_0 + \alpha_1 W + \varepsilon_2 \quad (\text{A.1.3})$$

$$\text{První stupeň: } W = \pi_0 + \pi_1 Z + \varepsilon_1 \quad (\text{A.1.4})$$

Předpoklady:

$$\text{cov}(\varepsilon_1, Z) = 0, \pi_1 \neq 0, \text{cov}(\varepsilon_2, Z) = 0$$

Uvedené předpoklady vycházejí z následujících požadavků na instrument  $Z$ :

1.  $\pi_1 \neq 0$ , tedy instrument  $Z$  musí mít vliv na ošetření  $W$ . Tomuto předpokladu se říká *relevance* a lze jej testovat.
2.  $Z$  musí být randomizované (nebo „podobné“ randomizovanému), proto  $\text{cov}(\varepsilon_1, Z) = 0$ .
3.  $Z$  nesmí mít žádný přímý vliv na  $Y$ , pouze skrze  $W$ . Proto se  $Z$  neobjevuje v rovnici druhého stupně a požadujeme  $\text{cov}(\varepsilon_2, Z) = 0$ . Tento předpoklad nazýváme *restrikce na výlučnost* (*exclusion restriction*) nebo v ekonometrické literatuře je požadavek  $\text{cov}(\varepsilon_2, Z) = 0$  označován jako *exogenita*<sup>35</sup>. Tento předpoklad bohužel nejde testovat.

**V uvedeném dvojstupňovém modelu lze identifikovat tři efekty:**

1. Efekt instrumentu  $Z$  na ošetření  $W$  (*efekt prvního stupně*)
2. Efekt instrumentu  $Z$  na odezvu  $Y$  (*ITT efekt*, intent-to-treat )
3. Efekt ošetření  $W$  na odezvu  $Y$  výlučně skrze exogenní variabilitu  $W$  indukovanou instrumentem  $Z$  (*IV efekt*, t.j. efekt ošetření pomocí instrumentu)

**Efekt prvního stupně  $Z$  na  $W$ :**

$$\begin{aligned} \hat{\pi}_1 &= \frac{\text{cov}(W, Z)}{\text{var}(Z)} \\ &= \frac{\text{cov}(\pi_0 + \pi_1 Z + \varepsilon_1, Z)}{\text{var}(Z)} \\ &= \frac{\pi_1 \text{cov}(Z, Z) + \text{cov}(Z, \varepsilon_1)}{\text{var}(Z)} \\ &= \pi_1 + \frac{\text{cov}(Z, \varepsilon_1)}{\text{var}(Z)} \end{aligned}$$

$$E(\hat{\pi}_1) = \pi_1 + E\left(\frac{\text{cov}(Z, \varepsilon_1)}{\text{var}(Z)}\right) = \pi_1$$

<sup>35</sup> V ekonometrii tradiční termín *exogenita* používá [Imbens \(2004\)](#) v jiném významu, a to pro předpoklad nezavádějícího mechanismu přiřazování. Není to zcela příhodné, jelikož  $X_i$  není v tradičním ekonometrickém významu exogenní vůči  $E[Y_i(1) - Y_i(0)]$ . Znalost rozložení  $X_i$  je informativní vzhledem k  $E[Y_i(1) - Y_i(0)]$ .

(Odhad  $\hat{\pi}_1$  lze v **R** získat buď pomocí příkazu pro lineární modely `lm`, nebo pomocí příkazu pro kovarianční matici `cov`.)

### ITT efekt:

Nejdříve v rovnici A.1.3 za  $W$  dosadíme z rovnice A.1.4 a získáme:

$$\begin{aligned} Y &= \alpha_0 + \alpha_1(\pi_0 + \pi_1 Z + \varepsilon_1) + \varepsilon_2 \\ Y &= (\alpha_0 + \alpha_1 \pi_0) + (\alpha_1 \pi_1) Z + (\alpha_1 \varepsilon_1 + \varepsilon_2) \end{aligned} \quad (\text{A.1.5})$$

$$Y = \gamma_0 + \gamma_1 Z + \varepsilon_3 \quad (\text{A.1.6})$$

kde  $\gamma_0 = \alpha_0 + \alpha_1 \pi_0$ ,  $\gamma_1 = \alpha_1 \pi_1$ ,  $\varepsilon_3 = \alpha_1 \varepsilon_1 + \varepsilon_2$

