

# Klasická testová teorie reliability v metodologii výběrových šetření\*

Jan Schubert\*\*

Institut sociologických studií

Fakulta sociálních věd, Univerzita Karlova v Praze

## The Classical Test Theory Concept of Reliability and the Methodology of Sample Surveys

**Abstract:** *This article outlines how different types of survey response reliability are estimated and used within the social sciences. Implementing the Classical Test Theory's concept of reliability in a panel survey exploring opinions and attitude crystallization in a Czech town. Estimates of the test-retest reliability of respondents' political orientation are estimated using four different approaches. Political orientation is measured using a standard 10 point left-right scale. It also estimates the 'true score' of a respondents' political orientation; and shows the implications of measurement precision for the estimation of this individual level trait. Inter-item reliability is calculated using a 4 item scale composed of dichotomous indicators. This analysis reveals that the estimated reliability of the political orientation scale depends on the number of items used to construct the scale. More concretely the 10 point left-right political orientation scale exhibits a satisfactory level of test-retest reliability demonstrating its validity in making inferences about the target population. The analyses presented in this study were estimated using the R statistical programming language and LISREL – a specialized software package for estimating structural equation models.*

*Data a výzkum - SDA Info 2010, Vol. 4, No. 2: 105-122.*

(c) Sociologický ústav AV ČR, v.v.i., Praha 2010.

---

\* Publikace článku byla podpořena Ministerstvem školství, mládeže a tělovýchovy v rámci projektu COMPDAT – Zdroje dat, výzkum standardů, kvality dat a metody harmonizace dat pro mezinárodní sociální komparativní výzkum a integraci do sítě CESSDA sponzorovaného MŠMT pod reg. č. LA09010. Autor děkuje oběma anonymním recenzentům za připomínky, které pomohly zvýšit kvalitu této stati.

\*\* Veškerou korespondenci posílejte na adresu: Jan Schubert, e-mail: honza.schubert@gmail.com.

## 1. Teorie reliability

### *Význam reliability*

Posudme následující příklad, v němž chceme změřit určitý znak osoby, který není přímo pozorovatelný (např. její znalost v určité oblasti). Chceme-li změřit dějepisnou znalost respondenta, navrhneme test a následně budeme znalost měřit jako schopnost respondenta obstát v tomto testu (řekněme počtem získaných bodů). Pokud navrhneme takovéto testy dva a necháme osobu, aby na všechny otázky odpověděla, a poté sečteme zvlášť výsledky obou testů, dojdeme nejspíše ke dvěma odlišným hodnotám, ačkoliv se vlastnost, kterou hodláme měřit (znalost dějepisu), nezměnila. Rozdíly budou nejspíše způsobeny chybami v měřicím nástroji (testu).

Tento příklad ukazuje, jak důležité je znát velikost chyby, kterou do měření vnáší měřicí nástroj. Chyba nastává při každém měření a užitečnost měřicí techniky (a odpovídajícího nástroje) spočívá do značné míry v její schopnosti přinášet přesné (stabilní) hodnoty. Každé měření je zkruseno kontextem (např. teplotou, atmosférickým tlakem nebo – v sociálněvědních šetřeních – náladou respondenta). Ve fyzice není změřená chyba nástroje obtížná, neboť je možné normalizovat, kontrolovat a mnohokrát opakovat podmínky měření. Odpovídající kontrola všech efektů v sociálních vědách není možná [Řehák 1968: 638]. Potřeba kvantifikovat chybu měřicího nástroje (dotazníku, testu) však zůstává. Reliabilita vyjadřuje, do jaké míry je měření konzistentní. Dostává-li výzkumník při každém použití měřicí metody velmi odlišné výsledky, je zřejmé, že je pro něj takto nepřesné měření nepoužitelné, neboť nemůže vědět, který výsledek se nejvíce blíží skutečné hodnotě.

Neodhadnutí reliability měření vede ke zřejmým důsledkům. Výzkumník může přijmout rozhodnutí o určitém znaku na základě výsledků, které neodrážejí skutečnou hodnotu tohoto znaku. I ty nejpropracovanější statistické metody dávají nepřesné výsledky, pokud jsou analyzovaná data zkrusena.

### *Chyba a skutečná hodnota*

V sociálních vědách obecně chceme měřit určitý znak, který není přímo pozorovatelný (např. sociální integraci, politickou orientaci, vyjadřovací schopnosti osoby). Tomuto znaku říkáme koncept. Jak jsme již uvedli, koncept obvykle není přímo měřitelný<sup>1</sup> a často je také vysoce abstraktní. Konkrétněji dochází k tomu, že v procesu konceptualizace rozvíjíme konstrukty. Definujeme určitou sadu položek (otázek), o nichž se domníváme, že souvisejí s konstruktem odrážejícím daný koncept. Věříme, že hodnoty pozorované u těchto položek představují hodnotu latentní proměnné. Měřitelnou hodnotu označíme  $X$ , hodnotu latentní proměnné  $T$  a chybu měření  $E$ .

---

1 Někde jej nazýváme latentní proměnná.

Pozorovaná hodnota  $X$  může být považována za funkci  $T$  a  $E$ , například v této podobě [Lord, Novick 1968: 34]:

$$X = T + E \quad (1),$$

kde pozorovaná hodnota  $X$  závisí na hodnotě skutečné (hodnotě latentní proměnné) a na chybě měření. Jak zdůraznil Lazarsfeld [Lazarsfeld, Merton 1950], chyby měření se mohou měnit v závislosti na kontextu (dočasný stav dané osoby, podmínky testování apod.). I pozorovaná hodnota  $X$  se tedy může měnit také v závislosti na kontextu. Z toho vyplývá, že pozorovaná hodnota  $X$  je jedna z možných pozorovaných hodnot  $X$  a závisí na nějakém nahodném procesu. Můžeme proto  $X$  považovat za náhodnou veličinu s určitým pravděpodobnostním rozdělením [Lord, Novick 1968: 29–30].

Klasická teorie testu spočívá na dvou předpokladech. Prvním z nich je, že skutečná hodnota  $T$  se mezi jednotlivými případy měření ( $M_1, M_2 \dots M_k$ ) nemění, a druhým, že měření jsou lineárně experimentálně nezávislá<sup>2</sup> [Lord, Novick 1968: 45]. Z těchto předpokladů můžeme odvodit tři základní východiska:  $E(E) = 0$ ,  $r(E, T) = 0$ ,  $r(E_1, E_2) = 0$  [důkaz viz Lord, Novick 1968: 37]. Ty znamenají, že očekávaná hodnota chyby měření je rovna nule, chyba nezávisí na skutečné hodnotě a chyby u dvou různých měření jsou nezávislé. Obecně Lord a Novick [1968] uvádějí, že chyba měření je skutečně nezávislá a její variabilita je náhodná.

