

Systematické chyby kvalifikačních prací, jejich obhajob

Zaměřeno na společenské vědy, ekonomii zejména
vydání 2014

Petr Houdek*

Obsah

1	K čemu tento text slouží?	2
2	Abstrakt a Úvod – neinformativní a bez výsledků	2
3	Vata v přehledu existující literatury či v tzv. »teoretické kapitole«	3
4	Kvalitativní analýza – chybějící polemika	5
5	Kvantitativní analýza – »něco mi vyšlo«	6
6	Závěr – formalistní a neobjevný	9

*Přípomínky k textu (verze 2014-09-01) prosím pošlete na: petr.houdek@gmail.com. Vizte rovněž dokument **Vzorová struktura semestrálních a kvalifikačních prací**.

1 K čemu tento text slouží?

„Komu není rady, tomu není pomoci.“

[citation needed]

Před započítím tvorby své kvalifikační práce byste měli mít přečteny desítky a možná i stovky odborných studií či knih.¹ Měli byste tušit, jak dobří autoři formulují argumenty, jak text strukturují a jak studie formálně vypadají. Pakliže stále tápete, nahlédněte do nějakého **průvodce, návodu či manuálu k psaní odborných či kvalifikačních prací**.² Rovněž byste měli znát postup pro tvorbu a odevzdávání prací uvedený na stránkách katedry, pod níž svoji práci vytváříte. Pravidelně byste svoji práci měli konzultovat minimálně se svým vedoucím.

Ač studenti tyto zdroje obvykle znají a využívají, řada chyb či nejasností je při obhajobách a v oponentských posudcích pravidelně kritizována. Tento pamflet jejich nevyčerpávající seznam přináší a krátce zmiňuje možná řešení. Respektování uvedených doporučení vám pomůže napsat kvalitní, smysluplný a užitečný text a vyhnout se kritickému posudku a nepříznivému hodnocení při obhajobě.

- Body s odrážkou se týkají úskalí vlastních obhajob. Pamatujte, že ne všichni členové komise vaši práci detailně prozkoumali, vycházejí proto z posudků, vaší prezentace a následných odpovědí. Studenti, bohužel, rétorickou i faktickou přípravu na obhajoby podceňují.

2 Abstrakt a Úvod – neinformativní a bez výsledků

„Dobré kázání by mělo být jako dámská sukně: krátké, aby vzbudilo zájem, ale dostatečně dlouhé, aby pokrylo to nejpodstatnější.“

Ronald A. Knox

Abstrakt je shrnutí práce v poměrně rigidní struktuře: problém – užitá metoda (není-li implicitně zřejmá) – **výsledky** – závěr a obvykle i doporučení. Abstrakt musí být pochopitelný a informativní, aniž by čtenář musel nahlížet do práce samotné (chybné jsou formulace typu: „*V třetí části je vysvětleno, proč se EU rozpadne.*“). Časté jsou chyby, kdy autor v abstraktu jen popisuje, na co se zaměří („*bude analyzována aktivní politika nezaměstnanosti*“). Abstrakt je poslední částí, která se píše, **hovoří tak o výsledcích, nikoliv o cílech** (proto není totožný se zadáním kvalifikační práce!). Abstrakt má být názorný, krátký a úderný (obsahuje-li přehršel příslovců, jde o nápovědu, že je příliš rozvláčný).

Nejčastějším lapsem úvodu je absence jasné formulace, v čem práce přispívá k poznání či k řešení problému. Po představení problému, krátkém popisu zkoumané tematiky a přehledu hlavních poznatků (tj. nejdůležitějších studií) studenské úvody obvykle končí: „*Tato práce chce prohloubit pochopení ... chce dále vysvětlit ... zaměřuje se na Českou republiku ... na Ústecko*“. Po těchto vágních tvrzeních však nenásleduje upřesnění, čeho bylo v práci dosaženo. V ekonomických oborech

Obsahuje abstrakt: problém, metodu, výsledky (a doporučení)?

Obsahuje úvod: popis problému, jeho zodpovězení či přínos práce?

¹Stát se dobrým autorem vyžaduje být dobrým čtenářem. Psaní je totiž velmi nepřírozená činnost, protože jde o komunikaci bez reakcí druhé strany. Jen málokdo dokáže přirozeně svou samomluvou–psáním srozumitelně, jednoznačně a vyčerpávajícím způsobem vysvětlit věc, o kterou mu jde. My ostatní se to musíme naučit, četbou.