Potom:

$$\hat{\gamma}_1 = \frac{\text{cov}(Y, Z)}{\text{var}(Z)} = \frac{\text{cov}(\gamma_0 + \gamma_1 Z + \varepsilon_3, Z)}{\text{var}(Z)}$$

$$E(\hat{\gamma}_1) = \gamma_1 + E\left(\frac{\text{cov}(Z, \varepsilon_3)}{\text{var}(Z)}\right) = \gamma_1$$

jelikož z  $\text{cov}(\varepsilon_1, Z) = 0 = \text{cov}(\varepsilon_2, Z)$  plyne, že i  $\text{cov}(\varepsilon_3, Z) = 0$ .

### IV efekt:

Nejdříve připomeňme, že rovnice A.1.3 je upravena na tvar A.1.5, tedy:

$$\begin{aligned} Y &= (\alpha_0 + \alpha_1 \pi_0) + (\alpha_1 \pi_1) Z + (\alpha_1 \varepsilon_1 + \varepsilon_2) \\ Y &= \gamma_0 + \gamma_1 Z + \varepsilon_3 \end{aligned}$$

Potom odhad parametru  $\alpha_1$  lze získat pomocí:

$$\begin{aligned} \alpha_1 &= \frac{\gamma_1}{\pi_1} = \frac{\text{Efekt } Z \text{ na } Y}{\text{Efekt } Z \text{ na } W} \\ &= \frac{\text{cov}(Y, Z) / \text{var}(Z)}{\text{cov}(W, Z) / \text{var}(Z)} = \frac{\text{cov}(Y, Z)}{\text{cov}(W, Z)} \\ \hat{\alpha}_1 &= \frac{\text{cov}(\alpha_0 + \alpha_1 W + \varepsilon_2, Z)}{\text{cov}(W, Z)} = \frac{\alpha_1 \text{cov}(W, Z) + \text{cov}(\varepsilon_2, Z)}{\text{cov}(W, Z)} = \\ \hat{\alpha}_1 &= \alpha_1 + \frac{\text{cov}(\varepsilon_2, Z)}{\text{cov}(W, Z)} \end{aligned}$$

$$E(\hat{\alpha}_1) = \alpha_1 + E\left(\frac{\text{cov}(\varepsilon_2, Z)}{\text{cov}(W, Z)}\right) = \alpha_1$$

Tedy na rozdíl od vychýleného odhadu  $\alpha_1, \hat{\alpha}_{OLS}$  A.1.2 je odhad  $\hat{\alpha}_1$  nevychýlený.

$$\text{IV efekt } \alpha_1 = \frac{\text{Efekt } Z \text{ na } Y}{\text{Efekt } Z \text{ na } W} = \frac{\text{cov}(Y, Z)}{\text{cov}(W, Z)} \quad (\text{A.1.7})$$

Nahradíme-li teoretické kovariance výběrovými, získáme *Waldův odhad* IV efektu. Realizace tohoto odhadu je numericky stejná, pokud použijeme 2SLS postup:

1. V prvním stupni A.1.4 odhadneme

$$\hat{W} = \hat{\pi}_0 + \hat{\pi}_1 Z$$

2. V druhém stupni A.1.3 dosadíme

$$Y = \alpha_0 + \alpha_1 \hat{W} + \varepsilon_2$$

$$Y = \alpha_0 + \alpha_1 (\hat{\pi}_0 + \hat{\pi}_1 Z) + \varepsilon_2$$

$$Y = (\alpha_0 + \alpha_1 \hat{\pi}_0) + \alpha_1 (\hat{\pi}_1 Z) + \varepsilon_2$$

$\alpha_1$  je nyní pomocí 2SLS identifikována výhradně na základě té části variability  $W$ , jež pochází ze  $Z$ . Hodnota bodového odhadu je stejná jako realizace odhadu IV efektu Waldovou statistikou. Bohužel ale směrodatná chyba odhadu je nyní zkreslená, jelikož nezohledňuje variabilitu  $\hat{\pi}_0$  a  $\hat{\pi}_1$ . (Waldův odhad pro  $\alpha_1$  lze v **R** získat pomocí příkazu `cov`; odhad pomocí 2SLS lze získat buď dvěma vnořenými `lm` modely, nebo po nainstalování balíčku **AER** pomocí příkazu `ivreg`.) Přístup 2SLS lze zobecnit i na modely s vektorem kovariát.