Dalším klíčovým pojmem klasické teorie testů jsou „paralelní měření“. Dvě měření jsou paralelní, pokud měří ten samý konstrukt. Může při nich být použito stejné měřicí metody (např. stejných otázek) nebo metod odlišných. Předpokládá se, že znak se během těchto měření ( $X_1$  a  $X_2$ ) nezmění, tedy:  $E(X_1) = E(X_2)$ ,  $\text{var}(X_1) = \text{var}(X_2)$ ,  $r(X_1, Z^3) = r(X_2, Z)$  a  $r(X_1, X_2) = r(X_2, X_3)$ , pokud je k dispozici více měření [Lord, Novick 1968: 59].

Reliabilitu můžeme považovat za tu složku pozorované hodnoty, která není ovlivněna chybou měření, tedy

$$\text{rel}(X, T) = \text{var}(T)/\text{var}(X) \quad (2).$$

Je-li tato složka měřené hodnoty shodná s hodnotou skutečnou, pak měřená hodnota přesně odpovídá hodnotě skutečné. Dále

$$\text{rel}(X, T) = [\text{var}(X) - \text{var}(E)]/\text{var}(X) = 1 - \text{var}(E)/\text{var}(X) \quad (3).$$

Je-li tato složka rozptylu chyby měření nízká v porovnání s měřenou hodnotou, považujeme měřenou hodnotu za dobrý odhad skutečné hodnoty  $T$ . Funkce  $X = T + E$  má lineární povahu, a tak můžeme obě nezjištěné složky vzorce odhadnout prostřednictvím měření jejich lineárního vztahu. Stačí nám odhadnout je z paralelních měření stejného znaku (stejně latentní proměnné

2 Dvě měření ( $M_1$  a  $M_2$ ) považujeme za lineárně experimentálně nezávislé, pokud jsou jejich očekávané hodnoty shodné, vezmeme-li v potaz druhé měření:  $E(x_1|x_2) = E(x_1)$  a  $E(x_2|x_1) = E(x_2)$  [Lord, Novick 1968: 45–46].

3 Z představuje třetí proměnnou.

nebo stejné skutečné hodnoty T). Pokud (2) a obě metody (otázky či položky) měří stejnou latentní proměnnou (stejnou skutečnou hodnotu), pak

$$r(X_1, X_2) = \text{cov}(X_1, X_2) / s(X_1) \cdot s(X_2) \quad (4).$$

Kovariance (X<sub>1</sub>, X<sub>2</sub>) je rozptyl skutečné hodnoty T, a proto

$$r(X_1, X_2) = \text{var}(T) / \text{var}(X) \quad (5),$$

což odpovídá výše definované reliabilitě. Pearsonův korelační koeficient mezi oběma paralelními měřeními představuje odhad reliability těchto měření [Lord, Novick 1968: 55–61].

### **Typy reliability**

Existují dva typy reliability: reliabilita test-retest a reliabilita mezipoložková. První typ je popsán výše; hlavním principem je srovnání pořadí pozorovaných hodnot<sup>4</sup> získaných měřeními provedenými ve dvou různých okamžicích. Druhý typ reliability se zaměřuje na odhad reliability odpovědí osoby na sadu položek (např. otázek) najednou. Konstruujeme otázky (položky) k měření stejného znaku (stejně latentní proměnné nebo stejné skutečné hodnoty), přičemž předpokládáme, že respondenti budou na tyto otázky odpovídat podobně. Oba typy reliability nám mají umožnit odhadnout, jak dobře dané otázky měří to, co měřit chceme, nicméně nacházíme pro ně různá uplatnění. Mezipoložkovou reliabilitu typicky uplatňujeme při odhadu reliability škálových položek, zatímco reliabilitu test-retest používáme při odhadu reliability jednotlivých otázek.

Výše zmíněný problém teorie reliability spočívá v tom, že reliabilitu nemůžeme změřit přesně. Její hodnotu můžeme přinejlepším odhadovat. Existují různé přístupy k odhadu hodnoty obou typů reliability. K odhadu reliability test-retest potřebujeme alespoň dvě paralelní měření. První odhad představuje základní přístup s použitím *korelačního koeficientu*. Reliabilitu dvou měření jsme popsali jako složku rozptylu, kterou mají obě měření společnou.<sup>5</sup> Můžeme oprávněně předpokládat, že rozptyly obou měření (pozorované proměnné) se vztahují k jediné latentní proměnné, a provést *konfirmační faktorovou analýzu (CFA)*. Můžeme chybu rozptylu měřených hodnot odhadnout jako podíl rozptylu, který není vysvětlen latentní proměnnou, a testovat, zda je tento model statisticky oprávněný [Lord, Novick 1968: 216].<sup>6</sup> Odhadneme-li pouze jednofaktorový model, ve kterém se nevysvětlí jedinečnost rovná rozptylu chyby, pak můžeme reliabilitu odhadovat pomocí *komunalit* [Lord, Novick 1968: 536]. Poslední odhad reliability používá jinou logiku. Nepracuje primárně s korelacemi mezi měřeními, nýbrž s *odhadem*

---

4 Které obdržíme od stejných respondentů.

5 Protože  $r(E_1, E_2) = 0$ , je tato společná složka rozptylu rovna skutečnému rozptylu hodnoty.

6 Odhadovaný rozptyl bude nejspíše roven čtvrtci faktorových zátěží, v závislosti na zvoleném modelu.

*rozptylu skutečné hodnoty*. Připomeňme, že  $\text{rel}(X,T) = \text{var}(T)/\text{var}(X)$ , a tedy můžeme-li odhadnout rozptyl skutečné hodnoty, můžeme reliabilitu vypočítat pomocí výše uvedeného vzorce. Lze oprávněně předpokládat, že existuje jedna latentní proměnná (skutečná hodnota), která vysvětluje významnou složku rozptylu hodnot měřených proměnných. Rozptyl tohoto faktoru (skutečné hodnoty) můžeme odhadovat pomocí CFA. Výhodou tohoto přístupu je, že nepředpokládá paralelnost měření [Lord, Novick 1968: 216–219]. Všechny čtyři přístupy by měly vést k odlišným (ale podobným) výsledkům.

V metodologii každodenních výzkumů je obvykle častěji odhadována reliabilita druhého typu (mezipoložková). Průkopníky tohoto přístupu byli Spearman a Brown [1910]. P. F. Lazarsfeld později ve Stoufferově *Americkém vojákovi* zformuloval a dokázal model latentní struktury [Lazarsfeld, Merton 1950]. Doložil, že měření určité latentní proměnné (latentního znaku) pomocí více otázek či položek vede k přesnějším výsledkům. Většinou pro své analýzy používal binární položky (ano-ne, plus-minus odpovědi). Jeho dílu předcházela práce L. Guttmana s klasickými škálami (odtud Guttmanovy škály). Lazarsfeldův model latentní struktury byl později rozšířen do analýzy latentních tříd [Jeřábek a Soukup 2008]. Pozdější výzkum mezipoložkové reliability byl silně ovlivněn Lee Cronbachem, jehož vzorec pro odhad mezipoložkové reliability (Cronbachovo alfa) je zřejmě v dnešních sociálních vědách používán nejčastěji.

