Z ekonomické literatury

NESTÁRNOUCÍ KNIHA UPOZORŇUJÍCÍ NA MOŽNOSTI MANIPULACE S FAKTY PROSTŘEDNICTVÍM EXAKTNÍCH METOD

Jaroslav Daňhel, Vysoká škola ekonomická v Praze

Darrell Huff: Jak lhát se statistikou – statistika vtípně a jinak

Praha : Brána, 2013. 189 stran. ISBN: 978-80-7243-623-1

S recenzovaným bestsellerem D. Huffa se čeští čtenáři setkávají poprvé, ačkoli byl vydán již v roce 1954, přeložen do většiny světových jazyků a prodalo se ho již přes půl milionu výtisků. Načasování opožděné edice knihy v českém jazyce nepochybně souvisí s celkovou změnou sentimentu odborné veřejnosti vůči dosti úporným, avšak z metodologického hlediska kontraverzním snahám hlavního teoretického proudu posunout společenskou ekonomickou vědu její matematizací k vyšší rigoróznosti a normativní predikci. Tento směr se svým zenitem v devadesátých letech a s demiurgem prof. Samuelsonem v čele začal v poslední dekádě významně ztrácet příznivce. Děje se tak zejména v souvislosti se selháváním matematických modelů na podporu rozhodování v ekonomických úlohách, zejména regulatorních projektů při identifikování a indikování předkrizových nerovnováh v minulé dekádě. Pro současnou ekonomickou vědu je symptomatické, že při hledání nových paradigmat reflektujících současné složité globální prostředí ji v poslední době nejvíce obohatili představitelé empirických a psychologických přístupů k ekonomii a zejména pak k rozhodování ekonomických

subjektů, významná jsou v tomto ohledu díla N. N. Taleba a D. Kahnemana, osobností moderního nekonvenčního myšlení, kteří ovšem na druhé straně mají blízko ke statistice a statistickému vidění světa.

Šedesát let stará a nestárnoucí publikace je opatřena úvodem a poznámkami ekonomky Evy Zamrazilové, které ji erudovaně a čtivě uvádějí do souladu se současným složitým globálním prostředím a talebovským děním „na trzích a v životě“. Zamrazilová vcelku logicky označuje Darrella Huffa za předchůdce těchto moderních přístupů, jejichž smyslem je oprostit matematiku a statistiku, tedy vědní disciplíny, které patří neoddělitelně ke struktuře moderního myšlení, od balastu aplikačních modelů a úloh, které tyto nedialektické vědní disciplíny nemohou ze své podstaty vyřešit. Na druhé straně mohou významně pomoci řešení ekonomické problémy vhodně strukturovat a vytvořit pro ně adekvátní logický rámec, ve kterém je možno kvantitativní výsledky správně kvalitativně interpretovat.

Kniha je rozdělena do deseti kapitol, každá z nich upozorňuje na jedno z poměrně často se vyskytujících možných úskalí, skrývajících se ve vybraném statistickém vzorku,

ve zvoleném statistickém průměru, v chybějících vysvětlivkách, v sugestivních grafech, chybně interpretované kauzality a dalších exaktně se tvářících institucích, které nás při našem úsudku o pravdivých stavech světa mohou vést na scestí.

V prvé kapitole Huff otevírá otázku reprezentativnosti statistických výběrových šetření. Upozorňuje, že vychýlenost vzorku pro tato šetření, ze kterého posléze provádíme zobecňující závěry, může být velmi snadno narušena. Je tudíž nezbytné při interpretaci výsledků zachovat si určitou míru pochybnosti, jestliže možnost vychýlení existuje a Huffova výstraha praví, že tato možnost existuje vždycky. Získání čistě náhodného vzorku, který jediné může být zkoumán statistikou s naprostou důvěrou, je většinou složité a drahé, takže zejména jeho nákladnost nutí k použití levnější, ale také méně věrohodný stratifikovaný výběr. Statistické průzkumy nemusejí být tudíž nutně rovnou zmanipulované za účelem vytvoření falešného dojmu, vzorek může být ze své podstaty nereprezentativní a tudíž může výsledek a tím i interpretaci zkreslit automaticky. Eva Zamrazilová dodává: Solidní zpráva o statistickém výběrovém šetření by měla vždy uvádět, jak vysoká je pravděpodobnost přibližné správnosti předkládaného výsledku, tedy statisticky řečeno v jakém intervalu spolehlivosti se výsledek pohybuje.