²Na většině univerzit existují podobné návody, skrývají se pod hesly „academic writing“, „scientific writing“ či „writing labs“.

platí úzus, že již z úvodu **by mělo být v krátkosti patrné, k čemu práce dospěla, kam posunula poznání či jaká je její přidaná hodnota** (což mohou být i unikátní data či třeba četba dosud neobjevené práce významného autora).

- Obhajoby kvalifikačních prací by neměly být zkouškou znalostí oboru. Nezbytné vědomosti byste již měli mít! Kvalifikační práce je důkazem, že dokážete aplikovat naučené poznání při vyřešení odborného problému. Obhajoba proto cílí na ověření, že rozumíte datům, užitým metodám a interpretaci výsledků. Studenti však mají tendence omílat teorie, učebnicové poučky či jiné studie, místo aby se soustředili na vlastní práci. Tím komisi jen signalizují, že je s jejich prací něco špatně...
- Typický úvod obhajob: „*Co Vaše práce přináší (nového)?*“ by měl být následován několika větami o uplatněné metodě a dosažených výsledcích. Nikoliv – jak je bohužel obvyklé – floskulami typu „*psal jsem o sociálních dávkách v ČR*“, taji motivace k práci, popisování oblasti, již se práce týká („*už von Bismarck zavedl sociální pojištění*“) či dokonce obtížemi při jejím vypracování.
- Oponentské posudky – ač nesporně velmi důležité – **nejsou** konečným hodnocením práce. Z textu samotného někdy nelze odvodit, nakolik autor užitě metody ovládá či je jen bezmyšlenkovitě přejímá. Komise se proto ptají na úzká místa, implicitní předpoklady analýzy či možnosti alternativních přístupů. Je varovné, odpovídá-li student dokola jen tím, co má napsáno ve „své“ práci.

3 Vata v přehledu existující literatury či v tzv. »teoretické kapitole«

„Pravda uvízne v našich hrdlech se všemi těmi omáčkami, s kterými je obvykle servírována; nikdy ji nespolkneme, pakliže omáčku zcela neodstraníme.“

George B. Shaw

Snad všechny kvalifikační práce trpí záplavou nadbytečného či i zcela irrelevantního textu. Studenti se domnívají, že každý pojem je nutné detailně vysvětlit, vedeni patrně příkladem vysokoškolských učebnic. **Předpokládejte vzdělaného čtenáře!** Obzvláště v diplomových pracích **je redundantní v několika odstavcích rozepisovat základní koncepty či jejich historii.**

Využíváte-li v analýze hodnoty změn HDP pro měření dopadu hospodářského cyklu na nemocnost, je zcestné pojednat v několika subkapitolách o historii zjišťování HDP či o metodách jeho měření. Nebo poté, co se na několika stranách čtenář dozví, jaké jsou druhy nezaměstnanosti (jako frikční, strukturální a cyklická či ne/dobrovolná, jak se mezi sebou liší a jak se měří), je v analýze pracováno jen s prostou mírou nezaměstnanosti.

Není-li podstatná část podobných obsáhlých expoé využita v analýze, jde o závažnou chybu – ukazuje, že student nemá schopnost odlišit nepodstatné informace od důležitých!

- Zeptá-li se oponent či komise, proč práce obsahuje kapitoly, které nejsou nijak využity, studenti obvykle odpovídají, „*že se takové věci učili*“. Odpověď stejně absurdní, jako kdyby na otázku po směru k nejbližší benzínové pumpě turistovi odpověděli, že první silnice na světě byla postavena roku 312 před Kristem římským cenzorem Appiem Claudiem...

Týká se text přímo řešeného problému?

Zaplevelenost textu je chyba! Nikoliv malá.

Každá delší část textu se musí **přímo týkat postulování problému, metody jeho zodpovězení a nebo výsledků a jejich diskuze**. Píšete-li o fiskálních dopadech morálního hazardu zemí v Evropské měnové unii, je skutečně zbytečné věnovat několikastránkovou kapitolu popisné historii zakládání Evropské unie! Zcela dostačuje historická vsuvka v pár souvětích, následována popisem mechanismu fungování Eurozóny vytvářejícího morální hazard.

Srozumitelné i bez odstavce? Smazat!

Zkrátka, můžete-li smazat nějakou část textu a práce přináší stále stejné pochopení problému, smažte ji, resp. ji nahraďte. Přestože studentské práce obsahují záplavu zbytečností, polemické sekce naopak chybějí (vizte dále část 4.).