Hlavní myšlenkou tohoto typu odhadů je, že pokud položky měří stejnou latentní proměnnou, měla by osoba na tyto položky odpovídat podobně, a proto by odpovědi měly být navzájem korelované. Pro odhad mezipoložkové reliability existuje řada vzorců, např. metoda „split-half“ (Spearmanův-Brownův vzorec), Cronbachovo alfa, McDonaldsovo omega, Revellsovo beta, Guttmanovo lambda. V naší analýze máme pouze baterii binárních položek, a proto budeme reliabilitu odhadovat pomocí *Kuderova-Richardsonova vzorce 20* [Ferjenčík 2000: 203] a *metody „split-half“*.<sup>7</sup> Vlastnosti baterie rovněž otestujeme pomocí *analýzy latentních tříd (LCA)*.

Dalším přístupem k odhadu reliability měření je metoda *paralelních měření ve stejném dotazníku* [Ferjenčík 2000: 202–203]. Spočívá v porovnání dvou měření stejného znaku různými metodami. Celkovou škálu složíme ze čtyř binárních položek dotazníku a vypočteme korelační koeficient s desetibodovou levo-pravou škálou měřící pozici osoby na levo-pravém kontinuu (podrobnosti viz následující oddíl). Výhodou tohoto přístupu je, že se nemusíme zabývat změnou skutečné hodnoty u každé osoby během zkoumaného časového úseku, neboť obě hodnoty jsou změřeny ve stejném čase. Zjevnou nevýhodou je, že obě měření by měla měřit stejnou latentní proměnnou (a rovněž stejné dimenze), což často v sociálních vědách neplatí.

---

7 Výpočet pomocí Spearmanova-Brownova vzorce. Korelaci mezi oběma rozdělenými částmi otázek změříme pomocí Spearmanova koeficientu.

## 2. Analyzovaná data a proměnné

Naše data (pracovně nazývaná „klatovská data“) pocházejí z panelového šetření krystalizace názorů a postojů ve městě Klatovy. Panelové šetření proběhlo ve čtyřech vlnách (více podrobností v tabulce P1 v příloze a ve studii od Jeřábka [1999]).

Na tomto místě bychom rádi rozlišili pojmy názoru, postoje, hodnoty a přesvědčení. Postoje můžeme chápat jako stabilnější druh životní orientace jednotlivce. Naproti tomu názory mohou být konceptualizovány jako soubor krátkodobých nálad, ovlivněných trendy [Šubrt 2006]. Postoje i názory nejspíše podléhají změnám v delším časovém období. Na rozdíl od nich hodnoty a přesvědčení můžeme chápat jako normativní systém jednotlivce, který vzniká od raného dětství v procesu socializace. Jedním z předpokladů klasické teorie testů je, že skutečná hodnota musí zůstat konstantní. Je obtížné nalézt takový sociální znak, který se v čase nemění.

V provedeném šetření jsme pro odhad různých typů reliability zvolili příklad levo-pravé politické orientace jednotlivce. Vágní pojmy „pravcový“ a „levicový“ nemají v žádném případě absolutní platnost, a mnozí lidé je chápou subjektivně [viz Vinopal 2006]. Toto subjektivní chápání může být zdrojem chyb, neboť zkresluje intersubjektivní srovnatelnost měření. Levo-pravou politickou orientaci chápeme jako postoj, který je krátkodobě stabilní, ale může se měnit v delších úsecích života a může být ovlivněn vnějšími faktory, jako je politická situace, obecné trendy ve společnosti atd.

Levo-pravou politickou orientaci lze chápat jako druh postoje, který je podmnožinou celkové politické orientace jednotlivce. Samo levo-pravé kontinuum může být multidimenzionální, nebo jednodimenzionální. Dimenzionalita kontinua je sama o sobě složitý problém, který překračuje rámec této stati [více viz Vinopal 2006].

Uveďme stručně, že v homogenních společnostech je kontinuum s větší pravděpodobností jednodimenzionální; tato dimenze obvykle odráží dimenzi ekonomickou; a lidé jsou obvykle schopni se na takovémto kontinuu snadno zařadit. V heterogennějších společnostech může být dimenzí více, v závislosti na počtu relevantních veřejných otázek. Jako příklad další dimenze je obvykle uváděno náboženství. Náboženství je již dlouho v rakouské politice považováno za relevantní otázku, a proto v rakouské společnosti existují dvě relevantní dimenze, levo-pravá a nábožensko-sekulární (viz kapitolu 11 v Sartoriho knize [2005]), přestože se tato situace v posledních několika letech mění. Počet dimenzí odráží složitost měření a měl by se projevit v designu měřicího nástroje.

Vedle problému dimenzionality kontinua naznačují jiné studie, že jednotlivci rozumějí pod pojmy „levicový“ a „pravcový“ různé obsahy a jejich odpovědi odrážejí rozdílné podmnožiny jejich celkové politické orientace [Vinopal 2006: 134–139]. Studie levo-pravé orientace jednotlivce naznačuje,

že na místní úrovni je v ČR pouze jediná relevantní dimenze,<sup>8</sup> a to ekonomický populismus – tržní liberalismus [Novák, Vlachová 2001; srov. Vinopal 2006]. K měření pozice jednotlivce na levo-pravém kontinuu existuje řada metod. Snad nejčastěji je používána škála se dvěma krajními body (krajní levice – krajní pravice). Tato metoda předpokládá, že respondenti jsou schopni představit si tuto škálu a jsou schopni se na ní umístit. Jiná metoda konstruuje celkovou škálu odrážející pozici jednotlivce na kontinuu prostřednictvím baterie otázek (položek) [Vinopal 2003: 18]. Dotazníky z šetření, které jsme zvolili, používají ke změření pozice jednotlivce na levo-pravém kontinuu desetibodovou škálu.<sup>9</sup> Obsahují rovněž baterii čtyř binárních položek,<sup>10</sup> které mohou být použity pro odhad mezipoložkové reliability. Můžeme proto otestovat oba typy reliability a prozkoumat obě metody měření pozice na levo-pravém kontinuu.

### 3. Příklady odhadu reliability test-retest

Nejprve zkontrolujeme *předpoklady paralelních měření*. Jak vidíme v tabulce 1, předpoklady stejných očekávaných hodnot, rozptylů (použili jsme směrodatné odchylky, abychom zachovali řád měření pro srovnání se změnou očekávané hodnoty) a korelací se třetí proměnnou (zde věkem) nebyly naplněny. Očekávaná hodnota, rozptyl (směrodatné odchylky) a korelace se významně odchylojí od předpokladu paralelního měření. Korelace mezi těmito čtyřmi měřeními nabývá nižších hodnot, než bychom očekávali, zejména u měření prvního.<sup>11</sup> Nekonzistentní jsou především první dvě měření. Uzavřeli bychom, že předpoklad, že skutečná hodnota pozice na levo-pravém kontinuu zůstane u všech měření konstantní, není naplněn. Na druhou

8 Údaje se vztahují k roku 2000, kdy byla studie provedena. Tato studia použila stejná data jako tato stať, a tak můžeme předpokládat, že se i naši respondenti umísťují na jednodimenzionálním levo-pravém kontinuu.