Druhá kapitola s názvem Co nám průměr zatajuje, se zabývá možnostmi zkreslování kvantitativní úrovně zkoumaných jevů: do kategorie středních hodnot zkoumaných statistických souborů patří více skalárů: aritmetický a geometrický průměr a také modus a medián, každá z těchto středních hodnot má jinou vypovídací hodnotu a je třeba při deklaraci výsledků statistického průzkumu korektně uvést, o který typ střední hodnoty se jedná. Třetí a čtvrtá kapitola tento problém dále rozvádějí: organizátoři průzkumů by měly vždy správně charakterizovat podmínky,

za kterých výběrové šetření probíhá. Přitom je třeba rozlišovat mezi úmyslnou, ev. neúmyslnou systematickou chybou, která je imanentní vadou průzkumu a chybou statistickou, která je vlastností metody výběrového šetření, je však změřitelná.

Kapitola pátá se věnuje možné dezinterpretaci faktů pomocí grafů. Na četné konkrétní ukázky autora publikace, ze kterých jednoznačně vyplývá, že k jednoznačně vyhlížejícím trendům na obrázcích je vhodné obstarat si další informace, navazuje sevržený komentář E. Zamrazilové: zejména s pomocí moderní techniky „dokážeme vytvořit graf, který skutečně ohromí a dokáže vyvolat tu představu, jakou chceme, i když původní surová data nám do karet moc nehrají. Posun měřítko osy, maxima či minima, může dokázat pravé divy.“ Toto téma pokračuje ještě v šesté kapitole zabývající se metodou neproporcionálního a tím pádem i sugestivního, na druhé straně ovšem i nenápadného zvětšování či zmenšování grafických symbolů podle toho, kterou stránku zobrazovaného jevu chceme náležitě zvýraznit. Obdobné téma probírá i kapitola sedmá: pojednává o jevu, který se dá nazvat „klonování statistiky“ (Huff kapitolu nazval anglickou slovní hříčkou *semiattached figures*), jde o případy, kdy jsou provedeny sice exaktní propočty, leč z nesprávných dat, výsledek se pak ovšem interpretuje jako naprosto věrohodné zjištění.

Poslední tři kapitoly knihy komentátorka Eva Zamrazilová považuje za klíčové. Osmá kapitola se věnuje problematice kauzality, které jsou senzitivním tématem filosofie počínaje a statistikou konče. Taleb zdůrazňuje, že my lidé jsme náchylní k přílišné interpretaci a zálibě v celistvých uzavřených příbězích a tedy i přeceňování příčin a následků (klam narativity). E. Zamrazilová v tomto duchu vidí příčinu sklonu lidí k narativitě v lidské obraně proti nejistotě: vše se snažíme vtěsnat do funkčních příběhů. Vzpomíná filosofa Davida Huma, který v pozici absolutního

