

2 Metodologie a použité metody analýzy

V sekundární analýze pracujeme s několika rozsáhlými datovými soubory. Nejvíce je využíván národní datový soubor z mezinárodního šetření TIMSS 2015, který zahrnuje celkem 159 českých škol a 5 202 žáků 4. ročníků vybraných z těchto škol. Pro mezinárodní srovnání využíváme kompletní mezinárodní datový soubor TIMSS 2015. Datové soubory mezinárodního šetření TIMSS 2015 byly dodatečně propojeny s národními daty shromažďovanými Českou školní inspekcí při hospitačních činnostech či realizaci inspekčního elektronického zjišťování pomocí systému InspIS DATA.¹⁰ V rámci přípravy dat pro zpracování sekundární analýzy bylo kvalitativně i kvantitativně posouzeno celkem 478 formulářů (datových souborů) z hlediska celkového počtu záznamů, celkového počtu otázek, úrovně agregace, úrovně dotazování, času dotazování a obecné míry využitelnosti pro další analytickou činnost ve zvolených zájmových oblastech sekundární analýzy. U souborů dle zmíněných parametrů vhodných pro párování s daty národního šetření se dodržovalo pravidlo co nejvyššího procenta napárovatelných dat a co nejbližší časové shody sběru národních dat i dat mezinárodního šetření TIMSS 2015. Soubor s nejnižší hodnotou napárování využitý v hlavních analýzách se v 96,9 % případů shodoval s daty škol, které se zúčastnily šetření TIMSS 2015 s časovou odchylkou max. 12 měsíců od data realizace testování TIMSS 2015. Navíc byly pro analýzu vybírány takové kontextuální proměnné, které se v krátkém časovém horizontu příliš nemění.

V rámci sekundární analýzy dat využíváme tradiční metody pro statistickou analýzu v pedagogických a společenských vědách. Mezi tyto patří deskriptivní statistiky s tříděním prvního a druhého stupně, které poskytují souhrnné informace o dílčích proměnných a jsou tak základem pro pokročilejší analýzy. Ve velké míře jsou využívány grafickou formou znázorněné t-testy¹¹, sloužící k porovnávání statistické podobnosti průměrů vždy mezi dvěma skupinami žáků (například mezi žáky disponujícími domácími zdroji pro výuku a žáky nedisponujícími těmito zdroji). Z pokročilých statistických metod využíváme hierarchické regresní modelování na dvou úrovních. Hierarchické regresní modely jsou namísto jednoduché lineární regrese využívány s ohledem na hierarchickou strukturu analyzovaných datových souborů (stát → škola → třída → žák), a to z toho důvodu, že jednoduchá lineární regrese nemusí být schopna přesně zachytit vztahy v takové hierarchické struktuře dat. Hierarchické regresní modely se běžně používají v edukačních vědách. Modely byly testovány jak s náhodnou konstantou na úrovni třídy, tak školy. V textu jsou zobrazeny ty výsledné modely, kde je druhou úrovní škola. Složitější vztahy testované regresními modely jsou interpretovány mj. za pomoci vizuálního znázornění interakčních efektů. Veškeré analýzy jsou prováděny s odpovídajícím vážením dat. Pro přípravu, kódování, propojování a základní analýzu dat byly využity softwarové programy IDB Analyzer a IBM SPSS Statistics. Pro pokročilejší analýzy a hierarchické regresní modelování byly využity programy Stata, R a Mplus. Příslušné metodologické a odborné statistické pojmy nalezne čtenář v příloze, stejně jako vysvětlení interpretace grafů a regresních modelů.

10 InspIS DATA je vlastní informační systém České školní inspekce, který slouží pro elektronický sběr dat. Do systému jsou vkládány formuláře různého typu, jež souvisí s hodnocením a monitorováním vzdělávání, které realizuje Česká školní inspekce. Nejčastěji zde najdeme záznamy o inspekční činnosti prováděné na školách, dále například hodnotící záznamy škol a školských zařízení, záznamy o školních úrazech apod. Podrobněji k jednotlivým informačním systémům využívaným Českou školní inspekcí viz web (ČŠI).

11 Alternativou je graf typu boxplot, který zobrazuje interkvartilové rozpětí, odlehle případy, průměr a medián. Z důvodu přehlednosti a srozumitelnosti je uvedena vizuálně jednodušší grafická forma.