9 Na této škále znamená 1 krajní levici a 10 krajní pravici. Zkratku „levo-pravý“ (LP) použijeme pro desetibodovou levo-pravou škálu použitou v prvním panelu (LP1), levo-pravou míru použitou ve druhém panelu (LP2), levo-pravou míru použitou ve třetím panelu (LP3) a levo-pravou míru použitou ve čtvrtém panelu (LP4).

10 Znění otázek (položek):

- 1) *Vláda se snaží dělat příliš mnoho věcí, které by lépe vyřešili jednotlivci nebo soukromé podniky. Ne–ano – pořadí odpovědí obráceno pro zachování škály (0 „ne“ – 1 „ano“).*
- 2) *Vláda by měla udělat vše nezbytné ke zlepšení životních podmínek všech chudých obyvatel České republiky (0 „ano“ – 1 „ne“).*
- 3) *Vláda se musí postarat o to, aby byla lidem poskytnuta zdravotní lékařská péče a úhrada za léky (0 „ano“ – 1 „ne“).*
- 4) *Vláda by měla zmenšit rozdíl mezi příjmy bohatých a chudých, nejspíše zvýšením daní pro bohaté rodiny nebo poskytnutím příjmové pomoci chudým (1 „ano“ – 1 „ne“).*

11 Korelace mezi těmito čtyřmi proměnnými se pohybuje od 0,52 do 0,69 a lineárně roste od prvního do čtvrtého měření.

**Tabulka 1. Předpoklady paralelního měření**

	Očekávané hodnoty	Směrodatné odchyly	Korelace se třetí proměnnou
LP1	<b>6,12</b>	<b>2,02</b>	<b>-0,25</b>
LP2	6,19	<b>1,91</b>	<b>-0,24</b>
LP3	6,15	1,84	-0,33
LP4	<b>6,23</b>	1,8	-0,3

Zdroj: Vlastní výpočet na základě klatovských dat.

Poznámka: Tmavá políčka obsahují významné odchyly od předpokladu paralelního měření. Hodnoty jsou odhadovány z respondentů, kteří se zúčastnili všech čtyř vln panelového šetření (n = 429).

stranu, poslední dvě měření (rozestup mezi třetí a čtvrtou vlnou činil tři měsíce) vykazují vyšší míru konzistentnosti. Změny hodnot mezi nimi pomohou odhadnout reliabilitu těchto měření přesněji než kterákoliv jiná kombinace měření ze čtvrté vlny panelového šetření.

Tabulka 2 ukazuje odhad reliability pomocí *korelačního koeficientu*. Z výsledků analýzy předpokladů paralelního měření vyplývá, že realističtější odhad reliability poskytují pouze korelační koeficienty třetí a čtvrté vlny panelového šetření. Ostatní pravděpodobně reliabilitu chybně interpretují, neboť změna rozptylu může být ovlivněna i změnou rozptylu skutečných hodnot.

*Konfirmační faktorová analýza*<sup>12</sup> (CFA) umožňuje výzkumníkovi testovat hypotézy o modelu. Všechny testy a kritéria ověřující vysvětlující schopnost modelu<sup>13</sup> vedou k závěru, že tento model má pro použitá data dobrou vysvětlující schopnost. Odhadovaný rozptyl chyby měření můžeme uvažovat jako podíl pozorované hodnoty, který není vysvětlen latentní proměnnou (skutečnou hodnotou)  $\text{var}(E)/\text{var}(X)$ . Reliabilitu můžeme spočítat jako  $\text{rel}(X,T) = 1 - \text{var}(E)/\text{var}(X)$ . Nebo můžeme vyjádřit čtverec odhadovaných standardizovaných faktorových zátěží jako podíl rozptylu, který mají pozorované proměnné společný se skutečnou hodnotou. Pak tedy  $\text{rel}(X,T) = \text{var}(T)/\text{var}(X)$ . Odhadované standardizované faktorové zátěže (korelace) a odhadované rozptyly chyby uvádíme v grafu 1 a opět v tabulce 2. Odhadované rozptyly chyby slábnou v čase (jejich autokorelace je rovna 0,74,<sup>14</sup> což naznačuje, že chyby měření jsou funkcí času). To může znamenat, že skutečná hodnota pozice na levo-pravém kontinuu se mění v čase (dalším možným

12 S použitím odhadu maximální pravděpodobnosti.

13 Chí kvadrát test modelu není významný ( $X^2 = 0,44$ , 2 (DF),  $p = 0,8$ ), což ukazuje, že předpovězené hodnoty se neliší od hodnot pozorovaných a že je takovýto model statisticky správný.

14 Autokorelační funkce měří závislost hodnot na čase a nabývá hodnot v rozmezí  $<-1; 1>$ . Můžeme ji interpretovat podobně jako korelační koeficient.

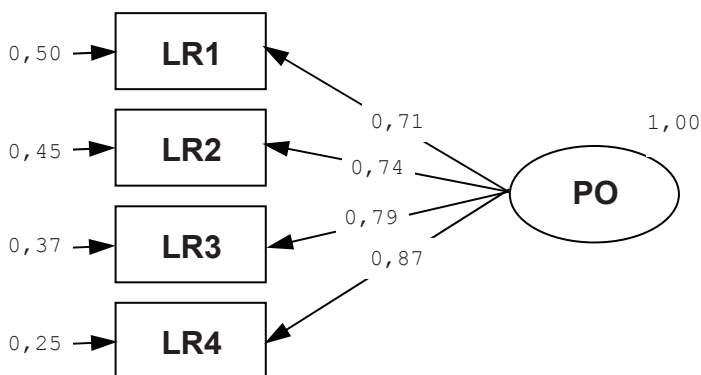


**Tabulka 2. Tři různé odhady reliability test-retest**

	CFA chyby měření	Korelační koeficient	CFA rozptylu skutečné hodnoty
LP1	0,5	0,57	0,53
LP2	0,55	0,59	0,55
LP3	0,63	0,61	0,68
LP4	0,75	0,65	0,69
Průměrná reliabilita	0,61	0,6	0,61

Zdroj: Vlastní výpočet na základě klatovských dat (n = 429).

Poznámka: Zvýrazněna je nejvhodnější metoda odhadu reliability. Hodnoty jsou spočítány na úplných případech.

**Graf 1. Orientovaný graf s odhadovanými standardizovanými faktorovými zátěžemi a rozptyly chyby**

Zdroj: Vlastní výpočet na základě klatovských dat (n = 429).

vysvětlením je „panelový efekt“, kdy se respondenti opakovaním učí zařadit se na levo-pravém kontinuu<sup>15</sup>).