Je varovné, když nejnovější citovaný článek je 10 let starý a průměrné stáří užitě literatury přesahuje věk studenta a někdy i věk vedoucího. Poznání se vyvíjí a ač je vhodné citovat texty, které přinesly původní teoretický koncept či silný empirický fakt, navazující výzkum téměř jistě upravil, rozšířil a často zneplatnil dosažené závěry. Kvalifikační práce přesto citují jen „otce zakladatele“ oblasti a nato přejdou k vlastní analýze, v níž s neomylnou pravidelností vykáží identická úzká místa, která jsou již mnoho let i desetiletí v odborné literatuře kritizována. Soudobá věda může být tažena i módností bez podstatné přidané hodnoty, abyste však neopakovali chyby minulosti či ve své analýze neopomněli důležité faktory, **je krucíální vycházet z aktuálního poznání**.³ Nejste-li si jisti aktuálností či relevantností užívaných studií, využijte konzultace svého vedoucího, popř. jiného odborníka.

Jde o aktuální poznání?

Pamatujte, že zevrubná kapitola přehled literatury vás ušetří putování po slepých uličkách, mnoha klopýtnutí a zklamání. Poznáním kvalitních studií nahlédnete, kudy cesty nevedou a kudy naopak dojdete k cíli rychle a bez zbytečné námahy.

Studium mnoha zdrojů = úspora času

Studenti mají tendence k uvažování „zdroj jako zdroj“. **Všechny studie si však nejsou rovny**. Závěry metastudií či výzkumů využívajících data sety detailních informací o rozhodnutích miliónů lidí za dlouhá období, popř. rozsáhlá mezinárodní data shromažďovaná uznávanými organizacemi jsou obvykle důvěryhodnější než výsledky tematicky shodných studií s horší kvalitou dat, shromažďovaných třeba neznámými firmami, společenskovedních laboratorních či terénních experimentů či studií využívajících dotazníková šetření.

Je zvažena různá kvalita studií?

Naprostá většina citací by měla být na uznávané zdroje dat a originální články v odborných časopisech. Takové zdroje prošly odbornou kontrolou (tzv. *peer-review procesem*) a jsou ověřitelné, tzn. data, metody i závěry lze (obvykle) znovu prozkoumat a potvrdit (či vyvrátit).

Vychází se z uznávaných zdrojů?

Citujete-li Wikipedii, noviny či společenské časopisy a zájmové servery – jak je u studentů bohužel stále zvykem – je téměř jisté, že jste hrozivě nezvládli práci s literaturou.⁴

Znáte využívanou referenci?

- **Uvádějte reference na literaturu, kterou jste přečetli!** Nikdo po vás, pochopitelně, nebude vyžadovat, abyste si pamatovali detailní obsah všech využívaných studií. Je však varovné, citujete-li výzkum (někdy opakovaně), přesto vaše práce přistupuje k stejnému problému jinak, opačně než citovaná práce doporučuje či využívá. Po zvědavé otázce oponenta či komise pak vyjde najevo, že jste dílo nečetli a referenci na něj „odněkud“ převzali. Zbytečně dodávat, že jde o podstatnou chybu.

Samozřejmostí by měla být zvládnutá práce s referencemi/citacemi. Bohužel se stále objevují kvalifikační práce, které přeberou text, myšlenku, argumentaci, data a nebo – jako uvedeno výše – citace

Není práce plagiátem?

³Samozřejmě můžete nalézt i staré, opomenuté klenoty ekonomické literatury, které jsou soustavně přehlíženy a vystavět na jejich základě aktuální a objevnou práci! Takové případy jsou však u kvalifikačních prací – bohužel – spíše vzácností.

⁴Některé novinové články přinášejí unikátní informace, proto je pochopitelně možné (či nutné) i noviny a podobná média citovat. Většinou se jedná o investigativní či jinak nedostupné informace. Citovat noviny coby zdroj dat třeba o nezaměstnanosti však přibližuje studenta spíše k hodnocení nevyhověl.

na další studie a **neodcitují původní práci**. Tato praxe je plagiátorstvím, vede k hodnocení nevyhověl a předvolání před disciplinární komisí fakulty.⁵

4 Kvalitativní analýza – chybějící polemika

„Nejobtížnější téma může být vysvětleno i úplnému hlupákovi, pakliže si na něj dosud nevytvořil žádný názor; avšak i ta nejjednodušší věc nebude moci být vysvětlena tomu nejchytřejšímu člověku, pakliže si je pevně jist, že ji – bez stínu pochybnosti – pochopil.“

Lev N. Tolstoj

Mnohé práce zaměřují analýzu problému s jeho popisem či jen prostým popisem faktů či teorií. (Zejména se to týká právě kvalitativních textů, takových, v kterých autor sám nepracuje s empirickými daty, ale vychází jen z přejeté literatury.)