skeptika kauzalitu mezi jevy zcela odmítal. Hume rovněž zpochybnil efektivnost statistické indukce v případě, že dojde k zásadní změně podmínek v časové řadě, tedy jinými slovy, kdy poznání, které fungovalo v minulosti, jednoho dne fungovat přestane, a to co jsme se dřív naučili, se ukáže v lepším případě jako irrelevantní a chybné, v horším případě jako nebezpečně zavádějící (Taleb: z minulosti nemůžeme získat více informací, než kolik jich dokáže poskytnout). Huff, abychom nepodlehli klamu následné příčiny, doporučuje každé tvrzení o příčinách a následcích podrobit důkladnému zkoumání, neboť se může jednat o souvislost některého z následujících typů: vztah vznikne náhodou, nebo vztah existuje, ale nelze určit, která z veličin je příčinnou a která důsledkem, anebo jevy se vůbec vzájemně neovlivňují, ale přesto tu existuje statistická korelace. Zamrazilová k tomu dodává: „V žádné jiné oblasti není statistik v tak vysokém riziku chybné interpretace svých výsledků jako při hledání souvislostí mezi statistickými znaky či jevy. Jestliže existuje statistická souvislost, neznamená to nutně, že jde o příčinnou věcnou souvislost“ (str. 136).

V deváté kapitole neortodoxně nazvané Jak statistikulovat se Huff zabývá otázkou, do jaké míry se dá dezinformování veřejnosti s využitím statistických materiálů považovat za konání s cílem klamat a odpovídá: může jít o záměr, ale i o neschopnost zpracovávatelů se závěrem, že si není jistý, která z těchto možností je pro statistiky méně hanlivá (Zamrazilová: je celkem jedno, jestli se dezinformování dopustil „nevinný břídl“ nebo „lstivý manipulátor“). Huff v knize uvádí velké množství konkrétních příkladů evidentní „statistikulace“, mezi ně počítá i vytváření falešného nádechu přesnosti i pro nepochybnější statistiky, kterým punc závažnosti dodá desetinná čárka ve výsledku. Pro nás je obzvláště úsměvné jeho doložení, že i Karel Marx podlehl pokušení stejného stříhu, když kalkuloval v ceně konkrétního výrobku míru nadhodnoty při deseti hodinové pracovní

době na základě údajů „přadláka z Manchesteru“ a přidanou hodnotu ze sbírky předpokladů, odhadů a zaokrouhlených čísel spočítal „přesně“ na 6 celých a dvě třiatřicetiny libry, tyto dvě třiatřicetiny vyvolávají příjemný pocit velké přesnosti, přestože jde evidentně o Marxův blaf (str. 147).

Závěrečná desátá kapitola je D. Huffem na rozdíl od předchozích devíti kapitol pojata veskrze pozitivně. Autor v ní formuluje základních pět otázek, které by si měl uživatel položit a v případě pozitivních odpovědí rozpoznat zdravá a použitelná statistická data. Otázky jsou tyto: 1) kdo to říká? 2) jak to ví? 3) co chybí? 4) nezměnil někdo předmět? 5) kardinální otázka: dává to smysl?

Touto kapitolou se prakticky vracíme na začátek našich úvah: Huff v úvodu poslední kapitoly konstatuje, že „statistické informace ... nemohou být prověřovány se stejnou spolehlivostí, s jako se provádí chemická analýza nebo postupy v testovacích laboratořích (str. 163). Jinými slovy řečeno, matematika a statistika nám mohou dát na základě jim vlastních metod jen takové informace, jaké jim dovolí zpracovat a interpretovat měkká společenská věda – ekonomie. Navíc pro statistické použití musí být pro činění závěru použita zdravá a použitelná data. Tento požadavek vypadá na první pohled triviálně, nicméně devět desetin Huffovy knihy se nás snaží upozornit na hojně dezinterpretace a provést všemi úskalími, na která můžeme při vyhodnocování statistických výstupů narazit. E. Zamrazilová k tomu dodává: Huff nám názorně ukazuje, jaký nesmysl či paskvil může vzniknout, používáme-li statistická data či metody bez logiky. Bez pevného logického rámce nemá interpretace žádných statistických vazeb de facto smysl.

Pokud si Huffovu knihu a inspirativní poznámky E. Zamrazilové přečtete, a to vám vřele doporučujeme, položíte si pět otázek a na základě svých odpovědí pozitivně za dohledu zdravého rozumu identifikujete seriózní a korektní data, pak vás statistika určitě neobelže.