*Odhad rozptylu skutečné hodnoty pomocí CFA spočívá na odlišné logice odhadu reliability, a proto nemusí naplňovat předpoklady paralelního měření. Zdá se, že toto platí pro naše data, a tento odhad by měl být nejrozmumnějším výpočtem reliability pro naše účely. Odhadované rozptyly se liší v závislosti na tom, podle které pozorované proměnné stanovíme škálu pro proměnnou latentní. Vypočetli jsme všechny čtyři odhady rozptylu latentní proměnné a rozptyl skutečné hodnoty jsme odhadli jako jejich průměr. Výsledky jsou znázorněny v tabulce 2.*

Na závěr lze říci, že odhad na základě korelačního koeficientu se nezdá být vhodným odhadem reliability, pokud měření nemohou být považována

15 Za tuto poznámku vděčím jednomu z recenzentů.

za paralelní. Vede k nadhodnocení spodní hranice reliability, která často leží v ohnisku naší pozornosti. Protože většinou nemáme měření, která mohou být považována za paralelní, měli bychom odhadovat variability skutečné hodnoty a ten považovat za reliabilitu. Tato metoda se zdá být realističtější. Pokud neodhadujeme reliabilitu konkrétních měření, nýbrž spíše jejich průměrnou reliabilitu, pak všechny odhady vedou k podobným hodnotám. Dle zvyklosti je za reliabilní hodnotu u reliability test-retest považována hodnota 0,7 [Litwin 1995: 8]. Obecně očekáváme, že se skutečná hodnota do určité míry změní (zejména mezi prvním a dalšími měřeními – nejméně devět měsíců). Uvažujeme-li pouze poslední dvě měření (kde byl rozestup mezi vlnami přibližně tříměsíční), činí průměrná reliabilita 0,69, což se blíží konvenční hranici přijatelnosti. Desetibodovou levo-pravou škálu proto považujeme za reliabilní metodu měření, u které se skutečná hodnota pravděpodobně mění v delším časovém období.

Stručně se zmíníme o důsledcích, které z takového závěru vyplývají. Zdá se, že nejvíce reliabilní je poslední měření. Ovšem i nadále pochází asi 30 % zjištěného rozptylu z chyby měření. Pokud bychom měli použít tuto proměnnou ve statistické analýze, mohli bychom očekávat, že skutečná korelace poklesne o 0,83 (druhá odmocnina odhadu reliability) [Řehák 1998: 55]. Například pozorovaná korelace mezi pozicí na levo-pravé škále a věkem činí -0,27. Očekávali bychom, že korelace mezi skutečnou hodnotou a věkem bude asi -0,33. Jak vidíme, čím nižší je odhad reliability, tím více je korelace ovlivněna chybou a tím nižších hodnot nabývá.

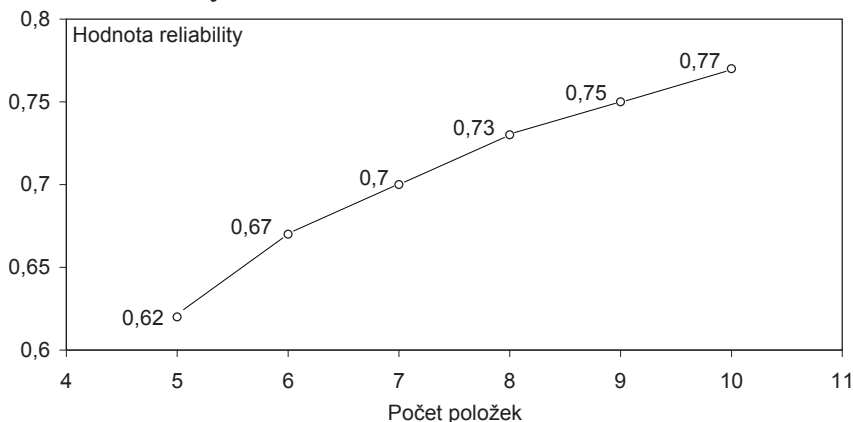
#### 4. Příklady odhadu mezipoložkové reliability

Mezipoložkovou reliabilitu používáme k odhadu reliability baterií. Máme za to, že baterie měří určitý latentní znak (nebo více latentních znaků v případě mnohodimenzionálních baterií). Logika této metody spočívá v tom, že použitím většího počtu otázek k měření jednoho znaku zvýšíme přesnost měření tohoto znaku. Je-li baterie složena z více položek, které mají měřit jeden skrytý znak, tak by jednotlivec měl na tyto položky odpovídat podobně a měla by mezi nimi existovat korelace. K představení druhého typu odhadu reliability použijeme baterii čtyř binárních položek (viz otázky/položky v poznámce 10). K odhadu reliability nemůže být použito Cronbachovo alfa, neboť je baterie zkonstruována z binárních položek. Proto jsme použili *Kuderova-Richardsonova vzorce 20* (KR 20) pro binární položky [Ferjenčík 2000: 203]. Rovněž jsme vypočetli reliabilitu „split-half“ pomocí *Spearmanova-Brownova vzorce* (SB)<sup>16</sup> [Ferjenčík 2000: 201]. Odhadovaná reliabilita podle KR 20 činí 0,57 a podle SB 0,53.<sup>17</sup> Jak vidíme, oba odhady leží pod zvykovou hranicí přijatelnosti 0,7. Problém odhadování reliability baterií spočívá v tom, že její hodnota je silně závislá na počtu položek. Graf 2 uka-

16 K výpočtu korelace mezi oběma částmi škály jsme použili Spearmanův koeficient.

17 U obou výpočtů  $n = 1282$ .

**Graf 2. Účinek přidání položek na vypočtenou hodnotu reliability**



Zdroj: Vlastní výpočet na základě klatovských dat.

zuje účinek přidání položek při konstantní hodnotě reliability (tj. 0,57 pro čtyřpoložkovou baterii).<sup>18</sup> Pokud bychom mohli přidat více otázek, přesáhla by vypočtená hodnota reliability zvykovou hranici 0,7, ačkoliv by skutečná reliability zůstala stejná.

K prozkoumání statistických vlastností baterií můžeme použít *analýzu latentních tříd* [McCutcheon 1987: 13]. Za předpokladu jediného latentního znaku model ukazuje, že takové očekávání není statisticky přiměřené.<sup>19</sup> Proto nemůžeme očekávat, že bude rozptyl položek vysvětlen jedinou latentní proměnnou. Model předpokládající dvě latentní třídy je statisticky významně lepší<sup>20</sup> a má pro tato data dobrou vysvětlující schopnost. V tabulce P2 v příloze vidíme podmíněné pravděpodobnosti jednotlivých položek modelu předpokládajícího dvě latentní třídy. Položka 1 (Vláda se snaží dělat příliš mnoho věcí, které by lépe vyřešili jednotlivci nebo soukromé podniky) má větší „rozptyl“ než všechny ostatní položky a použití jedné třídy nevysvětluje velký podíl tohoto rozptylu. Druhá latentní třída vysvětluje zbyvající rozptyl první položky. Můžeme proto předpokládat, že tato baterie nemá jednu, nýbrž spíše dvě dimenze. Jelikož předpokládáme jednodimenzionální kontinuum (viz oddíl Analyzovaná data a proměnné), navrhovali bychom, aby první položka nebyla do baterie zahrnuta. Vidíme, že pravděpodobnosti levicových odpovědí v třídě 2 jsou u položek 2–4 systematicky vyšší než u položky 1. Na druhé

18 Vzorec lze nalézt na [http://www.gower.k12.il.us/Staff/ASSESS/4\\_ch2app.htm#formula%2020](http://www.gower.k12.il.us/Staff/ASSESS/4_ch2app.htm#formula%2020).