Cílem kvalifikační práce je nalézt **důvody, proč k danému jevu došlo**, vyzdvihnout faktory, které sledovaný jev ovlivňují nebo formulovat podmínky, za kterých je jev možný. Tj. neměli byste psát „o“ ale i „protože“. Nestačí tedy pojednat jen „že v 60. letech byla sexuální revoluce“, ale je nutné dodat, že šlo o důsledek nadměrné nabídky mladších žen po poválečném baby boomu nebo o důsledek efektivnější antikoncepce (hypotéza „pilulky“) nebo další vysvětlení. Obsahuje-li text jen „došlo k události“ – je neúplný, musí být dopsáno „protože...“.

Vysvětlením solárního boomu v letech 2009 a 2010 v Česku jistě může být, že vstoupil v platnost zákon o podpoře využívání obnovitelných zdrojů. Takové vysvětlení je ale toliko proximální, ukazuje mechanismus, nikoliv důvod. **Odborné práce by měly nabízet spíše ultimátní vysvětlení** – ukazující hlubší důvody skutečností (v tomto případě by práce měla pojednat o lobbingu, veřejné volbě, informační asymetrii či korupci). Řečeno metaforicky, nikdy nevráždí zbraně, ale vždy jen lidé.

Při dostatečné snaze lze nalézt reference pro i proti téměř jakémukoliv tvrzení, v hospodářské politice zejména. Cti autora jistě nesvědčí, využívá-li pro svou práci jen argumenty či studie, které podporují jeho názor či teorii. Měly by být vyhledány **různorodé argumenty či důkazy** a na základě jejich diskuze či vzájemné konfrontace zodpovězena řešená otázka. Místo toho řada kvalifikačních prací jen předkládá argumenty pro předem vybranou odpověď. Píšete-li o blahodárném vlivu konkurence mezi zdravotními zařízeními a citujete-li studie dokazující dopad vyšší konkurence na snížení mortality, vyhledejte i studie uzavírající opak. Až na základě prozkoumání (diskuze) robustnosti, metod či kvality dat pluralitních studií formulujte svůj závěr.

Častou chybou studentů je nepochopení termínu „předpoklad“. S předpokladem k analýze přistupujeme (omezujeme nekonečnou komplexitu reality na její úzkou reprezentaci), proto jej nikdy nemůžeme „vyvrátit“ fakty. **Zkoumáme tvrzení, která jsou vystavěna za podmínky, že předpoklad platí**. „Za předpokladu, že děti jsou méněcenný statek, chci určit, proč s růstem důchodu klesá poptávka domácností po nich,“ je hodnotné teoretické zjišťování. Autor omezil své pojednání na situace, kdy děti jsou méněcenným statkem a hledá teoretické důvody jen tohoto jevu. I když empirické studie zjistí, že v některých zemích, v některých obdobích jsou děti normálním statkem, jeho analýzu to nezpochybní.

⁵Pro způsoby citování vizte Bratková, Eva (2008): Metody citování literatury a strukturování bibliografických záznamů podle mezinárodních norem ISO 690 a ISO 690-2: metodický materiál pro autory vysokoškolských kvalifikačních prací. Asociace knihoven vysokých škol ČR. Dostupné elektronicky na: <http://www.evskp.cz/SD/4c.pdf> [k 1. září 2013].

Nejde jen o popis? Odpovídá se na otázku „proč“ či „jak“?

Ultimátní a proximální příčiny

Má text polemický ráz?

Předpoklady v. hypotézy

Zjistíme-li však, že předpoklad silně limituje aplikovatelnost dosažených výsledků – tj. skoro vždy je v rozporu s fakty – můžeme jej poté změnit či uvolnit. Na druhou stranu, „za předpokladu, že děti jsou méněcenný statek, chci otestovat, zda s růstem důchodu klesá poptávka domácností po nich“ je zbytečné, jelikož z předpokladu vyplývá závěr (jsou to synonyma), není třeba nic testovat! Správná formulace by byla: „Chci otestovat, že děti jsou méněcenný statek čili že s růstem důchodu klesá poptávka domácností po nich“.