## A.2 Moderní přístup pomocí potenciálních odezev

Tradiční přístup předpokládá konstantní efekt ošetření na všechny subjekty populace. Často ale v praxi instrument ovlivňuje pouze část populace a informace o efektu ošetření na „neovlivnitelné“ subjekty není k dispozici. Tuto skutečnost reflektuje kauzální model s instrumentální proměnnou popsany v sekci 4.1. Níže je odvození souvislosti mezi 4.1.7 a 4.1.9.

Na základě 4.1.5 a 4.1.6 je

$$\tau_{LATE} = E[Y(1) - Y(0) \mid W(1) > W(0)] = \frac{ITT}{P_c}$$

Nyní ukážeme, že

$$\tau_{LATE} = \frac{E[Y | Z = 1] - E[Y | Z = 0]}{E[W | Z = 1] - E[W | Z = 0]}$$

$$\frac{E[Y | Z = 1] - E[Y | Z = 0]}{E[W | Z = 1] - E[W | Z = 0]} = \tag{A.2.1}$$

$$= \frac{E[Y(0) + (Y(1) - Y(0)) \cdot W(1) | Z = 1]}{E[W(1) | Z = 1] - E[W(0) | Z = 0]} - \frac{E[Y(0) + (Y(1) - Y(0)) \cdot W(0) | Z = 0]}{E[W(1) | Z = 1] - E[W(0) | Z = 0]} = \tag{A.2.2}$$

$$= \frac{E[Y(0) + (Y(1) - Y(0)) \cdot W(1)]}{E[W(1)] - E[W(0)]} - \frac{E[Y(0) + (Y(1) - Y(0)) \cdot W(0)]}{E[W(1)] - E[W(0)]} = \tag{A.2.3}$$

$$= \frac{E[(Y(1) - Y(0)) \cdot (W(1) - W(0))]}{E[W(1) - W(0)]} = \tag{A.2.4}$$

$$= \frac{E[(Y(1) - Y(0)) | W(1) > W(0)] \cdot P(W(1) > W(0))}{E[W(1) - W(0)]} - \frac{E[(Y(1) - Y(0)) | W(1) < W(0)] \cdot P(W(1) < W(0))}{E[W(1) - W(0)]} = \tag{A.2.5}$$

$$= \frac{E[(Y(1) - Y(0)) | W(1) > W(0)] \cdot P(W(1) > W(0))}{P(W(1) > W(0))} = \tag{A.2.6}$$

$$= E[(Y(1) - Y(0)) | W(1) > W(0)] = \tau_{LATE} \tag{A.2.7}$$

V A.2.2 a A.2.3 je využita skutečnost, že instrument  $Z$  je randomizovaný.

V A.2.5 a A.2.6 je využita skutečnost, že

$$W(1) - W(0) = \begin{cases} 1 & \text{pro } W(1) > W(0) \\ & \text{compliers} \\ 0 & \text{pro } W(0) = W(1) = 1 \vee W(0) = W(1) = 0 \\ & \text{always, never takers} \\ -1 & \text{pro } W(1) < W(0) \\ & \text{defiers} \end{cases}$$

a díky předpokladu monotonie je pravděpodobnost pro „defiers“ rovna nule.