19  $G^2 = 351, 11$  (DF) sig = 0,00001;  $X^2 = 748, 11$  (DF).

20 AIC: 4520; BIC: 4566. Model předpokládající 3 latentní třídy je statisticky horší (AIC: 4527; BIC: 4599).

straně, v třídě 1 je pravděpodobnost levicové odpovědi na otázku o zdravotnictví významně vyšší než u jiných otázek (viz opět tabulku P2 v příloze). Zdá se, že lidé častěji poskytují pravicové odpovědi ohledně hospodářství (i kdyby na všechny ostatní otázky odpověděli levicově). Proto nemůžeme očekávat, že lidé na všechny čtyři otázky odpoví podobně. Pokud bychom tedy použili pouze položky 2–4, činila by reliabilita podle KR 20 pro tyto tři položky 0,55. Pak bychom očekávali, že tato reliabilita (0,55 pro tři položky) vzroste na 0,62, použijeme-li ke konstrukci škály čtyři položky, což je více než původní odhad 0,57. Skutečnou reliabilitu škály jsme proto zvýšili.

Použití položek 1–4 ke konstrukci škály<sup>21</sup> může proto být problematické, neboť nemůžeme očekávat, že měří jediný latentní znak (latentní třídu). Obecně řečeno, před konstrukcí škály a odhadem její reliability bychom měli provést buď analýzu hlavních komponent (jsou-li všechny položky měřeny na ordinální nebo kardinální úrovni), nebo analýzu latentních tříd (jsou-li položky nominální nebo binární).

## **5. Odhad reliability pomocí paralelní metody ve stejném dotazníku**

Poslední metodou, kterou použijeme k odhadu reliability měření, je paralelní forma. Porovnááme hodnoty dvou současných měření na stejných respondentech. Pokud obě míry (otázky či položky) měří stejnou latentní proměnnou (stejný konstrukt nebo stejnou skutečnou hodnotu) různými metodami (různými otázkami nebo bateriemi položek), pak by jejich hodnoty měly být korelované [Ferjenčík 2000: 200]. Předpokladem tohoto modelu je, že obě míry měří stejné dimenze skutečné hodnoty. Nicméně sociální koncepty rozložené na konstrukty jsou příliš složité. Existuje příliš mnoho různých způsobů, jak dimenze chápat, a není realisticky možné nalézt dvě měření, která při použití různých metod měří přesně stejnou skutečnou hodnotu.

Tento typ odhadu můžeme demonstrovat na výpočtu korelace mezi čtyř-položkovou baterií a levo-pravou škálou. Celková škála je zkonstruována jako suma odpovědí na pravicové položky. Maximální hodnota, které může jednotlivec dosáhnout, tedy činí 4, což odpovídá krajní pravici, a hodnota 0 odpovídá krajní levici. Odhadovaná hodnota reliability činí 0,28. Tento odhad by pravděpodobně nebyl považován za dobrý odhad reliability. Myslíme si, že ukazuje nízkou reliabilitu při existenci více dimenzí politické orientace (nebo obecně jakéhokoliv konceptu), které dané míry nezachycují. Jelikož jsme u této baterie odhadovali přítomnost dvou dimenzí, očekáváme, že výsledek bude přiměřený. Proto se zdá, že tento typ reliability není pro většinu sociologického výzkumu, a zejména dotazníkových výběrových šetření, vhodný.

---

21 Pokud bychom konstruovali jednodimenzionální škálu (např. součtem všech kladných odpovědí).

## 6. Odhad hodnoty chyby a standardních chyb

Jakmile máme k dispozici odhad reliability, jsme schopni předpovědět skutečnou hodnotu. Lord a Novick [1968] popsali lineární regresní model pro odhad skutečné hodnoty z měřené hodnoty a z odhadu reliability [Lord, Novick 1968: 64–69]. Díky lineárnímu vztahu mezi X, T a E můžeme v lineárním regresním modelu odhadovat parametry funkce pomocí metody nejmenších čtverců. Regresní funkce má podobu [vzorec 3.2.7a v Lord, Novick 1968: 65]:

$$R(T|x) = \text{rel}_{(X,T)} * x + (1 - \text{rel}_{(X,T)}) * \mu_x \quad (6).$$

Skutečnou hodnotu odhadneme jako součet dvou odhadů. Prvním je vážený součet individuálních měřených hodnot ( $\text{rel}_{(X,T)} * x$ ), druhým vážený součet průměrů skupiny [ $(1 - \text{rel}_{(X,T)}) * \mu_x$ ]. Je-li měření reliabilní, pak dostane individuální hodnota větší váhu a průměr skupiny menší váhu – a naopak [Lord, Novick 1968: 65]. Je-li reliability nízká, pak se odhadovaná skutečná hodnota bude shlukovat okolo průměru pozorovaných hodnot ve skupině. Znázornění poskytuje graf P1 v příloze.

Matematicky by očekávaná hodnota skutečné ( $E(T)$ ) i měřené hodnoty ( $E(X)$ ) měla být shodná, pokud je skutečná hodnota funkcí hodnoty pozorované. To teoreticky odráží definici „limity skutečné hodnoty“, podle které průměrná hodnota velkého počtu měření konverguje k hodnotě skutečné [Lord, Novick 1968: 28–29]. Rovněž můžeme pozorovat, že rozptyl skutečné hodnoty je nižší než rozptyl hodnoty pozorované, pokud  $\text{var}(X) = \text{var}(T) + \text{var}(E)$ . Chceme-li odvozovat odpovědi ze vzorku celé populace, obvykle používáme statistický model založený na předpokladu normálního rozložení nebo jiné známé třídy rozložení. Klasická teorie testů předpokládá, že chceme-li odhadovat skutečnou hodnotu, je třeba tak činit s ohledem na reliability. Vzorec intervalu spolehlivosti (IS) má podobu [vzorec 3.8.3 v Lord, Novick 1968: 67]:

$$IS = \mu_x \pm t_{1-\alpha/2} / \sqrt{n} * \sigma_x * \sqrt{(1 - \text{rel}_{(X,T)})} \quad (7),$$

kde  $\sigma_x$  představuje směrodatnou odchylku a  $\sigma_x * \sqrt{(1 - \text{rel}_{(X,T)})}$  *standardní chybu měření*  $\sigma_e$ . Čím menší reliability, tím širší je interval, a naopak. Již jsme uvedli jeden způsob odhadu skutečné hodnoty pomocí regresního modelu. Máme-li pro tento odhad odhadnout interval spolehlivosti, vypočteme *standardní chybu odhadu*

$$\sigma_e = \sigma_x * \sqrt{(1 - \text{rel}_{(X,T)})} * \sqrt{\text{rel}_{(X,T)}} \quad (8)$$

a z ní vypočteme interval spolehlivosti [vzorec 3.8.4a v Lord, Novick 1968: 68]. Oba druhy chyb vedou v porovnání s klasickým modelem (který používá pouze standardní chybu průměru) k menšímu nebo rovnému intervalu spolehlivosti, neboť  $\text{var}(T) \leq \text{var}(X)$ . Standardní chyba odhadu bude vždy (kromě případů, kdy  $\text{rel}(X,T) = 1$  nebo  $\text{rel}(X,T) = 0$ ) menší než standardní chyba měření. Je tomu tak proto, že k odhadu skutečné hodnoty pomocí regresního modelu použijeme doplňkový údaj (průměr skupiny), který při