Předpoklady v.
fakta

Podobnou chybou je, že studenti „předpokládají“ něco, co ve skutečnosti **je předmětem jejich zkoumání a musí být zjištěno**. „Předpokládám, že po metanolové aféře koncem roku 2012 klesl prodej alkoholu. Chci otestovat, zda nižší spotřeba alkoholu vedla k poklesu nehod a zranění v Česku.“ Podobné zkoumání je cenné, nicméně pokles spotřeby či prodeje alkoholu zde nelze jen předpokládat. Je nezbytné ověřit, zda reálně spotřeba klesla, protože jen tak lze identifikovat vliv alkoholu na nehody.

Nevyřeší vše
trh nebo stát?

Závěry by měly být vyvážené, kupř. dospěl-li autor k výsledku, že konkrétní politika nevede k stanoveným cílům, není vhodný prostý závěr: „Bylo by žádoucí odstranění státního zásahu a ponechání vše volnému trhu“. Zneplatnění efektivnosti konkrétní politiky **nutně neznamená**, že její jiný mechanismus by nemohl být efektivní (zde by byla nutná mezinárodní či historická komparace), popř. že by „volný trh“ byl v dané oblasti optimální.⁶ Stejně tak, vede-li tržní selhání k neefektivitě, nelze jen napsat, že „stát by měl zasáhnout“. Státní řešení totiž může vést v krátkém i delším období k ještě horším dopadům.

Podobné závěry vyžadují podstatně hlubší – empirickou i teoretickou – analýzu, která v naprosté většině prací absentuje.

5 Kvantitativní analýza – »něco mi vyšlo«

„Jsem si jist, že toto je správné řešení. Nejsem si však jist jeho interpretací.“

Gary S. Becker

Užitečnost empirické analýzy závisí na kvalitě dat a na jejich vyhodnocení. Pravidlo „odpad do ní, odpad z ní“ platí beze zbytku a využití dat z uznávaných zdrojů či metod by mělo být samozřejmostí. Nad daty by autor měl vždy provést základní zamyšlení, **jsou-li srovnatelná**.

Jsou porovná-
vána srovnatelná
data?

Často se stává, že student využije data z několika zdrojů, které nato „analyzuje“, ačkoliv různé zdroje využívají různou metodiku. Kupř. testovat, jak míra nezaměstnanosti ovlivňuje počet násilných trestných činů v Evropě, je problematické, definuje-li každá země „násilný čin“ jinak, míra nezaměstnanosti se rovněž vyazuje odlišně, popř. se definice v průběhu sledovaných let měnila.⁷ Podobně, některá data jsou **upravena o inflaci** (vyjádřená v stálých cenách), jiná nikoliv (obvykle byste měli pracovat s reálnými proměnnými, nominální tedy vydělit nějakým cenovým indexem), některá jsou vyjádřena v paritě kupní síly, další ne.

Výhody a li-
mity užití me-
tody?

Ať již je zvolena jakkoliv (ne)sofistikovaná metoda, autor by měl znát její přednosti a limity a interpretovat výsledky podle toho. Na základě sloupcového grafu s dvěma číselnými údaji lapidárně uzavřít, že „politika se neukázala být efektivní“, je možná efektní, ale **bez diskuze o významnosti**

⁶Vypůjčíme-li si od Francise Fukuyami: „Tržní ekonomika a vysoká úroveň bohatství se magicky nezhmotní jakmile „se odstraní stát“; závisí na skrytých institucionálních základech majetkových práv, vlády práva a základního politického řádu.“ (The Origins of Political Order: From Prehuman Times to the French Revolution, New York: Macmillan (2011), str. 13).

⁷Pochopitelně existují ekonometrické nástroje, které po vhodné kalibraci dokáží pracovat i s takto „neporovnatelnými“ daty. Jejich využití však *ipso facto* značí, že si je autor problematičnosti dat vědom.

rozdílu těchto hodnot, mezinárodní či historické komparaci výsledků, směru vývoje, atd. je závěr venkoncem k ničemu.

Při nejčastěji užívané lineární regresi (OLS či „metoda nejmenších čtverců“): $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ také nestačí uvést: „Model ukázal, že $\beta_1 = 0.14$. Potvrdila se tedy má hypotéza“, kteráž věta ukončuje práci. Výsledky (či směr statistické závislosti) by měly být vždy interpretovány a nepostačuje kupř.: „Vztah mezi délkou zákona o daních z příjmu a daňovými úniky vyšel signifikantní“ či „vyšel pozitivně signifikantní“. **Napište prostým jazykem k čemu jste došli:** „Čím delší je zákon o daních z příjmu, tím větší jsou odhadované daňové úniky ($F(2, 234) = 3.42, p = 0.002$).“

Na prvním místě vždy interpretujeme výsledky, na které bylo testování zaměřeno, poté popřípadě výsledky nějakým způsobem extrémní či na základě teorie neočekávané. Vyšla-li elasticita s opačným znaménkem, než měla, je-li diskontní faktor záporný nebo parametr je v řádech odlišný než nalezly další studie a **nenásleduje-li vysvětlení, jde o chybu stejnou, jako kdyby byla interpretace špatně!**⁸ Interpretace by měla nabýt podoby objasnění, proč je parametr významný či nevýznamný, větší či menší než jsme očekávali (ať už na základě dříve diskutované teorie či minulých studií).