Platí tedy, že

$$\begin{aligned}\tau_{LATE} &= E[Y(1) - Y(0) \mid W(1) > W(0)] = \\ &= \frac{ITT}{P_c} = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[W \mid Z = 1] - E[W \mid Z = 0]}\end{aligned}$$

## Příloha B

# Balíky pro kauzální analýzu v R

**Tabulka B.1:** Přehled nejužívanějších balíků jazyka R vhodných pro metody kauzální analýzy popsané v kapitolách 3 a 4

Název balíku	Popis použití	Reference
<b>Balíky potřebné pro párovací procedury</b>		
Matching	Párování na základě vícerozměrných charakteristik či propenzitních skóru	<a href="#">Sekhon (2011)</a>
MatchIt	Párování na základě vícerozměrných charakteristik či propenzitních skóru	<a href="#">Ho, Imai, King a Stuart (2011)</a>
tableone	Přehled popisných charakteristik vhodný pro párovací procedury	<a href="#">Yoshida a Bohn (2018)</a>
sandwich	Robustní odhady směrodatných chyb	<a href="#">Zeileis (2006)</a>
<b>Balík pro metodu IPTW</b>		
ipw	Odhady vah pro metodu IPTW a následné analýzy	<a href="#">Van der Wal a Geskus (2011)</a>
<b>Balík pro instrumentální proměnné</b>		
AER	Aplikovaná ekonometrie	<a href="#">Kleiber a Zeileis (2008)</a>

## VĚDECKÁ REDAKCE MASARYKOVY UNIVERZITY

doc. Ing. Pavel Bobál, CSc.; prof. JUDr. Marek Fryšták, Ph.D.;  
Mgr. Michaela Hanousková; doc. RNDr. Petr Holub, Ph.D.;  
doc. Mgr. Jana Horáková, Ph.D.;  
prof. MUDr. Lydie Izakovičová Hollá, Ph.D.;  
prof. PhDr. Mgr. Tomáš Janík, Ph.D.; prof. PhDr. Tomáš Kubíček, Ph.D.;  
prof. RNDr. Jaromír Leichmann, Dr. rer. nat.;  
doc. RNDr. Lubomír Popelínský, Ph.D.; Ing. Zuzana Sajdlová, Ph.D.;  
Mgr. Kateřina Sedláčková, Ph.D.; prof. RNDr. Ondřej Slabý, Ph.D.;  
doc. Ing. Rostislav Staněk, Ph.D.; prof. PhDr. Jiří Trávníček, M.A.;  
doc. PhDr. Martin Vaculík, Ph.D.

## METODY KAUZÁLNÍ ANALÝZY PRO MĚŘENÍ EFEKTIVITY PODPORY PRODEJE

Maria Králová

Obálka: Václav Mekyska

Sazba: Vydavatelský servis, Plzeň

Vydala Masarykova univerzita, Žerotínovo nám. 617/9, 601 77 Brno

Vydání 1., elektronické / 2021

ISBN 978-80-210-9940-1

ISBN 978-80-210-9939-5 (brožováno)

<https://doi.org/10.5817/CZ.MUNI.M210-9940-2021>



Společnosti v ekonomicky rozvinutých zemích vynakládají na aktivity spojené s podporou prodeje stále více prostředků. Z řady studií však plyne, že mnohé z aktivit generují zanedbatelný, žádný, či dokonce záporný zisk. Jednou z možných příčin nežádoucího jevu jsou nedostatečné analytické nástroje, které by umožňovaly efektivitu podpory prodeje spolehlivě měřit. Běžně používané nástroje vedou díky nezapočítání vlivu mnoha v čase se měnících faktorů, jež mohou ovlivňovat ukazatele ziskovosti, k potencionálně nepřesným odhadům efektu podpůrných aktivit.

V monografii je představen nový přístup k odhadování základních tržeb (bez akcí na podporu prodeje) pomocí metod kauzální analýzy. Tento přístup je celkem jednoduchý, mohl by být implementován do běžné firemní praxe, a přitom odstraňuje nepřesnosti odhadu kvůli faktorům měnícím se v čase.

Kniha se zaměřuje také na metody KAUZÁLNÍ ANALÝZY, monograficky dosud v českém jazyce nezpracované, které mohou řešit mnohé problémy z oblasti marketingu i managementu a obecně sociálních věd, kdy jsou k dispozici pouze observační data. Jedná se o téma velmi aktuální, američtí ekonomové Joshua Angrist a Guido Imbens se stali v roce 2021 laureáty Nobelovy ceny za ekonomii právě za metodologický přístup k analýze kauzálních vztahů.