**Tabulka 3. Standardní chyba měření a odhadu**

	LP1	LP2	LP3	LP4
Standardní chyba měření	1,38	1,32	1	0,99
Standardní chyba odhadu	1,01	0,98	0,83	0,82

Zdroj: Vlastní výpočet na základě klatovských dat.

**Tabulka 4. Intervaly spolehlivosti 95 % standardní chyby měření a odhadu**

Věkové skupiny	Dolní hranice IS	Průměrná měřená hodnota	Horní hranice IS	Dolní hranice IS	Průměrná odhadnutá hodnota	Horní hranice IS
(0, 30]	6,87	6,94	7	6,67	6,72	6,77
(30, 50]	6,31	6,42	6,53	6,28	6,36	6,44
(50, 70]	5,63	5,72	5,81	5,81	5,88	5,96
(70+]	5,91	6	6,09	6	6,07	6,15

Zdroj: Vlastní výpočet na základě klatovských dat.

odhadu skutečné hodnoty pomocí standardní chyby měření nepoužijeme. Porovnání obou standardních chyb poskytuje tabulka 3.

Tuto vlastnost pozorujeme, když srovnáme průměrné měřené a odhadnuté hodnoty na levo-pravé škále u daných čtyř věkových skupin.<sup>22</sup> Průměrná velikost intervalu spolehlivosti (IS) standardní chyby měření činí 0,18, zatímco průměrná velikost IS standardní chyby odhadu činí 0,14. Odhad skutečné hodnoty pomocí regresního modelu je tedy přesnější. Další závěr zní, že čím reliabilnější měření, tím méně jednotek je zapotřebí pro stejnou šířku IS. Tato vlastnost může být obzvlášť užitečná s ohledem na náklady, které přináší dotazování každého dalšího respondenta. Tabulka 4 uvádí průměry a mezní hodnoty IS pro obě standardní chyby u jednotlivých věkových skupin.

## 7. Závěr

Reliabilitu měření můžeme považovat za míru, v jaké se rozptyl skutečné hodnoty shoduje s rozptylem měřené hodnoty. Měření je absolutně reliabilní, pokud se měřená hodnota rovná hodnotě skutečné, a naopak. Problém výpočtu hodnoty reliability spočívá v tom, že nemůžeme změřit skutečnou hodnotu a odpovídající hodnotu chyby. Hodnotu reliability můžeme pouze odhadovat na základě určitých předpokladů v rámci teoretického modelu.

Existují dva typy reliability: test-retest a mezipoložková. V této stati jsme použili příklad pozice jednotlivce na levo-pravém kontinuu k odhadu reliability určité metody měření (desetibodové levo-pravé škály) a k odhadu

<sup>22</sup> Byly sledovány čtyři věkové skupiny (18–30, 31–50, 51–70 a více než 70). Použili jsme levo-pravou škálu z poslední vlny s odhadem reliability  $rel = 0,69$  ( $n = 472$ ).

reliability baterie čtyř binárních položek. Předpoklad paralelního měření nebyl naplněn a očekávali jsme, že skutečná hodnota pozice na levo-pravém kontinuu se bude do určité míry měnit v delším časovém období.

Vezmeme-li měření ze dvou vln panelu s nejkratšími intervaly a nejvyššími mírami konzistence, pak jejich průměrná reliabilita činí 0,69, což je blízko konvenční hodnotě reliability 0,7. Pozorovaný rozptyl těchto dvou měření je nejspíše zapříčiněn rozptylem hodnoty chyby, a nikoli rozptylem skutečné hodnoty. Tento odhad rozptylu je proto méně zkreslený a vhodnější.

Odhadování reliability baterie závisí na počtu zkoumaných položek. Zvyková hodnota 0,7 je proto problematická a měla by být posuzována s ohledem na počet položek baterie. Ukázalo se, že naše baterie je reliabilní, ačkoliv vypočtené hodnoty reliability činily pouze 0,57. Přidáme-li další tři položky, zvýší se odhadovaná reliabilita škály na 0,7, což je pro sedmipoložkovou baterii uspokojivější výsledek [Soukup 2006].

Zkontrolovali jsme předpoklady pomocí LCA a ukázalo se, že naše baterie má více než jednu dimenzi. Proto bychom neměli očekávat, že respondenti budou na položky odpovídat podobně. Odstranili jsme problematickou položku, která způsobovala mnohodimenzionálnost, a odhadovaná reliabilita se při zachování stejného počtu položek zvýšila. Nakonec jsme odhadli skutečnou hodnotu pomocí lineárního regresního modelu, zahrnujícího odhadovanou reliabilitu. Výsledky ukázaly, že reliabilita má pozitivní účinek na přesnost prováděného usuzování. Průměrná velikost intervalu spolehlivosti při použití údaje o odhadované reliabilitě byla asi o 20 % menší než velikost intervalu spolehlivosti vypočtená bez tohoto údaje.

Měli bychom mít na paměti, že vedle klasické testové teorie, představené v této stati, existují i další odhady reliability, vhodné pro různé účely. Výzkumník by proto před výpočtem měl zvažovat situaci a kontext, ve kterém se nachází. Výzkumník by měl být motivován nejen zpřesněním výsledků, ale také snížením nákladů na výzkum. Pro metodologii výběrových šetření proto zůstává zlepšení měření důležitým cílem.

## Literatura

- Allen, M.J. and W. M. Yen. 2002. *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.
- Brown, W. 1910. „Some experimental results in the correlation of mental abilities“. *British Journal of Psychology* 3: 296-322.
- Ferjenčík, J. 2000. *Úvod do metodologie psychologického výzkumu*. Praha: Portál.
- Jeřábek, H. (ed.). 1999. *Utváření postojů obyvatel českého města*. Pracovní texty / Working papers 99:7. Praha: Sociologický ústav AV ČR.
- Jeřábek, H., P. Soukup (eds.). 2008. *Advanced Lazarusfeldian Methodology*. Praha: Karolinum.
- Kreidl, M. 2004. „Porovnání práce výzkumných agentur na základě srovnání validity standardizovaných měřících instrumentů“. Pp. 87-96 in: *Kvalita výzkumů volebních preferencí*, J. Krejčí (ed.). Praha: Sociologický ústav AV ČR.