Vyšla-li kupříkladu regrese:

$$Y(\text{poměr žáků se zkušeností s konopím}) = 0.2 + 0.14X(\text{roční četnost protidrogových přednášek na škole}) + \epsilon,$$

jaká je její interpretace? ... Protidrogové přednášky vyvolávají růst užívání konopí, po každé přednášce o 14 p. b.⁹ vzroste počet uživatelů marihuany – je možná technicky „správná“, ale fakticky téměř jistě chybná. Lze spíše očekávat, že na školách, kde je užívání drog častější, se budou i frekventovaněji konat protidrogové přednášky. Navíc, obě proměnné budou silně ovlivňovány dalšími faktory, druhem školy počínaje a velikostí sídelní obce konče.¹⁰

Jednoduchá regrese neodhalí kauzální závislost. Zvolená metoda a interpretace výsledků by proto měly vždy respektovat možnost **opačné závislosti či vliv dalších, „opomenutých“ proměnných a nebo zpožděných proměnných** (vliv zkoumaného faktoru se projeví až s odstupem). Nezávězení „dalších“, zásadních proměnných vždy pramenní z nezvládnuté kapitoly přehled existující literatury, v níž nejsou dostatečně diskutovány vlivy důležité pro zkoumanou otázku.

Společenské fenomény jsou spolu komplexně provázány, proto **je extrémně nesnadné nalézt „konečnou příčinu“**. Vždy byste se proto měli ptát, čím mohou být ovlivněny vámi sledované proměnné (či co ony dále ovlivňují). Kupř. mezinárodní analýzy ukazují, že po uzákonění zákazů kouření na veřejných místech klesá počet kuřáků v dané zemi. Zdá se tedy, že zákazy vedou k redukci kouření, že?

Nicméně schválení takového zákona vyžaduje podporu voličstva, ergo většina obyvatel země již začala být nakloněna omezování kouření. Mizení kuřáků je spíše důsledkem změny názorů obyvatel než samotného zákona. Prosazení zákazu v zemi, kde nenastal posun nálad proti kouření, by k žádoucímu výsledku proto vést nemuselo (spíše by vedlo k porušování zákona).

⁸Vycházejí-li parametry nečekaně, nemusí jít nutně o chybu (!), lze-li výsledky stále interpretovat ekonomickou logikou. Nicméně, ve studentských pracích se obvykle o chybu jedná, ať už ve specifikaci regresní funkce, využitím neporovnatelných dat či nezahrnutím podstatné kontrolní proměnné.

⁹Neodlišování procent a procentních bodů je rovněž pravidelné; jde o závažnou chybu. (Srovnejte: Za situace, kdy se míra nezaměstnanosti změnila z 10% na 11%, došlo k nárůstu o 1 procentní bod nebo o 10 % – vidíte, že v hodnotě čísla je rozdíl řádový!)

¹⁰Z těchto důvodů je jednoduchá regrese v řadě problémů společenských věd nevhodnou metodou. Volba řešené hypotézy by měla reflektovat i studentovy analytické (statistické, ekonometrické) schopnosti, aby vůbec byl schopen vhodnou metodu provést a problém zodpovědět.

Co znamená parametr, jakým směrem jde závislost?

Jsou interpretovány hlavní, extrémní a nečekané výsledky?

Korelace v. příčina a následek?

Je zvážena komplexita?

- U obhajoby můžete vždy očekávat otázky na interpretaci různých zjištěných parametrů, možnost obrácení „objevené“ závislosti a nebo na váš odhad změn výsledků, kdybyste do své analýzy zahrnuli i další relevantní proměnné, popř. kdybyste nějaký předpoklad uvolnili či nějaký fixní parametr upravili.¹¹

Data jsou, ať už z jakýchkoliv důvodů, transformována, obvykle logaritmicky (data nemají normální rozdělení, tj. jsou ohraničena nulou, vykazují pozitivní šikmost, atd.). Přistoupí-li autor k **transformaci dat**, měl by její adekvátnost odůvodnit a musí na operaci pamatovat při prezentaci výsledků. Ač se jedná o základy ekonometrie, jsou zejména zde chyby poměrně časté. Tab. 1 uvádí, jak by interpretace výsledků v textu či na obhajobě měla vypadat.