- Kreidl, M. 2005. „Zhodnocení vlivu práce výzkumných agentur na konstruktovou validitu škál“. *Sociologický časopis / Czech Sociological Review* 41 (2): 103-123.
- Lazarsfeld, P. F., R. K. Merton (eds.). 1950. *Continuities in Social Research: Studies in the Scope and Method of the American Soldier*. Glencoe: The Free Press.
- Litwin, M. S. 1995. *How to Measure Survey Reliability and Validity*. Thousand Oakes, CA, London, New Delhi: Sage.
- Lord, F. M., M. R. Novick. 1968. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison Westley.
- Matějů, P., K. Vlachová et al. 2000. *Nerovnost, spravedlnost, politika. Česká republika 1991–1998*. Praha: Sociologické nakladatelství (SLON).
- McCutcheon, A. L. 1987. *Latent Class Analysis*. Newbury Park, CA, London, New Delhi: Sage.
- Novák, M., K. Vlachová. 2001. *Linie štěpení v České republice. Komparace národní úrovně s příkladem konkrétní lokality*. Sociologické texty / Sociological Papers 01:8. Praha: Sociologický ústav AV ČR.
- Řehák, J. 1968. „Definice měření v sociologii.“ *Sociologický časopis* 4 (6): 638–647.
- Řehák, J. 1998. „Kvalita dat I. Klasický model měření reliability a jeho praktický aplikační význam.“ *Sociologický časopis* 34 (1): 51–60.
- Sartori, G. 2005. *Strany a stranické systémy. Schéma pro analýzu*. Brno: CDK.
- Soukup, P. 2006. „Čím větší, tím lepší (aneb mýty o reliability).“ *Socioweb* [online] (7) [cit. 4. 1. 2011]: 11–12. Dostupné z: <http://www.socioweb.cz/index.php?disp=teorie&shw=242&lst=112>.
- Spearman, Charles, C. 1910. „Correlation calculated from faulty data“. *British Journal of Psychology* 3: 271-295.
- Šubrt, J. 2006. „Veřejné mínění z pohledu teorie a praxe.“ Pp. 13–24 in D. Kunštát (ed.). *České veřejné mínění: výzkum a teoretické souvislosti*. Praha: Sociologický ústav AV ČR, v.v.i.
- Vinopal, J. 2003. „Kdo je ‚levičák‘ a kdo ‚pravičák‘?“ *Naše společnost* 1 (1–2): 18–21.
- Vinopal, J. 2006. „Empirická přístupnost levo-pravé politické orientace.“ *Sociologický časopis / Czech Sociological Review* 42 (1): 129–147.



## PŘÍLOHA

**Tabulka P1. Panelové šetření krystalizace názorů a postojů ve městě Klatovy**

Vlna	Sběr dat	Odstup mezi vlnami	Velikost výběru
První	Září 1999		2037
Druhá	Květen 2000	9 měsíců	672
Třetí	Září 2000	5 měsíců	562
Čtvrtá	Listopad 2000	3 měsíce	521

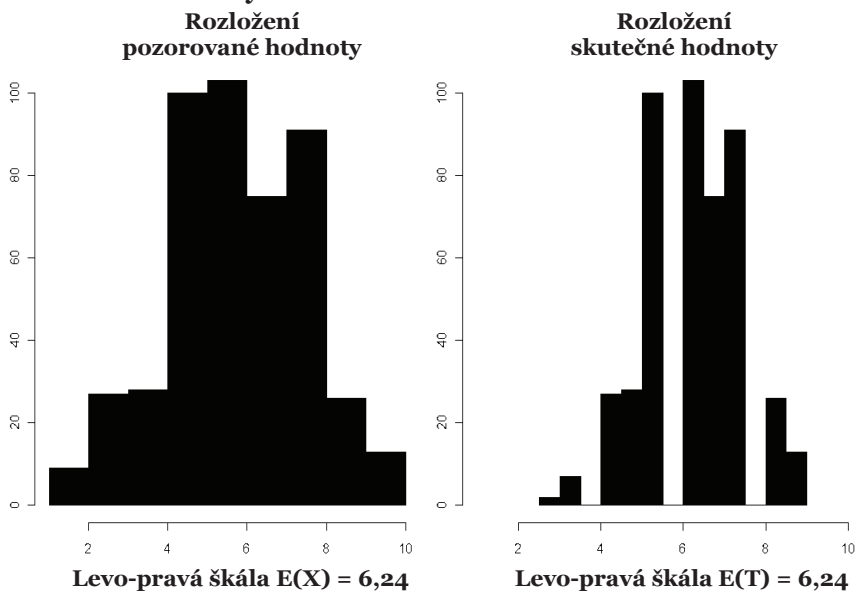
Poznámka: Hypotézu o zkreslení v souvislosti s chybějícími odpověďmi v dalších vlnách (2 až 4) jsme testovali pomocí chí kvadrát testu vzdělanostních a věkových skupin u osob, které odpověděly na studované otázky. Žádný test nebyl významný (významnost se pohybovala od 0,42 od 0,95). Rozdělení věku se liší mírně více než rozdělení vzdělání, což ukazuje, že rozdělení odpovědí různých věkových a vzdělanostních skupin bylo ve všech čtyřech vlnách prakticky stejné.

**Tabulka P2. Podmíněné pravděpodobnosti jednotlivých položek v modelu dvou latentních tříd**

<i>Položka 1: Vláda se snaží dělat příliš mnoho věcí, které by lépe vyřešili jednotlivci nebo soukromé podniky.</i>		
	Pravděpodobnost pravicové odpovědi	Pravděpodobnost levicové odpovědi
Třída 1	0,68	0,32
Třída 2	0,24	0,76
<i>Položka 2: Vláda by měla udělat vše nezbytné ke zlepšení životních podmínek všech chudých obyvatel České republiky.</i>		
	Pravděpodobnost pravicové odpovědi	Pravděpodobnost levicové odpovědi
Třída 1	0,57	0,43
Třída 2	0,09	0,91
<i>Položka 3: Vláda se musí postarat o to, aby byla lidem poskytnuta zdravotní lékařská péče a úhrada za léky.</i>		
	Pravděpodobnost pravicové odpovědi	Pravděpodobnost levicové odpovědi
Třída 1	0,29	0,71
Třída 2	0,01	0,99
<i>Položka 4: Vláda by měla zmenšit rozdíl mezi příjmy bohatých a chudých, nejspíše zvýšením daní pro bohaté rodiny nebo poskytnutím příjmové pomoci chudým.</i>		
	Pravděpodobnost pravicové odpovědi	Pravděpodobnost levicové odpovědi
Třída 1	0,7	0,3
Třída 2	0,06	0,94

Zdroj: Vlastní výpočet na základě klatovských dat (n = 1282).

## Graf P1. Rozdělení pozorované a odhadované skutečné hodnoty



Zdroj: Vlastní výpočet na základě klatovských dat (n = 429).

Poznámka: Použili jsme levo-pravou škálu z předchozího kola s odhadem reliability rel = 0,69 (n = 472).