Jak transformace dat ovlivní interpretaci?

Tabulka 1: **Logaritmická transformace proměnných a jejich interpretace**

Specifikace regresní funkce	Interpretace β_1
$Y_i = \beta_0 + \beta_1 \ln(X_i) + \epsilon_i$	„1% změna X souvisí se změnou Y o $0.01\beta_1$ “
$\ln(Y_i) = \beta_0 + \beta_1 X_i + \epsilon_i$	„Změna X o 1 jednotku souvisí se $100\beta_1\%$ změnou Y“
$\ln(Y_i) = \beta_0 + \beta_1 \ln(X_i) + \epsilon_i$	„1% změna X souvisí s $\beta_1\%$ změnou Y“ či „ β_1 je elasticita Y vzhledem k X“

Podobně, upravujeme-li jakkoliv data, od odstranění extrémů po dopočet chybějících hodnot, vždy musí být uvedeno, že k úpravě došlo a její racionální odůvodnění („některé údaje se ukázaly příliš velké“ skutečně nepostačuje). Žádoucí je i zmínka, jak úprava dat ovlivnila výsledky.

Je diskutována úprava dat?

Velmi často se chybí v interpretaci konfidenčních intervalů, obvykle se objevují v pracích formulace typu: „*Daný interval obsahuje skutečnou hodnotu parametru v 95 % případech*“. Je kriticky důležité si uvědomit, že „skutečná hodnota“ parametru β_1 je konstanta, kterou neznáme (a asi nikdy nepoznáme), co zjistíme, je její bodový či intervalový odhad na základě užitého vzorku dat. Jsou tedy vybrány hodnoty Y_i pro odhad regresní rovnice s tím, že my stanovíme kupř. 95% interval spolehlivosti pro skutečnou hodnotu parametru. Takto může být získáno (nekonečně) mnoho intervalových odhadů parametru a 95 % z nich bude pokrývat jeho skutečnou (nám neznámou) hodnotu. Výše uvedená interpretace je proto nesprávná, námi získaný interval nezahrnuje skutečnou hodnotu v 95 % případech, ale **95 % takových intervalů zahrnuje skutečnou hodnotu**.

Co znamená interval spolehlivosti?

V textu jsou obvykle udávány p-hodnoty¹² (p-values) bez další interpretace. P-hodnota znamená pravděpodobnost, že testovací kritérium nabyde stejné hodnoty či extrémnější, za předpokladu že nulová hypotéza platí. Nulová hypotéza (tedy „X na zkoumanou vlastnost Y nemá vliv“) je v ekonomii a společenských vědách obecně (kde „vše souvisí se vším“) často chybná a často ji zamítáme.

Co znamená p-hodnota?

¹¹Protože byste faktory, které jsou pro zkoumaný vztah podstatné, měli jistě v práci diskutovat, cílem otázek je ověřit, zda problematice rozumíte, popř. se očekává zdůvodnění, proč jste jiné proměnné nezahrnuli (nedostupnost dat, jejich pochybná kvalita, atd.). Cílem otázek není vás motivovat k nepodloženým spekulacím, komise otázkou obvykle cílí na nějaký standardní vztah v daném tématu, není proto strategické začít odpověď: „*Já nevím, asi to dopadne takto ...*“.

¹²A měly by být udány coby přesné číslo, nikoliv zastaralé coby p-hodnota je menší než arbitrární hladina významnosti, coby $p < 0.05$ či $p < 0.01$. Testujete-li však mnoho parametrů, využijte v tabulce výsledků úzus hvězdičkového označení významnosti (***, **, * pro hladiny významnosti na úrovni 1%, 5%, 10%; někdy na úrovni 0,1%, 1%, 5%).

P-hodnota je proto spíše znakem, do jaké míry je část užitých dat v souladu s nulovou hypotézou. Nemůžeme-li nulovou hypotézu zamítnout, je možné stále očekávat, že může být nepravdivá, jen pro to nemáme dostatek dat.

- Komise někdy zaměří otázky i na tyto „technikálie“ a velmi často získá odpovědi, že: „*p-hodnota je pravděpodobnost, že nulová hypotéza platí*“, resp. že jde o „*pravděpodobnost, že výsledek je chybný*“ a nebo o „*pravděpodobnost chybného zamítnutí nulové hypotézy*“. Všechny tyto odpovědi jsou chybné.

V každém případě nebuďte bezmyšlenkovými otroky p-hodnot. Zda je vztah „významný“ je arbitrárně (byť konsenzuálně) určeno, zároveň i způsoby stanovení p-hodnot mají své limity. Především však není natolik důležité, zda je vztah statisticky významný, ale i **nakolik silný efekt má v realitě**. Zvažte, že fáze Měsíce mají statisticky významný vliv na ceny akcií na burzách (p-hodnoty testů jsou nízké), nicméně reálný efekt – dopad na změny cen – jsou nanejvýše desetiny procenta. Řádově významnější proměnné ovlivňující ceny jsou pochopitelně jiné (očekávaná ziskovost firmy, ad.).

V hospodářské, sociální politice, ale i v kampaních firem se obvykle sleduje vliv nějakého opatření na zvýšení **nebo** snížení sledovaného ukazatele, nikoliv na jeho změnu. Testuje se proto explicitní jednostranná hypotéza, kupř. „*po snížení daní z lihu se zvýší počet sebevražd*“, nikoliv „*snížení daní z lihu ovlivní počet sebevražd*“, měly by být proto užívány jednostranné testy. Používáním dvoustranných se zbytečně ztrácí síla testu. Pokud by se počet sebevražd snížil, je naše hypotéza prostě chybná, stejně jako kdyby se počet sebevražd nezměnil vůbec.

V případě užití vlastních dotazníkových a podobných šetření, musí být diskutováno, jak charakter (či počet) respondentů a jejich výběr ovlivnil dosažené výsledky. Jelikož studenti téměř nikdy nedosáhnou relevantní reprezentativnosti svého vzorku (jejich rodina a přátelé obvykle nejsou reprezentativním vzorkem), je podobná diskuze nezbytná! Je zbytečné realizovat vlastní šetření, jsou-li volně dostupná kvalitnější data.

Než začnete s vlastním shromažďováním dat, **ověřte, zda-li není možné využít již existující data!**

Nejde jen o p-hodnoty!

Je nutný dvoustranný test?

Je vzorek reprezentativní?

Nejsou dostupná lepší data?

6 Závěr – formalistní a neobjevný

„Dějiny chyb lidstva jsou cennější a zajímavější než dějiny jeho objevů. Pravda je uniformní, neměnná, úzká, stále existuje a nezdá se, že by vyžadovala mnoho aktivní síly, než pasivní schopnost duše se s ní potkat. Ale chyba je nekonečně rozmanitá, není reálná, je ryzím a prostým dílem mysli, která ji tvoří. Zde má duše konečně prostor k rozmachu, zde ukáže své bezmezné schopnosti a všechny její krásné a zajímavé výstřednosti a absurdity.“

Benjamin Franklin

Závěry¹³ studentských kvalifikačních prací často duplikují obsah úvodu a někdy obsahují i doslovné kopie vět z předchozích částí. Obé je, pochopitelně, chybou. V závěru krátce shrňte hlavní či

Není závěr zbytečný?

¹³Závěr (Conclusion nebo Final remark) nezaměňujte se „závěry“, tj. vašimi výsledky či jejich interpretací. Tyto se nacházejí v předchozích sekcích vaší práce – vizte doporučenou strukturu uvedenou třeba v dokumentu **Vzorová struktura semestrálních a kvalifikačních prací**.

nečekaná zjištění, zejména ale **představte aplikaci výsledků**, doporučení pro firmu, obec, hospodářskou politiku či pro další výzkum. Měly by zde být stanoveny limity práce (vycházíte-li z dat v ČR, téměř vždy tím bude krátká časová osa) a nebo její přednosti – to, co ji odlišuje od ostatních podobných děl. Závěr by měl nabídnout směry, kam by další zkoumání mělo směřovat – v závěru je dovoleno odborně (!) spekulovat.

- Neobsahuje-li závěr výše uvedené, komise se na ně téměř jistě optá. Vždy je varovné, není-li student schopen odpovědět, jak lze jeho práci využít (jak či kam posunuje poznání), popř. jak by práci rozšířil, vylepšil či v čem vidí její úzká místa.

Víte-li, že jste prozkoumali jen část složitého fenoménu či využívali ne-až-tak-kvalitní data, popř. vámi zjištěné parametry jsou nízké s marginální hladinou významnosti, buďte ve svých formulacích opatrní, využívejte spíše obrátů „zdá se“, „vypadá“, „naznačuje“, „může“ než „prokázáno“, „dokázáno“, „je (nade vší pochybnost) zřejmé“.

Buďte pokorní.